

Statistical Modelling and Analysis of Time-to-Event Data under Bayesian Paradigm

काशी हिन्दू
विश्वविद्यालय



BANARAS HINDU
UNIVERSITY

THESIS SUBMITTED FOR THE DEGREE OF
Doctor of Philosophy
in
Statistics

By
Vikas Barnwal

Under the Supervision of
Dr. Mahaveer Singh Panwar

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

Enrolment No.: 411186

December 2023

Dedicated to
“My Mother and Elder Sister”

Copyright © Institute of Science, 2023
Banaras Hindu University
Varanasi – 221005, India

All rights reserved

UNDERTAKING

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. ordinance of the Banaras Hindu University, Varanasi, India and that the research work embodied in this thesis entitled "**Statistical Modelling and Analysis of Time-to-Event Data under Bayesian Paradigm**" 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

(**Vikas Barnwal**)

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

CANDIDATE'S DECLARATION

I, **Vikas Barnwal**, s/o Shri Pancham Barnwal, 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**, Department of Statistics, Banaras Hindu University for a period of September 2018 to December 2023 at the Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi. 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 devotedly 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
(**Vikas Barnwal**)

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

COURSE WORK COMPLETION CERTIFICATE

This is to certify that **Mr. Vikas Barnwal**, a bonafide research scholar of this department, has successfully completed the Ph.D. course work, which is a part of his Ph.D. programme.

Date:
Place: Varanasi

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

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

PRE-SUBMISSION SEMINAR COMPLETION CERTIFICATE

This is to certify that **Mr. Vikas Barnwal**, a bonafide research scholar of this department, has successfully completed the pre-submission seminar requirement on topic "**Statistical Modelling and Analysis of Time-to-Event Data under Bayesian Paradigm**" dated December 01, 2023, which is a part of his Ph.D. programme, .

Date:
Place: Varanasi

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

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

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: **Statistical Modelling and Analysis of Time-to-Event
Data under Bayesian Paradigm**

Candidate's Name: **Vikas Barnwal**

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 utmost respect and gratitude to the revered founder of our esteemed university, **Pt. Madan Mohan Malviya Ji**. It is an immense privilege for me to have such opportunity to pursue research in Statistics at Banaras Hindu University.*

*I am deeply indebted to my supervisor, **Dr. Mahaveer Singh Panwar**, for his unwavering guidance, invaluable insights, and continuous support throughout my research journey. His caring nature and amicable attitude have been a constant source of motivation and encouragement for me, giving me the courage to overcome the hurdles encountered throughout this research journey. I will be forever grateful to him for the invaluable impact he has had on my life and academic journey.*

*I express my heartfelt gratitude to **Prof. S. K. Singh**, Head, Department of Statistics and **Prof. B. B. Khare**, the former Head of the Department of Statistics, Institute of Science, Banaras Hindu University. Their keen interest and unwavering support, along with the provision of necessary departmental facilities, have been instrumental in the successful completion of my research work. I would also like to extend my special thanks to my Research Programme Committee, comprising **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. Their kind support and constructive suggestions throughout the research process have significantly contributed to the refinement and quality of my work. I wish to acknowledge my profound gratitude to **Dr. Nirpeksh Kumar**, Associate Professor, Department of Statistics, Institute of Science, Banaras Hindu University for his unwavering belief in my abilities and constant encouragement throughout my research journey.*

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 to **Dr. Rashmi Bundel**, Assistant Professor, Rajasthan University and **Dr. Himanshu Rai**, Guest Faculty, TISS, Mumbai for their discussion of core concepts of Survival analysis and Bayesian inference. I feel pleasure in recording my special thanks to **Dr. Tribhuvan Singh**, Assistant professor, DIT University, Dehradun for his brotherly support whenever I seek help from him during my research work. I would like to extend my thanks to my seniors **Dr. Mridulesh Kumar Yadav** and **Mrs. Sonam Jaiswal** for*

*their support and motivation at the very beginning and later at various stages of the research journey. I extend my sincere thanks to **Dr. Chandra Prakash Yadav**, Research Fellow, Saw Swee Hock School of Public Health, NUS, Singapore for his invaluable help and insightful discussions during the formulation of the problem for my research. I would also like to thank **Annapurna, Suraj Yadav** and **Vikas Baranwal** Research Scholars, Department of Statistics, Banaras Hindu University for their support and encouragement during my entire research journey. Moreover, I am thankful for my beloved juniors **Mr. Sanjiv Singh**, **Mr. Avinash Shukla**, **Mr. Ankush Sharma**, Research Scholars, Department of Statistics, Banaras Hindu University and **Mr. Sushil Yadav** for their multilevel support during the last stages of my research journey in Department of Statistics. Along with it, I am also thankful to all other research scholars of Department of Statistics, Banaras Hindu University, for their wholehearted support and kind cooperation throughout the research period.*

*I would like to express my deepest gratitude to my parents, **Smt. Asha Devi & Ms. Priti Baranwal** for their unconditional love, unwavering support, and countless sacrifices they have made to ensure my success in every step of my life. Their constant encouragement, guidance, and belief in me have been the driving force behind my accomplishments. I would also like to extend my heartfelt thanks to my loving sister **Ms. Priya Baranwal** and brother **Mr. Manish Baranwal**, for pushing me up, motivating and cheering whenever I felt down.*

*Last but certainly not least, I would like to extend my heartfelt gratitude to my close friends **Dr. Naveen Gupta**, **Ms. Yashasvi Jaiswal**, **Mr. Rahul Dev Gupta** and **Mr. Mohd. Hasnain**. Their unconditional support, encouragement, and companionship throughout this academic endeavor have been invaluable to me. It is with the collective support of these individuals that I have been able to embark on and complete this thesis successfully.*

December 2023

Vikas Baranwal

Preface

“Uncertainty is the only certain thing in nature.”

Time-to-event studies are always concerned with the duration of the study and/or the resources required for its implementation. Consequently, the time to event of interest for all individuals is not always observed and this incompleteness related to the time variable is referred to as censoring. There exists various censoring schemes such as type-I censoring, type-II censoring, random censoring, interval censoring and progressive hybrid censoring which commonly arise in different fields. These censoring schemes represent different ways in which time-to-event data is incompletely observed. Among them, the extreme form of interval censoring is known as current status data, where information is available only regarding the occurrence of the event of interest. The recall-based data is one step ahead of the current status data, as it involves both the status of the occurrence of the event and the time-to-event recorded for the individuals who are able to recall it.

Competing risks survival data refers to the observed data that includes the time-to-event along with its cause of failure. In such type of data, individuals are exposed to multiple risks, resulting in the occurrence of the event attributed to one of the possible risks. Competing risks data are often analysed with the independence assumption among the risks. However, incorporating the potential dependency among the causes under consideration can lead to more comprehensive and accurate analyses of these data. If the cause of failure is not known for all the individuals who have experienced the event, this incomplete observation regarding the true cause of failure is referred to as masking. The masked competing risks data poses several problems for the estimation process and needs careful consideration of the masked competing risks data to ensure accurate and efficient estimation.

Longitudinal data consists of repeated measurements of a response variable over a period of time from the same individuals in a study. The trend of the response variable over time and any changes in its pattern may have an impact on the different risks in competing risks survival data. Therefore, the joint analysis of longitudinal data and competing risks data may produce more accurate and efficient results in the estimation process. This thesis

is dedicated to analysing the time-to-event data, with a specific focus on incorporating the dependency among the potential risks for competing risks data and lastly the joint modelling of longitudinal and competing risks data with dependent masking, particularly within the Bayesian paradigm.

In Chapter 2 of this thesis, a parametric model for recall-based data is developed and the analysis is performed using the objective Bayesian approach. The non-recall probabilities are chosen under the phenomena that the chance of recalling the time to the event decreases as the duration between the occurrence of the event and the monitoring point increases. Under the objective Bayesian approach, the reference priors for different groups of parameters are derived. These groups of parameters are ordered based on the inferential importance of these parameters. The Bayesian estimates are obtained under squared error and linear exponential loss functions to account for both overestimation and underestimation in the estimation process and an extensive simulation study is performed for illustration purposes. The duration of breastfeeding for the recent child is estimated using the proposed methodology under the Bayesian paradigm using data from National Family Health Survey (NFHS) IV data.

In Chapter 3, a novel bivariate distribution called the Marshall-Olkin bivariate generalized lifetime distribution (MOBGLD) is presented to account for dependency in competing risks data. The dependency arises when there is a non-zero probability of simultaneous occurrence of two causes of failure. The derived MOBGLD distribution consists of five bivariate distributions having different types of hazards. The unknown parameters of the distribution are ordered into different sets according to their inferential importance and corresponding to each set, a reference prior is derived. The estimation process is carried out for the obtained reference priors and to illustrate the effectiveness of our approach, we apply it to two real datasets: prostate cancer data and diabetic retinopathy data. Furthermore, a comprehensive simulation study is conducted to explore the nature of hazards of the proposed MOBGLD for dependent competing risks data.

In Chapter 4, Marshall-Olkin bivariate Weibull (MOBW) distribution is considered to address the issue of dependency in competing risks data analysis. The reference priors are

derived for the ordered group of parameters when there is prior information for part of the parameters of the model under study. The model compatibility is checked using posterior predictive p -value for diabetic retinopathy data. A complete simulation study is performed for showing the practical implementation of the proposed methodology under the Bayesian paradigm. For illustration purposes, we utilized the same datasets used in the previous chapter.

In Chapter 5, the survival model is extended to account for the longitudinal data outcome into the competing risks data outcome analysis. A flexible joint model is derived for both outcomes incorporating cause-dependent masking. The association structure in joint modelling is represented as shared random effects. The estimation of unknown parameters in the joint model is done under the Bayesian paradigm. An extensive simulation study is conducted to depict the performance of the joint model. For the practical implementation of the proposed model under the Bayesian paradigm, prostate cancer data from Prostate, Lung, Colorectal, and Ovarian Cancer (PLCO) Screening trial is utilized.

The papers based on Chapter 3 and Chapter 4 in the thesis have been published in the esteemed journal *Communications in Statistics - Theory and Methods* in 2022 and 2023, respectively. It is expected that the findings and outcomes presented in the thesis will be helpful for researchers who have an interest or currently engaged in the field of time-to-event survival data analysis under the Bayesian framework, particularly in the context of competing risks with cause dependency.

Contents

1. Introduction	1
1.1 Time-to-Event Study	2
1.1.1 Censoring	2
1.1.2 Recall-Based Time-to-Event Data	3
1.1.3 Competing Risks Masked Data	4
1.1.4 Cause-Specific Hazard Model	5
1.2 Longitudinal Study	6
1.2.1 Repeated Measures Analysis of Variance	7
1.2.2 Generalized Estimating Equations	8
1.2.3 Linear Mixed-Effect Modelling	8
1.3 Estimation Methods	9
1.3.1 Bayesian Estimation	9
1.4 Advance Computational Techniques	13
1.4.1 Markov Chain Monte Carlo Method	13
1.4.2 Metropolis-Hastings Algorithm	14
1.4.3 Metropolis Adjusted Langevin Algorithm	16
1.4.4 Gibbs Sampling Algorithm	17
1.4.5 Highest Posterior Density Interval	17
1.4.6 Model Compatibility	18
1.4.7 Posterior Predictive p -Value	19
2. Objective Bayesian Approach for Recall-Based Time-to-Event Data	20
2.1 Introduction	20
2.2 Likelihood Construction for Recall-Based Data	24
2.3 Objective Bayesian Inference	29
2.3.1 Reference Priors	29
2.3.2 Posterior Propriety and Estimation	35
2.4 Simulation Study	38
2.5 Real Data Application	43
2.5.1 Model Compatibility based on PPV	43
2.5.2 Breastfeeding Data	44

2.6 Conclusion	46
3. Objective Bayesian Analysis of Cause-Dependent Competing Risks Data using Marshall-Olkin Bivariate Generalized Lifetime Distribution	48
3.1 Introduction	48
3.2 Model Building for Dependent Competing Risks	51
3.2.1 Generalized Lifetime Distribution	51
3.2.2 Marshall-Olkin Bivariate Generalized Lifetime Distribution	52
3.2.3 Ageing Properties for MOBGLD	56
3.2.4 Dependent Competing Risks for MOBGLD	56
3.3 Inference for Competing Risks Model	57
3.3.1 Fisher Information Matrix	57
3.3.2 Estimation under Re-parametrization	61
3.4 Bayesian Inference for Competing Risks Model	63
3.4.1 Reference Prior for MOBGLD	64
3.4.2 Posterior Estimation	71
3.5 Simulation Study	74
3.6 Real Data Analysis	75
3.6.1 Prostate Cancer Data	76
3.6.2 Diabetic Retinopathy Study Data	78
3.7 Conclusion	81
4. Objective Bayesian Analysis of Cause-Dependent Competing Risks Data with Partial Information	83
4.1 Introduction	83
4.2 Preliminaries	86
4.3 Likelihood Construction for Dependent Competing Risks Model	87
4.4 Bayesian Estimation	88
4.4.1 Reference Prior with Partial Information	89
4.4.2 Posterior Propriety	92
4.4.3 Posterior Estimation	94
4.4.4 Testing under Bayesian Paradigm	95
4.4.5 Posterior Predictive Compatibility of Model	99
4.5 Simulation Study	100
4.6 Real Data Application	103
4.6.1 Prostate Cancer Data	103
4.6.2 Diabetic Retinopathy Study Data	105
4.7 Conclusion	108

5. Bayesian Joint Modelling of Longitudinal and Competing Risks Data with Cause-Dependent Masking	109
5.1 Introduction	109
5.2 Motivation of the Work	112
5.3 Joint Modelling	113
5.3.1 Sub-model for Longitudinal Data	114
5.3.2 Sub-model for Masked Competing Risks Data	116
5.3.3 Joint Inference of Longitudinal and Competing Risks Data	120
5.4 Bayesian Estimation	121
5.4.1 Prior Elicitation	122
5.4.2 Posterior Inference	123
5.5 Simulation Study	127
5.6 Conclusion	133
References	135

List of Tables

2.1	All reference prior for the parameter Θ	34
2.2	Bayes estimates of parameters along with their AB and MSE for uniform monitoring.	40
2.3	Bayes estimates of parameters along with their AB and MSE for exponential monitoring.	41
2.4	AL and CP of 95% HPD interval estimates for uniform monitoring.	42
2.5	AL and CP of 95% HPD interval estimates for exponential monitoring.	42
2.6	Bayes estimates of parameters of Weibull model for breastfeeding data of Arunachal Pradesh under SELF.	46
3.1	Distributions for specific forms of $g(x)$ and parameters for GLD	51
3.2	Ageing classification of $MOBGLD$ for specific distributions.	57
3.3	Reference priors for parameters, Λ , of $MOBGLD(\Lambda)$ and for parameters, Θ , of $MOBGLD(\Theta)$	68
3.4	Reference priors for parameters, (β, Λ) , of $MOBGLD(\beta, \Lambda)$ and for parameters, (β, Θ) , of $MOBGLD(\beta, \Theta)$	69
3.5	All reference priors for the parameter Λ when β is known and for (β, Λ) when β is unknown with corresponding triplets (l_1, l_2, l_3)	70
3.6	AB , $RMSE$ and CP of the Bayes estimators for $MOBED(0.87, 1.30, 1.50)$. . .	75
3.7	AB , $RMSE$ and CP of the Bayes estimators for $MOBRD(0.33, 0.66, 0.76)$. . .	75
3.8	AB , $RMSE$ and CP of the Bayes estimators for $MOBPD(1.15, 1.65, 1.71)$ when $a = 0.5$	76
3.9	AB , $RMSE$ and CP of the Bayes estimators for $MOBBD(1.25, 1.64, 1.72)$ when $b = 1.5$	76
3.10	AB , $RMSE$ and CP of the Bayes estimators for $MOBWD(0.75, 0.55, 1.25, 1.35)$. .	77
3.11	Patients summary statistics of prostate cancer data.	77
3.12	Bayes estimates (95% HPD Interval) of parameters for $MOBWD$ of prostate cancer data.	79
3.13	Time to blindness in days and causes for 71 white male diabetic patients. . .	79
3.14	Bayes estimates (95% HPD interval) of parameters for $MOBWD$ of Diabetic Retinopathy study data.	80

4.1	ARB and MSE of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$ under $SELF$	101
4.2	AL of 95% HPD interval and coverage probability of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$	101
4.3	ARB and MSE of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$ under $SELF$	102
4.4	AL of 95% HPD interval and coverage probability of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$	102
4.5	Bayes estimates of parameters of $MOBWD$ for prostate cancer data under $SELF$	105
4.6	Bayes estimates of parameters of $MOBWD$ for DRS data under $SELF$	107
5.1	Bayes estimates and their bias, MSE , AL and CP of the parameters of the joint model for $(q_1, q_2) = (0.85, 0.78)$	130
5.2	Bayes estimates and their bias, MSE , AL and CP of the parameters of the joint model for $(q_1, q_2) = (0.65, 0.55)$	131
5.3	Bayes estimates and their bias, MSE , AL and CP of the parameters of the joint model for $(q_1, q_2) = (0.35, 0.25)$	132

List of Figures

3.1	TTT plot for (a) X_1 (time to death of patients in group 1) and (b) X_2 (time to death of patients in group 2)	78
3.2	TTT plot for (a) time to blindness of eye under laser treatment observations, X_1 , and (b) time to blindness of the eye with no treatment observations, X_2	80
3.3	Survival function estimate for DRS data	81
4.1	Estimated survival plots for observed prostate cancer data (solid line) and predictive datasets.	104
4.2	Estimated survival plots for observed DRS data (solid line) and predictive datasets.	107
5.1	Trajectory of the longitudinal response for 20 randomly selected individuals from simulated data	129

List of Abbreviations

AB:	Absolute Bias
AIC:	Akaike Information Criterion
AL:	Average Length
BIC:	Bayesian Information Criterion
BF:	Bayes Factor
$\mathcal{B}(\cdot, \cdot)$:	Beta Distribution
χ^2 :	Chi-square Statistic
CFR:	Constant Failure Rate
CP:	Coverage Probability
cdf:	Cumulative Distribution Function
DFR:	Decreasing Failure Rate
Dir :	Dirichlet Distribution
$\mathcal{E}(\cdot)$:	Exponential Distribution
$\mathcal{G}(\cdot, \cdot)$:	Gamma Distribution
GLD:	Generalized Lifetime Distribution
HPD:	Highest Posterior Density
IFR:	Increasing Failure Rate
$\mathcal{IG}(\cdot, \cdot)$:	Inverse Gamma Distribution
LMM:	Linear Mixed-Effect Model
MCMC:	Markov Chain Monte Carlo
MOBGLD:	Marshall-Olkin Bivariate Generalized Lifetime Distribution
MSE:	Mean Square Error
M-H:	Metropolis-Hastings
\mathcal{MD} :	Multinomial Distribution
$\mathcal{N}(\cdot, \cdot)$:	Normal Distribution
Θ :	Parameter Vector
\mathbb{R}^+ :	Positive Real Line
PPV:	Posterior Predictive p -Value
pdf:	Probability Density Function
RPPI:	Reference Prior with Partial Information
SELF:	Squared Error Loss Function
$\mathcal{T}\mathcal{N}(\cdot, \cdot)$:	Truncated Normal Distribution
$\mathcal{U}(\cdot, \cdot)$:	Uniform Distribution
$\mathcal{W}(\cdot, \cdot)$:	Weibull Distribution

Chapter 1

Introduction

Human life is subject to various events those are inherently linked to individuals. These events encompass multiple aspects of human life such as birth, death, marriage and so forth. Additionally, the failure of systems that individuals rely on also has a pivotal role in their life. Researchers are always interested in analysing the time to the onset of these events. In statistical inference, the idea of random sampling from a population of interest leads to drawing inferences about the characteristics of that population. The researcher deals with random quantities on which statistical models are based and has different properties than the deterministic models. The events discussed above are random in nature in the sense that one can not be certain about their occurrence. The uncertainty associated with those events should be addressed by making an appropriate choice of the statistical model. From the perspective of the estimation process of the parameters of a statistical model, the classical framework and the Bayesian framework are two different approaches in the literature. The Bayesian paradigm provides flexibility of incorporating any kind of non-sample information regarding the characteristics which are often represented by the parameters themselves or function of the parameters. The time-to-event phenomena deal with random quantities measured on positive real lines, hence they have required special statistical techniques apart from the usual ones. Time-to-event data primarily arises in clinical research, demography, epidemiology, public health research and reliability engineering. In these areas of study, individuals or systems are observed over a specific time period and their status of the event is recorded along with the time-to-event.

1.1 Time-to-Event Study

Time-to-event studies are predominantly observed in the fields of survival and reliability analysis. These two analyses consist of statistical methods which are specifically designed to analyse data related to the occurrence of events over time. The event of interest related to living being is mainly studied under survival analysis while the event of interest related to systems/components is explored within the reliability analysis framework. This encompasses a wide range of events including the birth of a child, the death of an individual, the diagnosis of a disease, hospital admission, the failure of a component, the breakdown of a system and many more. The variable of interest i.e., time of an event has different terminologies such as survival time, diagnosis time, death time, infection time, and failure/breakdown time according to the field from which the observations are recorded. Consider a real-life data study where individuals are enrolled in a trial. The objective of the study is to analyse the risk of having been diagnosed with a certain disease during the study. The time to diagnose the disease differs from individual to individual. In such studies, one may not observe the time-to-event for some individuals due to the termination of the study at a prefixed time. The data obtained in such a manner is subject to censoring, an incompleteness of observing time variable.

1.1.1 Censoring

In the study of lifetime variables, the units (systems/components) or individuals are observed over a certain period of time. The lifetime studies are subject to certain limitations imposed by time constraints and the resources required for their implementation. It is not possible to observe the exact time-to-event for each unit/individual due to relying on such constraints. It may be the case that some of them are still alive at the end of the study period while others may have left the study at any time during the study due to unavoidable circumstances. The observed time-to-event data is known as right-censored observations in this scenario as one has only information that the individual is still alive by time, say T. Right-censored observations are a special case of type-I censoring.

In type-I censoring, the units are put on the test and observed until either they fail by some time T or still functioning by the prefixed termination time (T_0). In such a case, the termination time of the test is fixed before the experiment begins.

It is sometimes possible that a sufficient number of failures are not observed in type-I censoring which results in inconsistent estimation. To overcome this issue, the number of failures, say r , is prefixed before starting the experiment and the units are observed until the r number of units get failed. This type of censoring where the test is terminated at the time of the r^{th} failed unit is known as type-II censoring. For more details about different censoring schemes, one may refer to Balakrishnan (1990) and Balakrishnan and Aggarwala (2000).

In retrospective studies, it is observed that for some individuals, the event has already occurred before the time of monitoring. Such type of censoring where it is only known that the time-to-event is less than some observed time say t , is called left censoring.

Many times, it is not possible to observe the units or individuals continuously. In this case, they are observed at pre-fixed times. In this scenario, the observed time-to-event is only known to be lying between two consecutive monitoring times say, t_1 and t_2 . This type of censoring is known as interval censoring (see Chen et al. 2012) and it occurs frequently in clinical settings.

1.1.2 Recall-Based Time-to-Event Data

In retrospective studies, individuals are monitored at specific time points to record the status of an event of interest. The obtained data in such cases are referred to as current status data. If the individuals are asked to recall the time-to-event in case they were already experienced it by the time of monitoring, the experimenter has one more layer of data. In this framework, the obtained time-to-event data regarding the individuals who recall it is known as recall-based time-to-event data. For example, in *NFHS* Round IV (IIPS and ICF 2017), individual mothers were monitored and asked about the status of breastfeeding of their recently born child. The observations corresponding to mothers who are still breastfeeding their children are considered right censored. Some of the mothers are currently not breastfeeding their

children and the time to breastfeeding for them was recorded retrospectively by asking them to recall the particular time they stopped breastfeeding their babies. This is an excellent example of recall-based time-to-event data which is further utilised in this thesis for analysis purposes.

1.1.3 Competing Risks Masked Data

In time-to-event studies, it is often observed that individuals may experience failure with different causes, with only one cause (risk) associated with each individual leading to their failure. For instance, in clinical trials aimed at assessing the effectiveness of treatment on a specific disease, it is possible for subjects to experience mortality due to a cause different from that disease. In such studies, in addition to observing the time-to-event for the individual, the cause of the event is also recorded. This type of data is commonly referred to as competing risks data within the statistical literature. For a comprehensive understanding of the concepts and analysis of competing risks data, one can see Sinha (1986), Crowder (2001) and Lawless (2003).

Competing risks data are commonly encountered in fields such as reliability engineering, public health studies, clinical trials, demography and epidemiology. In reliability engineering, complex systems consist of multiple components in a series configuration. The failure of one of the components leads to the failure of the entire system. In this case, the observed data will be the time of failure of a system along with the corresponding cause (component) of the failure. In studies concerned with competing risks, the true cause of failure for some individuals is not known sometimes due to some unavoidable reasons such as diagnostic limitations, cost of the procedure and many more. The data obtained in this framework becomes incomplete with respect to the cause of failure. This problem of incompleteness within competing risks data is known as masking. Miyakawa (1984) analysed the competing risks data with the masking problem and provided parametric and non-parametric estimation procedures to deal with it. Usher and Hodgson (1988) generalized the problem with a three component series system and analysed the competing risks data with complete and partial masking. Sen et al. (2010) considered the masking problem under the Bayesian paradigm

while modelling the risk under proportional hazard assumptions and illustrated the established methodology using breast cancer data. In a recent study by Rai et al. (2021), the authors relaxed the assumption of symmetry for the probability of masking and considered it as a function of cause and time separately. They analysed the masked competing risks data with Lindley distribution under classical and the Bayesian paradigm.

Most of the studies conducted so far assume independency among the competing risks which is not feasible always. Since the components of complex systems share the same working environment (e.g., temperature, shocks, voltage etc.), this may substantially result in dependency among them. For instance, a battery being used for supplying electric power may fail due to ageing or a result of a random shock, such as overheating. In this case, the two risks of failure cannot be assumed to be independent of each other. In cancer studies, treatment efficacy is the primary event of interest corresponding to the cancer patients in the study. The treatment may fail either due to relapse or the death of the patient under study. These two endpoints of the event of interest (i.e. treatment failure) constitute competing risks data. One can not avoid the potential interdependence between these two risks of failure. Guan et al. (2013), Feizjavadian and Hashemi (2015), Xu and Zhou (2017) and Shen and Xu (2018) analysed the cause dependency in competing risks data using Marshall-Olkin bivariate set-up when the simultaneous occurrence of different risks has non-zero probability. The simultaneous occurrence of risks induces the dependency among causes in this case and needs to be incorporate into the model.

1.1.4 Cause-Specific Hazard Model

The competing risks data considered so far does not include any covariate information which may potentially have an effect on the time-to-event process of risks. To include the covariate information into the model and to analyse their effects on the risks process, Cox (1972) introduced a model known as the Cox-proportional hazard model. For a given set of covariates, say z , the model for the time-to-event process is given by

$$h(t \mid z) = h_0(t) \exp(\beta' z),$$

where $h(t | z)$ is the hazard function at time t conditioning on z , $h_0(t)$ is unspecified baseline hazard function in absence of z and β is vector of regression coefficients corresponding to covariates. On a similar ground, the cause-specific Cox-proportional hazard model is defined as

$$h_k(t | z) = h_{0k}(t) \exp(\beta'_k z), \quad k = 1, \dots, K,$$

where index k refers to the k^{th} cause.

The hazard function is always of interest to researchers as it provides a summary of how the risk of failure changes with respect to time. In Cox-proportional hazard modelling, the estimation of regression coefficients can be done without specifying $h_{0k}(t)$. For doing this, the estimation procedure is generally carried out using partial likelihood as the name suggests it is only a function of β while treating $h_{0k}(t)$ as a nuisance parameter.

1.2 Longitudinal Study

In follow-up studies, the researchers often collect data on a response variable, say Y , at prefixed time points over the individuals (or participants) in addition to time-to-event data. The data collected in this manner constitutes repeated measurements on Y over different time points and are said to be longitudinal data (LD). It is different from cross-sectional data as in the latter, observations are taken on Y at a single time point unlike in LD for individuals. LD can be predominantly seen in public health, social studies, biological, economics and behavioural sciences. In AIDS clinical studies (Goldman et al. 1996), individuals who are infected with HIV are visited at a number of time points and the researchers collect data on CD4 cell counts (a potential biomarker for AIDS) at each visited time. The trajectory of CD4 cell counts is traced to monitor the progression of the HIV virus in the individual's body and consequently help in diagnosing AIDS. Another interesting example can be found in cancer trials. In diagnosing prostate cancer in male participants, the LD are recorded on response variable prostate-specific antigen (PSA) over a number of prefixed time points. The progression of PSA values can help in diagnosing prostate cancer. To model the mean trajectory of longitudinal response Y and analyse the effect of covariates such as age, gender,

body mass index (BMI) and race, different estimations methods have been proposed in the statistical literature. We discuss some of them as follows:

1.2.1 Repeated Measures Analysis of Variance

Repeated measures analysis of variance (ANOVA) is the traditional method to deal with longitudinal measurements. It is generally used for balanced LD i.e., individuals share common time points of observation having an equal number of repeated measurements. Longitudinal responses can be seen as analogous to the randomized block design (Scheffe 1999; Yates 1935) in which individuals are assumed to be blocks and repeated measurements are made on each block. The univariate repeated measures ANOVA model for n individuals each with m_i responses can be written as

$$Y_{ij} = \mathbf{X}'_{ij}\gamma + b_i + \epsilon_{ij}; \quad i = 1, \dots, n; j = 1, \dots, m_i,$$

where γ is vector of regression coefficients corresponding to design matrix X , b is random subject effect with normality assumption i.e., $b_i \sim \mathcal{N}(0, \sigma_b^2)$ and measurement error $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. The repeated measure ANOVA is used to compare the group means for individuals on different time points and correspondingly testing of hypotheses and confidence intervals estimation can be performed (Sullivan 2008). It makes strong assumptions regarding the correlation structure among repeated measurements of an individual. This assumption is called compound symmetry correlation structure which assumes that there is a constant correlation among repeated measurements of the same individual: constant variance $Var(Y_{ij}) = \sigma_b^2 + \sigma_\epsilon^2$ and constant covariance $Cov(Y_{ij}, Y_{ik}) = \sigma_b^2$. The normality assumption for response variable Y limits its applicability in analysing LD. It can only handle time-independent covariates and has limitations to handling only balanced longitudinal responses.

1.2.2 Generalized Estimating Equations

Generalized estimating equations (GEE) are one of the statistical methods used for modelling LD. In the GEE method, one models the mean response with a separate correlation structure accounting for correlation among repeated measurements on the same individual (Zeger et al. 1988). It comprises of two-step method. In the first step, a regression model is defined for the mean response of Y for estimating the population-averaged effect of covariates and in the second step, a working correlation structure is selected. In this case, the form of the regression model is flexible based on the type of longitudinal response. It can be a linear model, logistic regression model or any generalized linear model. For LD, it utilizes the marginal model approach and specifies the structure of mean response as a function of covariates only, i.e., $E(Y_{ij}) = f(X_{ij})$. Estimating the parameters of interest, it does not involve the likelihood function and relies on the solution of regression equations for mean and covariance structures. The GEE method considers the interdependency among the repeated measurements on the same individual while avoiding any other dependency among individuals e.g., the individuals may belong to the same family or clusters inducing correlation among them.

1.2.3 Linear Mixed-Effect Modelling

The LMM is a statistical method by which once can analyse the longitudinal trajectory for an individual over time giving a summary of the process through it evolves. To incorporate the within and between individuals variation, the mixed-effect model approach is perhaps one of the most popular methods (Laird and Ware 1982; Elashoff et al. 2016). The mixed-effect model uses a conditional model approach where the mean response of repeated measurements is modelled as a function of covariates conditioning on random effects. It incorporates the fixed effect which accommodates the population-level effect and the random effect which accommodates individual-level variation. The general form of LMM for

longitudinal response Y_i is given by

$$Y_i = \mathbf{X}'_i \gamma + \mathbf{Z}'_i b_i + \epsilon_i; \quad i = 1, \dots, n; j = 1, \dots, m_i,$$

where \mathbf{X}_i is a vector of explanatory variables which may or may not consist of time-varying covariates, γ is the vector of fixed-effects corresponding to \mathbf{X}_i , Z_i is the design matrix corresponding to random effect b_i . The measurement error $\epsilon_i(\tau)$ is accountable for the variation in Y_i which cannot be explained by the model itself. It is assumed to be normally distributed with mean zero and variance σ^2 , i.e., $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$.

The LMM model considers the complete joint distribution of repeated measurements and utilizes likelihood method for the estimation of unknown parameters of the model. It can be extended to accommodate any outliers or skewness in response variable Y and provides flexibility to model any kind of LD.

1.3 Estimation Methods

To draw inferences about the population characteristics, statistical models are established on the basis of random samples taken from the population of interest. In the parametric modelling of derived samples, the unknown parameters or some well-defined functions of the parameters are of interest. These parameters resemble the population characteristics in which researchers are primarily interested. In statistical inference, the estimation procedure is carried out mainly under two paradigms: classical and Bayesian. We will discuss the estimation method under the Bayesian paradigm as follows:

1.3.1 Bayesian Estimation

The classical method of estimation relies primarily on the observed data for making inferences considering parameter θ as an unknown but fixed quantity. In many situations, it becomes inevitable to incorporate non-sample information about the unknown parameters of the model. The Bayesian method of estimation provides flexibility to incorporate such information by assuming θ as a random quantity. The non-sample information about the

parameter of interest is generally available in the form of prior information. The prior information is the key to drawing inferences about θ under the Bayesian paradigm. The simple way to incorporate the prior information into inferences is to translate this information into a probability distribution. The probability density, say $\pi(\theta)$, depicts the prior belief about θ and formulate the uncertainty in values of θ through probability distribution. In the Bayesian framework, the prior belief about θ is updated on behalf of the observed data \underline{t} through likelihood function $L(\theta | \underline{t})$ and results in posterior distribution. The posterior density, say $p(\theta | \underline{t})$, is utilised to draw any inference about the parameter under the Bayesian paradigm.

Using the renowned Bayes theorem, one obtains the posterior density of θ for given data \underline{t} as

$$p(\theta | \underline{t}) = \frac{L(\theta | \underline{t})\pi(\theta)}{\int_{\Theta} L(\theta | \underline{t})\pi(\theta) d\theta},$$

where the notations have their usual meaning. The denominator term represents the marginal density of random variable T which is marginalised with respect to θ .

The Bayesian method of estimation comprises of specifying the prior density, $\pi(\theta)$ and then drawing posterior samples from $p(\theta | \underline{t})$ defined above. The posterior density does not come into analytical form always which makes the process of generating posterior samples quite tedious in form of computational burden. The choice of priors and their specification is always a fundamental problem in the Bayesian framework and there is no unique way to specify the prior distribution. The specification of priors is usually done on the grounds of available prior information and misspecification can lead to incorrect inference. If the information regarding the parameter θ is informative in the sense that it can be a resemblance to any well-known probability distribution, one has to proceed with informative priors. If the information is very little or vague, one should proceed with non-informative priors to conduct a posterior analysis. The whole Bayesian method of estimation can be bifurcated into two dimensions: (i) objective Bayesian method, and (ii) subjective Bayesian method of estimation. One can refer to Jaynes (1968); Tiao and Box (1973) and Bernardo (1979) for a detailed study of the choice of priors and their applicability in different scenarios. We mainly focused on the objective Bayesian approach for the estimation of parameters of considered

models throughout the entire thesis except in the last chapter.

Informative Prior

In case, when the available information is not vague, it can be translated into a probability distribution to specify the prior. Thus, the Bayesian inference is carried out with informative priors. In this framework, the prior distribution of the parameters is characterized by some unknown quantities. These are called hyper-parameters with respect to the prior distribution. Informative priors are based on existing knowledge or information available from expert opinion, previous studies, historical data or other reliable sources. Subjective priors are somewhat different from informative priors in the sense that they are based on an individual's personal belief about the parameter of interest. One of the popular informative prior is conjugate prior. If the defined prior distribution and obtained posterior distribution belong to the same family of distribution then the prior is said to be conjugate prior. The well-known distribution such as normal, beta and gamma distributions are a few examples of conjugate priors.

Non-informative Prior

The non-informative priors are useful in conducting the Bayesian analysis when the available prior information regarding the parameter is quite vague in nature. The formulation of priors in this case is performed in the sense that the Bayesian inference based on the obtained posterior density is dominated by the observed data through likelihood function most and the prior has least dominance onto it. The non-informative prior specification leads to posterior inference under objective Bayesian method of estimation. Consider a situation where the prior belief of a researcher for a one-dimensional parameter θ is suspicious and one can not assign weights to specific values of θ , a uniform prior can be defined as

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

Here, the vagueness in the information available for θ , does not permit to assign weights to distinct values in support of θ , hence a uniform prior is most suitable to formulate the prior distribution which equally assigns weight in the entire range of θ . Jeffreys (1946) formulated one of the non-informative prior in the sense that the observed Fisher information matrix holds the information contained in the sample about the parameter θ . It has invariant properties under reparametrization. The Jeffreys priors is defined as the square root of the determinant of Fisher information matrix as

$$\pi(\theta) \propto |I(\theta)|^{1/2},$$

where $I(\theta)$ is the Fisher information matrix based on observed sample. For multi-parameter case, Jeffreys prior has limitations in its applicability.

Reference priors are one of the most popular non-informative prior available in statistical literature of objective Bayesian method. Reference prior is analogous to Jeffreys prior in case of one-dimensional parameter. The idea behind the formulation of reference priors is that it derived in such a way the Kullback-Leibler divergence between the prior and obtained posterior is maximum resulting in the data dominated Bayesian inference by diffusing prior's effect onto the posterior. The key idea behind the formulation of reference prior is that it maximizes the expectation of Kullback-Leibler divergence between the prior and posterior density asymptotically, i.e.

$$\pi_r(\theta) = \arg \max_{\pi(\theta)} K(\theta, X_n),$$

where

$$K(\theta, X_n) = E_X \left[\int \pi(\theta|X_n) \log \left\{ \frac{\pi(\theta|X_n)}{\pi(\theta)} \right\} d\theta \right],$$

which is the expected Kullback-Leibler divergence.

In statistical inference, sometimes researchers are more interested in few of the parameters. The reference prior provides a flexible framework under the objective that the inference can be done efficiently for the parameters of interest by treating the remaining parameters as a nuisance one without compromising with the efficacy of results. For detailed review and discussion on derivation and usefulness of reference priors, one can see Bernardo (1979);

Berger and Bernardo (1992); Ghosh and Mukerjee (1992) and Berger (2006). Before deriving the reference priors, the parameters of the model is ordered in a set. The ordering is done on the basis of inferential importance of the parameters. There can exist more than one reference priors for a particular vector of parameters based on the different ordered set of those parameters. For example, consider a ordered set of parameters, say $\theta = \{\theta_1, \theta_2, \theta_3\}$; here the parameter θ_1 is the most important one and θ_3 is the least important parameter on the basis of inferential objective. One can make different ordering of the parameters according to one's own inferential objective.

1.4 Advance Computational Techniques

In time-to-event data observed from various fields of study, the complexity of the data can pose challenges. Analysing such data sets often requires the use of complex statistical models. These models may involve one or multi-dimensional integrals that cannot be solved analytically. Moreover, within the objective Bayesian paradigm, the resulting posteriors do not take a closed form or resemble any standard distribution. As a result, obtaining posterior samples adds an additional computational burden to the estimation process. Fortunately, in this fast-developing era of technology, computers offer a solution to tackle such complex problems through simulation process. Computational techniques serve as powerful tools to solve these challenges through step-wise, repeated, and iterative procedures, offering increased efficiency compared to other methods.

1.4.1 Markov Chain Monte Carlo Method

In the Bayesian method of estimation, the form of posterior does not always has an analytical form which poses a challenge for generating samples. This leads to a restriction on estimating the parameter or any function of the parameter of interest. Suppose the objective is to estimate $g(\theta)$, a function of parameter θ , then the Bayesian estimator of it under the squared

error loss function is simply defined as

$$E(g(\theta)) = \int_{\theta \in \Theta} g(\theta) p(\theta | \underline{t}) d\theta$$

The Markov chain Monte Carlo (MCMC) method is a class of algorithms for generating posterior samples from the density $p(\theta | \underline{t})$. One can refer to Metropolis et al. (1953) and Hastings (1970) for the historical development of MCMC methods. The samples generated under this method follow the Markovian property, hence the name Markov chain. The stationary distribution of generated samples converges to the desired probability distribution (posterior density in our case) after a large number of iterations. To start the algorithm for generating a correlated chain of observations from target density, a starting point is needed. The generated posterior samples are very sensitive to this initial value and hence, few observations from the starting point are often discarded to remove the impact of the initial value on the chain. Now, suppose we have generated a sequence of Markov chain, say $\{\theta_i; i = 1, \dots, M\}$ from its posterior density by choosing a suitable candidate density, then the Bayesian estimator of $g(\theta)$ under SELF is given by

$$E(g(\theta)) \approx \frac{1}{M} \sum_{i=1}^k g(\theta_i).$$

The above result is ensured by the ergodic theorem and the chain constructed is aperiodic and irreducible in nature. We can summarize the MCMC method as consisting of selecting a suitable candidate density and an initial value to initiate the algorithm.

1.4.2 Metropolis-Hastings Algorithm

The Metropolis-Hastings (M-H) algorithm is one of the MCMC methods to generate samples from the non-standard target density. In the M-H algorithm, to generate samples from a density which is not in standard form, one first chooses a proposal density keeping in mind the support of the target density. The chain of random samples is generated from the proposal density and each observation is accepted or rejected on the ground of probability of acceptance. Thus, the Markov chain generated from the proposal tends to behave in such

a way that it is sampled from the target density. Now, suppose the target density is $p(\theta)$ and the proposal density is $q(\theta)$, then generating a chain of sample observations from the one-dimensional target density, the M-H algorithm is defined as

- Start with initial value θ^0 . For iteration $k = 1, \dots, M$
- Generate a sample observation from proposal density. Symbolically $\theta \sim q(\cdot | \theta^{(k-1)})$
- Calculate the probability of acceptance

$$\varphi(\theta, \theta^{(k-1)}) = \min \left(1, \frac{q(\theta^{(k-1)} | \theta)}{q(\theta | \theta^{(k-1)})} \cdot \frac{p(\theta | \theta^{(k-1)})}{p(\theta^{(k-1)} | \theta)} \right)$$

- Generate uniform random number, say u from $\mathcal{U}(0, 1)$

If $u \leq \varphi(\theta, \theta^{(k-1)})$, accept θ by setting $\theta^{(k)} = \theta$,

otherwise set $\theta^{(k)} = \theta^{(k-1)}$.

If the proposal density is symmetric, i.e., $q(x | y) = q(y | x)$, e.g., normal distribution, then the probability of acceptance becomes simplified to

$$\varphi(\theta, \theta^{(k-1)}) = \min \left(1, \frac{p(\theta | \theta^{(k-1)})}{p(\theta^{(k-1)} | \theta)} \right).$$

On executing the M-H algorithm defined above, we get a sample of Markov chain $(\theta^{(1)}, \dots, \theta^{(M)})$ whose stationary distribution converged to the target distribution. It is common practice to discard some of the starting values of the Markov chain as it does not follow the stationary distribution in general. So, we discard the first K sample observations and the reduced chain $(\theta^{(K+1)}, \dots, \theta^{(M)})$ is used for further analysis. The M-H algorithm can also be used for multi-parameter cases in the same way as it is used for single-parameter cases with the limitation that the acceptance rate decreases with the increasing dimensionality of the target density.

1.4.3 Metropolis Adjusted Langevin Algorithm

The Metropolis adjusted Langevin algorithm (MALA) is an alternative to the M-H algorithm for drawing random samples from a probability distribution for which direct sampling is difficult. The MALA was first given by Rossky et al. (1978) to generate random moves in a physical system. It uses proposal distribution based on the target density. The MALA has a better acceptance rate than the random-walk M-H algorithm as the dimensionality of the parameter vector increases. To proceed with MALA, the proposal point is generated as

$$\theta^{(k)} = \theta^{(k-1)} + \frac{\sigma^2}{2} \nabla \log p(\theta^{(k-1)} | \underline{t}) + \sigma \epsilon_t,$$

where $\nabla \log p(\theta | \underline{t})$ is the gradient vector of natural logarithmic of the target density $p(\theta | \underline{t})$ defined as

$$\nabla \log p(\theta | \underline{t}) = \left[\frac{\partial \log p(\theta | \underline{t})}{\theta_1}, \dots, \frac{\partial \log p(\theta | \underline{t})}{\theta_d} \right]^T,$$

σ^2 is the step of variance, also known as discretization size and random quantity $\epsilon_t \sim \mathcal{N}_d(0, I_d)$; d being the dimension of the parameter vector. To generate posterior samples from the target density using MALA, the steps of the algorithm are the same as defined for the M-H algorithm except for the candidate point generation. After obtaining proposal point, $\theta^{(k)}$ is accepted with probability

$$\varphi(\theta^{(k)}, \theta^{(k-1)}) = \min \left(1, \frac{q(\theta^{(k-1)} | \theta^{(k)})}{q(\theta^{(k)} | \theta^{(k-1)})} \cdot \frac{p(\theta^{(k)} | \theta^{(k-1)})}{p(\theta^{(k-1)} | \theta^{(k)})} \right),$$

where

$$q(\theta^{(k)} | \theta^{(k-1)}) = (2\pi\sigma^2)^{-1/2} \exp \left\{ -\frac{\|\theta^{(k)} - \theta^{(k-1)} - \frac{\sigma^2}{2} \nabla \log p(\theta^{(k-1)} | \underline{t})\|_2^2}{2\sigma^2} \right\}$$

and $\|\cdot\|_2^2$ is usual L_2 -norm. In case if proposal point is not accepted then we set $\theta^{(k)} = \theta^{(k-1)}$ with probability $1 - \varphi$.

The MALA is a modification of the random-walk M-H algorithm in which the random

moves from proposal distribution are drawn incorporating the structure of the target density through gradient vector.

1.4.4 Gibbs Sampling Algorithm

The Gibbs sampler is another MCMC method to construct a Markov chain from the joint probability distribution when it does not have any explicit form. The name of this algorithm is after the physicist Josiah Willard Gibbs, due to the analogy between sampling algorithm and statistical physics. For more details about the Gibbs sampling algorithm, one can refer to the work of Geman and Geman (1984). The Gibbs sampling algorithm is useful in cases when the direct generation from the joint density is not possible rather than generating from the conditional densities. Consider a joint posterior distribution $p(\theta, \eta)$. Suppose, it does not have any analytical form and as a result, it hinders the process of sample generation from the joint density directly. But, for some reason, sample generation from conditional densities $p(\theta | \eta)$ and $p(\eta | \theta)$ is relatively possible. Thus, the Gibbs sampling algorithm enables one to draw samples from their conditional densities which forms a Markov chain.

In this case, the steps of the Gibbs sampling algorithm are defined as

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

- For $k = 1, \dots, M$

$$\text{Generate } \theta^{(k)} \sim p(\theta | \eta^{(k-1)}) ,$$

$$\text{Generate } \eta^{(k)} \sim p(\eta | \theta^{(k)})$$

- In the end, a Markov chain $(\theta^{(k)}, \eta^{(k)}) ; k = 1, \dots, M$ of size M is obtained.

The obtained Markov chain can be used for estimating the function of parameters under a suitable loss function.

1.4.5 Highest Posterior Density Interval

In the Bayesian framework, credible intervals are formed for the parameter of interest analogous to the confidence interval in the classical framework. The highest posterior density

(HPD) interval (Chen and Shao 1999) is the shortest length credible interval obtained for a given level of confidence, say α . The points in the HPD interval have higher probability density than the points excluded from the interval. The probability statement about the parameter in the Bayesian paradigm makes more sense than in the classical paradigm as the parameter is considered a random quantity in the former rather than a fixed quantity in the latter. To obtain the HPD interval for θ , we use the ordered sample, say $\theta^{(1)}, \dots, \theta^{(K)}$ after discarding some of the initial observations from posterior density through MCMC method. That is, the sample ordered in such a way, $\theta^{(1)} \leq \theta^{(2)} \leq \dots \leq \theta^{(K)}$ and use the following steps:

- Obtain all $100(1 - \alpha)\%$ credible intervals, i.e.,

$$(\theta^{[i]}, \theta^{[i+K(1-\alpha)]}) ; \quad i = 1, \dots, \alpha K \text{ and } [r] \text{ is integer part of } r.$$

- Calculate the length of each credible interval as

$$L_i = \theta^{[i+K(1-\alpha)]} - \theta^{[i]}.$$

- Select the interval for which $L_i; i = 1, \dots, \alpha K$ is smallest.

1.4.6 Model Compatibility

In statistical inference, the goodness of fit of the model considered for a given set of data is a significant step in the model-building process. The model compatibility is used to validate the model constructed for the given data in hand. In the classical framework, there are different criteria to check for the compatibility of the model based on some graphical measures and test statistics. The Bayesian framework also provides a number of tools for model compatibility (i.e., likelihood and prior) for the considered dataset. It utilizes the predictive simulation idea based on the posterior density to see whether the predictive observations resemble the real data observations. A primary intuition about the model compatibility can be obtained through a graphical approach by plotting the survival or hazard function for the predictive datasets based on the model considered and the same for the real dataset. A superimposition of the survival/hazard curves for the predictive

dataset onto the curve for the real dataset depicts some realization of the compatibility of the model under consideration. One can refer to the work of Gelman et al. (1996) for a detailed discussion on model compatibility. A tail area-based measure known as the posterior predictive p-value (PPV) is available in the Bayesian paradigm for model compatibility.

1.4.7 Posterior Predictive p -Value

In the thesis, to check for the model compatibility of the real datasets under the Bayesian paradigm, we use PPV under the Bayesian paradigm. The PPV is based on the χ^2 test statistic which measures the discrepancies between the expected value of a statistic calculated from the predictive dataset and the observed dataset. For calculating the value of PPV , we use a discrepancy measure based on chi-square statistic (Upadhyay and Smith 1993, 1994) is given below

$$\chi^2 = \sum_{i=1}^n \frac{(t_i - E(t_i^* | \theta))^2}{V(t_i^* | \theta)},$$

where t_i^* is the predictive observation for i^{th} individual. The PPV based on the discrepancy measure defined above can be obtained as

$$PPV = \int_{\Theta} P[\chi_2^2 > \chi_1^2 | \theta] p(\theta | \underline{t}) d\theta,$$

where χ_1^2 and χ_2^2 are the calculated values of χ^2 for the observed data and predictive data sets, respectively. The PPV is defined as the expected value of the classical p-value, where the expectation is taken with respect to posterior density. A large PPV indicates towards the model compatibility for the observed data.

Note: As highlighted by Upadhyay et al. (2001), caution is advised against the use of PPV for studying model compatibility. Its primary utility lies in providing tentative guidance on the model's compatibility with the observed dataset. The limitation however arises from its double use of observed data, constraining its applicability to some extent though its simplicity and versatility for various prior choices make it a viable option.

Chapter 2

Objective Bayesian Approach for Recall-Based Time-to-Event Data

2.1 Introduction

In a parametric model, the specification of its parameters elucidates important characteristics of the random variable being studied. In statistical inference, the focus is often on inferring about the unknown parameters of the considered model. The estimation of these unknown parameters is carried out by considering them as random variable rather than fixed quantity under the Bayesian paradigm unlike in the classical method of estimation. The prior information about these unknown parameters plays an important role while executing the estimation process. The main concern in Bayesian inference is the elicitation of possibly suitable prior for the parameters. The parameters assumed to have probability distribution are further characterized by their own parameters called hyper-parameters. To elicit the appropriate form of priors, one may use historical data or expert advice. But historical data or expert advice is mainly used for model building with which they are more familiar as mentioned by Berger (2006). In many situations, researchers may not be able to encapsulate the prior information into a probability distribution, e.g., when the parameter space has a large dimension. In this regard, the objective Bayesian analysis comes out in the role. The term objective has broader meaning in this context. In literature, other names such as non-informative, reference, conventional and non-subjective are used interchangeably for

it. In the objective Bayesian analysis, the prior elicitation is done in such a way that the posterior inference is primarily dominated by the likelihood of the observed data at hand.

One of the most popular objective priors is reference prior introduced by Bernardo (1979) and further developed by Berger and Bernardo (1992). The idea behind using the reference prior for objective Bayesian inference is that it maximizes the divergence between the prior and posterior so that the effect of prior on the posterior inference can be diffused. In order to derive the reference prior, the model parameters are kept in a group where the order of the parameter matters according to its inferential importance. Thus if someone is more interested in a parameter than the rest of the others, the Bayesian inference can be done using reference prior without losing the precision related to the parameter of interest. Hence, a number of reference priors can be derived for a given vector of parameters according to the inferential order of the parameters. Also, the derived reference priors come out to be improper often, so propriety of the obtained posterior distributions is needed to be established. Sun and Ye (1995) derived two reference priors for the product of means of n normal distributions and established the propriety of one of them. Similarly, Berger et al. (2001) obtained reference prior for spatially correlated data under the objective Bayesian paradigm. Further, Fu et al. (2012) analysed Pareto distribution using an objective Bayesian approach under progressive type-II censoring and derived reference priors. Thus, the objective Bayesian approach has found its application in different fields of data making it a possibly good choice when one has no prior information or expertise about the model parameters.

Time-to-event studies are primarily concerned with the analysis of data related to the time corresponding to an event of interest. The event of interest depends on the field from which the data is observed. For example, in demographic surveys, the event of interest may be the age at first marriage and the age at first child born, etc. In the survival field, it may be a diagnosis of a disease, hospital infection, death of a patient, etc. In the reliability field, it may be a failure of a system or a breakdown of a system. Due to many limitations such as time and/or cost restrictions, it is not always possible to observe complete data in an experiment mainly in clinical trials, demographic surveys, etc. One obtains censored data in such experiments when we do not have failure time or time to the event of interest related

to each subject in the experiment. A well-known form of censoring is interval censoring when scheduled or pre-fixed visits are done. In interval censoring, we only observe that the time to an event corresponding to subjects lies in an interval where the lower and upper end of the interval are their two consecutive visits. In this case, the information available is limited to a range or interval indicating when the event occurred, rather than an exact time. It is worth noting that exact or right (left)-censored failure times can be considered a subset of interval-censored failure times. This is because, in the case of exact or right(left)-censored failure times, the interval reduces to a single point or extends infinitely towards the right(left). Independent interval censoring refers to mechanisms where the occurrence of censoring is unrelated to the event of interest.

To reduce time and cost, follow-up studies often utilize a cross-sectional sampling approach to collect survival data. However, this technique may lead to case I interval-censored data (Sun 2006), a simplest type of interval censoring. These data are also referred to as “current status” data, which is a term originating from demographical studies. In current status censoring, each subject is only observed at one point in time. The observed time intervals “include” either time zero or infinity, i.e. the observed time-to-event is either left-censored or right-censored with respect to monitoring time. There are two common types of studies that produce current status data: cross-sectional studies and tumorigenicity experiments on non-lethal tumours. In cross-sectional studies, current status data are often used to study demographic trends. On the other hand, in tumorigenicity experiments, current status data are observed due to the inability to measure the variable directly and accurately. Despite the different origins of these two types of current status data, they are analysed in the same way. For example, Diamond et al. (1986) discussed the estimation of the distribution of age at weaning for a given current status data, where the age of the child at weaning is considered as time to the event and the age of the child is considered as monitoring time. In addition, Keiding et al. (1996) estimated the age-specific immunization under the current status data set-up. One can go through Sun and Kalbfleisch (1993) and Jewell and van der Laan (2003) for detailed works related to current status data. Specifically, for current status data, independent censoring means that the monitoring time and time-to-event are independent of

each other.

In statistical analysis, researchers often collect data at predetermined or random monitoring time points to observe the time-to-event of interest. However, in some cases, recall-based data is necessary. In such studies, individuals are asked to recall the time-to-event at monitoring time if they have experienced the event of interest. The chance of recall decreases with the passage of time, and individuals who have experienced the event recently are more likely to recall the event than those who have experienced the event in the past. Recall-based time-to-event is a type of data commonly used in retrospective studies where participants are asked to recall the time of the event that occurred in the past, such as the onset of a disease or the occurrence of a medical condition. Applications of such studies include case-control studies and retrospective cohort studies. As an example, consider the data on menarche discussed in Salehabadi et al. (2015). The dataset contains details on whether menarche occurred at the time of observation and the age at which it happened. The study is focused on the onset of menstruation, which is the event of interest. The age at which menarche took place is documented only if the individual can recall the timing of the event. However, the accuracy of recall-based time-to-event data can be influenced by several factors, such as the length of the recall period, the nature of the event, and the characteristics of the study population. Therefore, it is important to carefully consider the limitations and potential biases associated with this type of data when analysing and interpreting the results of a study.

This necessitates the use of specialized statistical techniques, such as those employed in the recall-based studies of Gillespie et al. (2006), who assessed the recall accuracy of breastfeeding mothers regarding the age of weaning, and Cooney et al. (2009), who estimated the time to pregnancy and validated the estimate against previous studies. To address the challenges posed by recall-based data, Salehabadi et al. (2015) proposed a parametric model taking exponential and piecewise functional form for the chance of recalling the event of interest, which led to improved estimator performance. In a recent study, Yadav et al. (2022) analysed recall-based competing risks data using both classical and Bayesian approaches, ultimately estimating the age at menarche. However, there is currently no literature available

that focuses on the objective Bayesian analysis of recall-based time-to-event data. The present study aims to address this deficiency.

In the current study, we apply objective Bayesian analysis to recall-based time-to-event data using Weibull distribution. The reference priors for different sets of ordered parameters are derived in presence of latent variables and consequently, we established the propriety of the obtained posterior distributions. To provide the Bayesian estimation under squared error loss function (*SELF*), a simulation study is conducted on different proportions of censored and non-recall observations. The model is applied to breastfeeding data, illustrating its practical applicability.

Further, the chapter is structured in a systematic and rigorous manner, with a comprehensive and illuminating discussion of various key aspects of Bayesian inference in the context of recall-based data scenarios. In Section 2.2, a complete likelihood function is skillfully constructed, accounting for latent variables that arise due to incomplete observations, and the Fisher-information matrix is elegantly derived. Building upon this foundation, an objective Bayesian analysis is undertaken in Section 2.3, whereby reference priors are derived for different ordering groups of the parameters, and the propriety of the posterior distributions is established. In Section 2.4, a simulation study is conducted, carefully designed to elucidate the performance of the Bayes estimators, thus providing valuable insights into the efficacy of the proposed methodology. To showcase the practical utility of the Bayesian paradigm, Section 2.5 presents an application to real-world data, specifically estimating the duration of breastfeeding of a recent child. Finally, the chapter is concluded in Section 2.6, bringing together the various aspects of the study and summarizing its key contributions.

2.2 Likelihood Construction for Recall-Based Data

We first define the structure for recall-based data before constructing the likelihood function for it. Let S be a random variable representing the monitoring time having cumulative distribution function (*cdf*) and probability density function (*pdf*) $G(\cdot)$ and $g(\cdot)$, respectively. Similarly, let T be the random variable for the time to onset of the event of interest with

cdf and *pdf* $F(\cdot)$ and $f(\cdot)$, respectively.

The independent assumption between monitoring time (S) and time-to-event (T) is assumed which implies that the occurrence of the event is not influenced by the length of time the subject is being monitored. This assumption provides justification for selecting the monitoring points based on a separate set of criteria from those that determine the time-to-event of interest. For instance, if we are interested in studying the effect of a treatment on survival time, we may choose monitoring points based on factors such as cost, labour or feasibility, which may not be directly related to the patient's response to the treatment.

Let us consider the recall-based study with n individuals each monitored at a single time point. The status of the event is recorded for each individual. At any monitoring time, more than one individual can be visited but here we are assuming that there is unique monitoring time corresponding to each individual. The time-to-event, T corresponding to an individual is only recorded if (s)he is able to recall it. If an individual is not able to recall the time-to-event of interest, the time of event associated with that individual is said to be left-censored with respect to that monitoring time. There is one more possibility that at the monitoring time, the event has not happened yet for the individual and hence one gets right-censored observation for that individual. Let $(S, \delta, \epsilon, T\delta\epsilon)$ be the observed data vector for recall-based set-up. Here $\delta = I(T \leq S)$ takes value 1 if the event of interest has occurred before monitoring time S otherwise it takes value 0. Also, another indicator ϵ takes value 1 if the individual is able to recall the time of event otherwise it takes value 0. Indicator ϵ is defined for those individuals corresponding to whom the event of interest had occurred before the monitoring time. For i^{th} individual, $(S_i, \delta_i, \epsilon_i, T_i\delta_i\epsilon_i); i = 1, 2, \dots, n$ be the observed data.

Under the current setting, the likelihood contribution of i^{th} individual can be written as

$$L_i(\theta | \cdot) = \begin{cases} f(t_i; \theta)(1 - \psi(s_i, t_i)) & \text{if } \delta_i = 1, \epsilon_i = 1, \\ \int_0^{s_i} f(u; \theta)\psi(s_i, u)du & \text{if } \delta_i = 1, \epsilon_i = 0, \\ 1 - F(s_i; \theta) & \text{otherwise,} \end{cases}$$

where $\psi(s, t)$ represents the non-recall probability which is a function of monitoring time and time-to-event. Alternatively, $1 - \psi(s, t)$ denotes the recall probability.

Here, we assume the functional form of recall probability as exponential in nature. It is reasonable in the sense that the chance of recall reduces exponentially as the difference between monitoring time and time-to-event becomes larger. Also, this functional form is used earlier by Salehabadi et al. (2015) while studying the age of menarche data under recall-based set-up. Hence, the non-recall probability with parameter λ , for an individual who is unable to recall the time-to-event, takes the form

$$\psi(s_i, t_i) = 1 - \exp\{-\lambda(s_i - t_i)\}, \quad (2.1)$$

where $\lambda > 0$ is non-recall parameter. Alternatively, $1 - \psi(s_i, t_i) = \exp\{-\lambda(s_i - t_i)\}$ gives the recall probability.

The recall probability becomes 1 if the monitoring and time-to-event coincide. But it is not suitable under the assumption of continuous time-to-event and monitoring time; as the probability of simultaneous occurrence of both times is zero.

Let us denote $n_r = \sum \delta_i \epsilon_i$, and $n_{nr} - n_r = \sum \delta_i (1 - \epsilon_i)$; $i = 1, 2, \dots, n$. After properly re-grouping the collected time-to-event data, without loss of generality, one can assume that the first n_r data points are subject to recall, middle $(n_{nr} - n_r)$ as non-recall, and the rest $(n - n_{nr})$ as right censored observations, respectively. The total number of individuals in recall-based data can be denoted as: $\{1, \dots, n_r, n_r + 1, \dots, n_{nr}, n_{nr} + 1, \dots, n\}$.

Using the expression of recall and non-recall probabilities, the likelihood function for parameter vector θ given observed data can be written as

$$L(\theta, \lambda | \cdot) = \prod_{i=1}^{n_r} \left[f(t_i; \theta) \exp\{-\lambda(s_i - t_i)\} \right] \prod_{i=n_r+1}^{n_{nr}} \left[\int_0^{s_i} f(u; \theta) \left(1 - \exp\{-\lambda(s_i - t_i)\}\right) du \right] \prod_{i=n_{nr}+1}^n \left[1 - F(s_i; \theta) \right]. \quad (2.2)$$

Under the objective Bayesian approach, the derivation of reference priors relies on an explicit calculation of the Fisher information matrix. However, this task can be challenging

due to the intractable nature of the likelihood function (2.2). To circumvent this problem, we employ a direct calculation of the Fisher information matrix by introducing latent variables and expressing the model in a more general form. In this framework, non-recall and right-censored observations can be treated as missing quantities as partial information on associated time-to-event is available only. For the case of non-recall observations, it is observed that T lies in the interval $(0, S)$. For the right-censored case, T lies in the interval (S, ∞) . Dealing with these quantities, we used the concept of equivalent quantity approach (Tan 2007, 2009; Wang 2016). The concept of equivalent quantity approach aims to find out the missing quantity based on available information. Let us introduce two latent variables T_l^* and T_r^* for non-recall and right-censored observations respectively.

Using these quantities, the likelihood can be re-written as

$$L(\theta, \lambda | \cdot) = \prod_{i=1}^{n_r} \left[f(t_i; \theta) \exp\{-\lambda(s_i - t_i)\} \right] \prod_{i=n_r+1}^{n_{nr}} \left[f(t_{li}^*; \theta) \left(1 - \exp\{-\lambda(s_i - t_{li}^*)\}\right) \right] \\ \prod_{i=n_{nr}+1}^n \left[f(t_{ri}^*; \theta) \right]. \quad (2.3)$$

Here, it is to be noted that monitoring time S is observed corresponding to all individuals. Next, with respect to S , let us denote the observed time vector for recall individuals as $\underline{D} = T$. Further to make the non-recall probability into a more convenient form, another latent variable W_l^* , following an exponential variate with mean $1/\lambda$ truncated at point $(S - T_l^*)$, is introduced. Thus, the missing time vectors for non-recall and right-censored categories are denoted by $\underline{D}^* = (T_l^*, W_l^*)$, and $\underline{\mathcal{D}} = T_r^*$, respectively. Further, let $D_i = T_i$; $i = 1, \dots, n_r$ be the observed time corresponding to individuals under recall category; $D_i^* = (T_{li}^*, W_{li}^*)$; $i = n_r + 1, \dots, n_{nr}$ be the missing observations corresponding to individuals in non-recall category, and similarly $\mathcal{D}_i = T_{ri}^*$; $i = n_{nr} + 1, \dots, n$ denotes the missing observations corresponding to the individuals in the right-censored category. Finally, the complete data vector is written as $\underline{Y} = (\underline{D}, \underline{D}^*, \underline{\mathcal{D}})$.

In light of complete data, the complete likelihood function can be written as

$$L(\theta, \lambda | \underline{Y}) = \prod_{i=1}^{n_r} \left[f(t_i; \theta) \exp\{-\lambda(s_i - t_i)\} \right] \prod_{i=n_r+1}^{n_{nr}} \left[f(t_{li}^*; \theta) \lambda \exp\{-\lambda w_{li}^*\} \right] \prod_{i=n_{nr}+1}^n \left[f(t_{ri}^*; \theta) \right]. \quad (2.4)$$

Since, we have obtained the complete likelihood, thus the next step is to obtain the estimates under suitable choices of time-to-event distribution. The widely known Weibull distribution is used to model datasets that are utilized generally in life testing, medicine and reliability engineering. With a shape parameter greater than 1, it has an increasing rate of failure and thus makes it suitable to model life testing experiments having components that age with time. If the shape parameter is less than 1, the nature of the hazard is decreasing. If the shape parameter is equal to 1, the hazard rate becomes constant, mainly used for modelling the failure of electronic devices. Thus having flexible nature to capture different types of hazard behaviour, Weibull distribution is assumed for time-to-event.

Let us assume that the time-to-event variable follows Weibull distribution, i.e. $T \sim W(\beta, \alpha)$ (β -shape and α -scale). Also, the conditional densities of latent variables T_l^* and T_r^* are assumed to be truncated Weibull (β -shape and α -scale) distributions with restricted support $(0, S)$ and (S, ∞) respectively. Updating the parameter vector as $\Theta = (\beta, \alpha, \lambda)$, the complete likelihood becomes

$$L(\Theta | \underline{Y}) = \prod_{i=1}^{n_r} \left[\alpha \beta t_i^{\beta-1} \exp\{-\alpha t_i^\beta\} \exp\{-\lambda(s_i - t_i)\} \right] \prod_{i=n_r+1}^n \left[\alpha \beta t_{ri}^{*\beta-1} \exp\{-\alpha t_{ri}^{*\beta}\} \right] \\ \prod_{i=n_r+1}^{n_{nr}} \left[\alpha \beta t_{li}^{*\beta-1} \exp\{-\alpha t_{li}^{*\beta}\} \lambda \exp\{-\lambda w_{li}^*\} \right]. \quad (2.5)$$

Taking the natural logarithm of (2.5), the complete log-likelihood function can be written as

$$l(\Theta | \underline{Y}) = n \ln(\alpha) + n \ln(\beta) + n_{nr} \ln(\lambda) - \alpha \sum_{i=1}^{n_r} t_i^\beta - \alpha \sum_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta} - \alpha \sum_{i=n_{nr}+1}^n t_{ri}^{*\beta} \\ - \lambda \sum_{i=1}^{n_r} (s_i - t_i) - \lambda \sum_{i=n_r+1}^{n_{nr}} w_{li}^* + (\beta - 1) \sum_{i=n_{nr}+1}^n \ln(t_{ri}^*)$$

$$+ (\beta - 1) \sum_{i=1}^{n_r} \ln(t_i) + (\beta - 1) \sum_{i=n_r+1}^{n_{nr}} \ln(t_{li}^*). \quad (2.6)$$

The idea of deriving reference prior is solely based on the Fisher information matrix. Since, it is too cumbersome to derive reference prior using original likelihood (2.2), so we use the complete likelihood (2.5). The Fisher information matrix is obtained by

$$I(\beta, \alpha, \lambda) = -E_{D^*, \mathfrak{D}} \left[E_D \left(\frac{\partial^2}{\partial \Theta^2} l(\Theta | \underline{Y}) \right) \right]. \quad (2.7)$$

To calculate the Fisher information matrix, we first obtain the expectation of observed variables given latent vectors and then take the expectation with respect to missing data (Chen et al. 2010; Li et al. 2020). The expectation terms required later are calculated using the conditional densities of latent variables T_l^* , T_r^* and W_l^* .

2.3 Objective Bayesian Inference

In this section, we carry out the objective Bayesian inference for the recall-based observations. The reference priors based on the ordered group of parameters are derived and the propriety of the obtained posteriors is established.

2.3.1 Reference Priors

Reference priors first derived by Bernardo (1979) is one of the most popular non-informative priors available in the literature in multi-parameter cases. In case of the single parameter problem, it is equivalent to Jeffrey's prior. Berger and Bernardo (1992) further extended the work and proposed an algorithm to derive the reference prior. The same algorithm is followed to establish the priors in this chapter.

Let us consider the different ordered grouping of unknown parameters $\Theta = (\beta, \alpha, \lambda)$ based on their inferential importance. Four ordered grouping of parameters $\{\beta, \alpha, \lambda\}$, $\{(\beta, \alpha), \lambda\}$, $\{\alpha, \beta, \lambda\}$ and $\{\lambda, (\beta, \alpha)\}$ are being considered for inferential analysis. It is essential to order the parameters by descending interest before deriving the reference priors. For example,

consider the ordered group of parameters $\{\beta, \alpha, \lambda\}$; here the parameter β is of the highest importance while the parameter λ is of the least importance for the inferential purpose of the researchers. Similarly, for the set $\{\lambda, (\beta, \alpha)\}$, the parameter λ is the most important while (β, α) are of equal but least importance for the same. In case, if a researcher is more interested in the shape of the distribution being considered for time-to-event, (s)he can carry out the Bayesian analysis with reference prior derived for ordered group $\{\beta, \alpha, \lambda\}$ of parameters. The objective Bayesian approach with reference prior provides more flexibility to carry out the analysis based on in terms of the inferential interest on parameters. The formulation of reference prior has been done on the same line as given by Berger and Bernardo (1992). To derive the reference priors for the five considered ordered groups of parameters based on their inferential importance, the Fisher information matrix is obtained considering the complete data log-likelihood function (2.6) which is given below

$$I(\beta, \alpha, \lambda) = \begin{bmatrix} \frac{n}{\alpha^2} T(\alpha) & \frac{n}{\alpha\beta} (1 + r_1 - \ln \alpha) & 0 \\ \frac{n}{\alpha\beta} (1 + r_1 - \ln \alpha) & \frac{n}{\alpha^2} & 0 \\ 0 & 0 & \frac{n_{nr}}{\lambda^2} \end{bmatrix}, \quad (2.8)$$

where $T(\alpha) = 1 + 2r_1 + r_2 - 2(r_1 + 1) \ln(\alpha) + (\ln(\alpha))^2$ and $r_q = \int_0^\infty (\ln(z))^q \exp\{-z\} dz$.

(a) The reference prior for $\{\beta, \alpha, \lambda\}$

In this case, the number of groups is $m = 3$ with the group sizes $g_1 = g_2 = g_3 = 1$ with $\Theta_{(1)} = \beta$, $\Theta_{(2)} = \alpha$ and $\Theta_{(3)} = \lambda$. Also $\Theta_{[1]} = \alpha$, $\Theta_{[2]} = (\beta, \alpha)$ and $\Theta_{[3]} = (\beta, \alpha, \lambda)$ along with $\Theta_{[\sim 0]} = (\beta, \alpha, \lambda)$, $\Theta_{[\sim 1]} = (\beta, \lambda)$ and $\Theta_{[\sim 2]} = \lambda$, $\Theta_{[0]}$ and $\Theta_{[\sim 3]}$ are vacuous.

The variance-covariance matrix for $\{\beta, \alpha, \lambda\}$ is

$$S = \begin{bmatrix} \frac{\beta^2}{n(r_2 - r_1^2)} & -\frac{\alpha\beta}{n} \frac{(1+r_1-\ln\alpha)}{(r_2-r_1^2)} & 0 \\ -\frac{\alpha\beta}{n} \frac{(1+r_1-\ln\alpha)}{(r_2-r_1^2)} & \frac{\alpha^2}{n} \frac{T(\alpha)}{(r_2-r_1^2)} & 0 \\ 0 & 0 & \frac{\lambda^2}{n_{nr}} \end{bmatrix}.$$

Hence, $S_1 = \frac{\beta^2}{n(r_2 - r_1^2)}$, $H_1 = \frac{n(r_2 - r_1^2)}{\beta^2} = h_1$,

$$S_2 = \begin{bmatrix} \frac{\beta^2}{n(r_2 - r_1^2)} & -\frac{\alpha\beta}{n} \frac{(1+r_1 - \ln \alpha)}{(r_2 - r_1^2)} \\ -\frac{\alpha\beta}{n} \frac{(1+r_1 - \ln \alpha)}{(r_2 - r_1^2)} & \frac{\alpha^2}{n} \frac{T(\alpha)}{(r_2 - r_1^2)} \end{bmatrix},$$

This implies

$$H_2 = S_2^{-1} = \begin{bmatrix} \frac{n}{\beta^2} T(\alpha) & \frac{n}{\alpha\beta} (1 + r_1 - \ln \alpha) \\ \frac{n}{\alpha\beta} (1 + r_1 - \ln \alpha) & \frac{n}{\alpha^2} \end{bmatrix},$$

$h_2 = \frac{n}{\alpha^2}$ and $S_3 = S$, $H_3 = I(\beta, \alpha, \lambda)$ and $h_3 = \frac{n_{nr}}{\lambda^2}$. Now consider the compact sets of parametric space i.e. $\Theta_l = \{(\beta, \alpha, \lambda) \mid a_{1l} < \beta < b_{1l}, a_{2l} < \alpha < b_{2l}, a_{3l} < \lambda < b_{3l}\}$; where $a_{1l}, a_{2l}, a_{3l} \rightarrow 0$ and $a_{1l}, a_{2l}, a_{3l} \rightarrow \infty$ as $l \rightarrow \infty$.

Now we have to compute the following integral $\int_{\Theta(\Theta_{[j-1]})} |h_j|^{1/2} d\Theta_{(j)}$ for $j = 1, 2$ and 3 as

$$\begin{aligned} \int_{a_{1l}}^{b_{1l}} |h_1|^{1/2} d\beta &= \{n(r_2 - r_1^2)\}^{1/2} (\log b_{1l} - \log a_{1l}); \\ \int_{a_{2l}}^{b_{2l}} |h_2|^{1/2} d\alpha &= n^{1/2} (\log b_{2l} - \log a_{2l}); \\ \int_{a_{3l}}^{b_{3l}} |h_3|^{1/2} d\lambda &= n_{nr}^{1/2} (\log b_{3l} - \log a_{3l}). \end{aligned}$$

Since h_1 , h_2 and h_3 satisfies the lemma 1 of Berger and Bernardo (1992), so the reference prior for $\{\beta, \alpha, \lambda\}$ is given by

$$\pi_1(\Theta) \propto (\alpha\beta\lambda)^{-1}. \quad (2.9)$$

The reference prior given in (2.9) is same for grouping $\{(\beta, \alpha), \lambda\}$ and can be obtained by following the same procedure as followed for $\{\beta, \alpha, \lambda\}$.

(b) The reference prior for $\{\lambda, (\beta, \alpha)\}$

The parameter λ is orthogonal to nuisance parameters (β, α) . In this case, we have

$$h_1 = \frac{|I(\beta, \alpha, \lambda)|}{|I_1(\beta, \alpha)|} = \frac{n_{nr}}{\lambda^2}; \quad h_2 = |I_1(\beta, \alpha)| = \left(\frac{n}{\alpha\beta}\right)^2 (r_2 - r_1^2)$$

$$\implies \int_{a_{1l}}^{b_{1l}} \int_{a_{2l}}^{b_{2l}} |h_2|^{1/2} d\alpha d\beta = Q_1; \quad \text{where } Q_1 \text{ is a constant.}$$

The conditional prior of (β, α) is given by

$$\begin{aligned} \pi_2^l((\beta, \alpha) | \lambda) &\propto \frac{|h_2|^{1/2}}{\int_{a_{1l}}^{b_{1l}} \int_{a_{2l}}^{b_{2l}} |h_2|^{1/2} d\alpha d\beta} \\ &\propto \frac{1}{\alpha\beta}. \end{aligned}$$

$$\begin{aligned} E^l\{\log h_1 | \beta, \alpha\} &= \int_{a_{1l}}^{b_{1l}} \int_{a_{2l}}^{b_{2l}} \log\left(\frac{n_{nr}}{\lambda^2}\right) \pi_2^l(\beta, \alpha | \lambda) d\alpha d\beta \\ &= \log\left(\frac{n_{nr}}{\lambda^2}\right). \end{aligned}$$

The marginal prior of λ is

$$\pi_1^l(\lambda) \propto \exp\left(\frac{1}{2} E^l\{\log h_1 | \beta, \alpha\}\right) = \frac{n_{nr}^{1/2}}{\lambda}.$$

The reference prior for $\{\lambda, (\beta, \alpha)\}$ is given by

$$\pi_1(\Theta) = \lim_{l \rightarrow \infty} \frac{\pi_2^l(\beta, \alpha | \lambda) \pi_1^l(\lambda)}{\pi_2^l(\beta_\circ, \alpha_\circ | \lambda_\circ) \pi_1^l(\lambda_\circ)} \propto (\alpha\beta\lambda)^{-1}, \quad (2.10)$$

where $\beta_\circ, \alpha_\circ, \lambda_\circ$ are the interior points of interval $(0, \infty)$.

(c) The reference prior for $\{\alpha, \beta, \lambda\}$

In this case, the number of groups is $m = 3$ with the group sizes $g_1 = g_2 = g_3 = 1$ with $\Theta_{(1)} = \alpha$, $\Theta_{(2)} = \beta$ and $\Theta_{(3)} = \lambda$. Also $\Theta_{[1]} = \alpha$, $\Theta_{[2]} = (\alpha, \beta)$ and $\Theta_{[3]} = (\alpha, \beta, \lambda)$ along with $\Theta_{[\sim 0]} = (\alpha, \beta, \lambda)$, $\Theta_{[\sim 1]} = (\beta, \lambda)$ and $\Theta_{[\sim 2]} = \lambda$, $\Theta_{[0]}$ and $\Theta_{[\sim 3]}$ are vacuous.

The variance-covariance matrix for $\{\alpha, \beta, \lambda\}$ is

$$S_1 = \begin{bmatrix} \frac{\alpha^2}{n} \frac{T(\alpha)}{(r_2 - r_1^2)} & -\frac{\alpha\beta}{n} \frac{(1+r_1-\ln\alpha)}{(r_2 - r_1^2)} & 0 \\ -\frac{\alpha\beta}{n} \frac{(1+r_1-\ln\alpha)}{(r_2 - r_1^2)} & \frac{\beta^2}{n(r_2 - r_1^2)} & 0 \\ 0 & 0 & \frac{\lambda^2}{n_{nr}} \end{bmatrix}.$$

Hence $S_1 = \frac{\alpha^2}{n} \frac{T(\alpha)}{(r_2 - r_1^2)}$, $H_1 = \frac{n(r_2 - r_1^2)}{\alpha^2 T(\alpha)} = h_1$,

$$S_2 = \begin{bmatrix} \frac{\alpha^2}{n} \frac{T(\alpha)}{(r_2 - r_1^2)} & -\frac{\alpha\beta}{n} \frac{(1+r_1-\ln\alpha)}{(r_2 - r_1^2)} \\ -\frac{\alpha\beta}{n} \frac{(1+r_1-\ln\alpha)}{(r_2 - r_1^2)} & \frac{\beta^2}{n(r_2 - r_1^2)} \end{bmatrix}, H_2 = S_2^{-1} = \begin{bmatrix} \frac{n}{\alpha^2} & \frac{n}{\alpha\beta}(1 + r_1 - \ln\alpha) \\ \frac{n}{\alpha\beta}(1 + r_1 - \ln\alpha) & \frac{n}{\beta^2} T(\alpha) \end{bmatrix}';$$

$$h_2 = \frac{n}{\beta^2} T(\alpha) \text{ and } S_3 = S_1, H_3 = I(\alpha, \beta, \lambda) \text{ and } h_3 = \frac{n_{nr}}{\lambda^2}.$$

Now consider the compact sets of parametric space i.e. $\Theta_l = \{(\alpha, \beta, \lambda) \mid a_{1l} < \alpha < b_{1l}, a_{2l} < \beta < b_{2l}, a_{3l} < \lambda < b_{3l}\}$; where $a_{1l}, a_{2l}, a_{3l} \rightarrow 0$ and $a_{1l}, a_{2l}, a_{3l} \rightarrow \infty$ as $l \rightarrow \infty$.

Now we have to compute the following integral $\int_{\Theta(\Theta_{[j-1]})} |h_j|^{1/2} d\Theta_{(j)}$; for $j = 1, 2$ and 3 as

$$\begin{aligned} \int_{a_{1l}}^{b_{1l}} |h_1|^{1/2} d\alpha &= Q_2; \text{ where } Q_2 \text{ is a constant;} \\ \int_{a_{2l}}^{b_{2l}} |h_2|^{1/2} d\beta &= n^{1/2} (\log b_{2l} - \log a_{2l}); \\ \int_{a_{3l}}^{b_{3l}} |h_3|^{1/2} d\lambda &= n_{nr}^{1/2} (\log b_{3l} - \log a_{3l}). \end{aligned}$$

Since h_1, h_2 and h_3 satisfies the lemma 1 of Berger and Bernardo (1992), so the reference prior for $\{\alpha, \beta, \lambda\}$ is given by

$$\pi_2(\Theta) \propto T(\beta)^{-1/2} (\alpha\beta\lambda)^{-1}. \quad (2.11)$$

Since three of the reference priors turned out to be identical, so the number of different reference priors for the four ordered groups is only two. It is apparent that all the reference priors are proportional to the power of $T(\alpha)$ and $(\alpha\beta\lambda)^{-1}$, hence a general form for reference prior can be written. The subsequent theorem presents all the derived reference priors in a

general form.

Theorem 2.3.1. The generalized form of reference prior for an ordered group of unknown parameters is

$$\pi(\Theta) \propto (\alpha\beta\lambda)^{-1} T(\alpha)^c, \quad c \in \mathfrak{R}. \quad (2.12)$$

The reference prior $\pi(\Theta)$ defined in (2.12) takes different form of priors according to the real number c , where $c \in \{-\frac{1}{2}, 0\}$. All the considered reference priors for different ordered grouping of parameters are given in Table 2.1. As we can see from Table 2.1, for the first three considered ordered group of parameters, the derived reference priors are identical. It shows that the reference prior $\pi_1(\Theta)$ is robust to the ordered groups $\{\beta, \alpha, \lambda\}$, $\{(\beta, \alpha), \lambda\}$ and $\{\lambda, (\beta, \alpha)\}$ of parameters.

The generalized form of posterior density using generalized form of reference prior in (2.12) and complete likelihood in (2.5) is given by

$$\begin{aligned} \Pi(\Theta) &\propto L(\Theta | \underline{Y}) \cdot \pi(\Theta) \\ &\propto T(\alpha)^c \alpha^{n-1} \beta^{n-1} \lambda^{n_{nr}-1} \left(\prod_{i=1}^{n_r} t_i^{\beta-1} \right) \left(\prod_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta-1} \right) \left(\prod_{i=n_{nr}+1}^n t_{ri}^{*\beta-1} \right) \\ &\quad \exp \left\{ -\alpha \left(\sum_{i=1}^{n_r} t_i^\beta + \sum_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta} + \sum_{i=n_{nr}+1}^n t_{ri}^{*\beta} \right) \right\} \exp \left\{ -\lambda \left(\sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^* \right) \right\}. \end{aligned} \quad (2.13)$$

Table 2.1: All reference prior for the parameter Θ .

Grouping order	Reference prior	c
$\{\beta, \alpha, \lambda\}$	$\pi_1(\Theta)$	0
$\{(\beta, \alpha), \lambda\}$	$\pi_1(\Theta)$	0
$\{\lambda, (\beta, \alpha)\}$	$\pi_1(\Theta)$	0
$\{\alpha, \beta, \lambda\}$	$\pi_2(\Theta)$	-1/2

2.3.2 Posterior Propriety and Estimation

The reference priors given in (2.12) come out to be improper and the posterior density given in (2.13) is obtained using these improper reference priors. Hence before drawing any inference using this posterior distribution, we need to establish its propriety. In this subsection, we have proved that the posterior density obtained under improper prior (2.12) is proper for selected values of c . The following theorem establishes the propriety of the posterior density:

Theorem 2.3.2. For $n > 0$, the posterior density $\Pi(\Theta)$ is proper under the reference priors $\pi_1(\Theta)$ and $\pi_2(\Theta)$.

Proof. Let us consider the reference prior given by (2.12) for the case when $c = 0$. Also let the observed data be $y_i = (t_i, t_{li}^*, t_{ri}^*)$. For reference prior $\pi_1(\Theta)$, we have to show that the integral $\int \Pi(\Theta) d\Theta < \infty$.

Consider the integral

$$\begin{aligned} I_1 &= \iiint_{\Theta} \beta^{n-1} \alpha^{n-1} \lambda^{n_{nr}-1} \left(\prod_{i=1}^n y_i^{\beta-1} \right) \exp \left(-\alpha \sum_{i=1}^n y_i^\beta \right) \\ &\quad \exp \left\{ -\lambda \left(\sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^* \right) \right\} d\lambda d\alpha d\beta \\ &= A \int_0^\infty \left[\int_0^\infty \alpha^{n-1} \exp \left(-\alpha \sum_{i=1}^n y_i^\beta \right) d\alpha \right] \beta^{n-1} \left(\prod_{i=1}^n y_i^{\beta-1} \right) d\alpha \\ &= C \int_0^\infty \beta^{n-1} \frac{\prod_{i=1}^n y_i^\beta}{(\sum_{i=1}^n y_i^\beta)^n} d\beta \leq C \int_0^\infty \beta^{n-1} \prod_{i=1}^n \left(\frac{y_i}{y_{(n)}} \right)^\beta d\beta < \infty, \end{aligned}$$

where $y_{(n)} = \max_{1 \leq i \leq n} y_i$, $A = \Gamma(n_{nr}) \left(\sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^* \right)^{-n_{nr}}$ and $C = \frac{A \Gamma(n)}{\prod_{i=1}^n y_i}$ are constants. Hence, the posterior density $\Pi(\Theta)$ under reference prior $\pi_1(\Theta)$ is proper for $n > 0$.

Under the reference prior $\pi_2(\Theta)$, consider the form of posterior $\Pi(\Theta)$ when $c = -\frac{1}{2}$.

Now consider the integral given by,

$$I_2 = \iiint_{\Theta} T(\alpha)^{-1/2} \beta^{n-1} \alpha^{n-1} \lambda^{n_{nr}-1} \left(\prod_{i=1}^n y_i^{\beta-1} \right) \exp \left(-\alpha \sum_{i=1}^n y_i^\beta \right)$$

$$\begin{aligned}
& \exp \left\{ -\lambda \left(\sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^* \right) \right\} d\lambda d\alpha d\beta \\
& = A \int_0^\infty \left[\int_0^\infty T(\alpha)^{-1/2} \alpha^{n-1} \exp \left(-\alpha \sum_{i=1}^n y_i^\beta \right) d\alpha \right] \beta^{n-1} \left(\prod_{i=1}^n y_i^{\beta-1} \right) d\beta \\
& = K \int_0^\infty \beta^{n-1} \frac{\prod_{i=1}^n y_i^\beta}{(\sum_{i=1}^n y_i^\beta)^n} d\beta \leq K \int_0^\infty \beta^{n-1} \prod_{i=1}^n \left(\frac{y_i}{y_{(n)}} \right)^\beta d\beta < \infty,
\end{aligned}$$

where $y_{(n)} = \max_{1 \leq i \leq n} y_i$, $A = \Gamma(n_{nr}) \left(\sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^* \right)^{-n_{nr}}$.

If $Q = \int_0^\infty T(\alpha)^{-1/2} \alpha^{n-1} \exp \left(-\alpha \sum_{i=1}^n y_i^\beta \right) d\alpha < \infty$, then $K = AQ$ is constant because

$$\lim_{\alpha \rightarrow \infty} |T(\alpha)^{-1/2}| = 0.$$

Hence the posterior density $\Pi(\Theta)$ under reference prior $\pi_2(\Theta)$ is proper for $n > 0$. \square

Using (2.13), the full conditional of β , α and λ can be written as

$$\begin{aligned}
p_1(\beta | \alpha) & \propto \beta^{n-1} \left(\prod_{i=1}^{n_r} t_i^{\beta-1} \right) \left(\prod_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta-1} \right) \left(\prod_{i=n_{nr}+1}^n t_{ri}^{*\beta-1} \right) \\
& \exp \left\{ -\alpha \left(\sum_{i=1}^{n_r} t_i^\beta + \sum_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta} + \sum_{i=n_{nr}+1}^n t_{ri}^{*\beta} \right) \right\}, \tag{2.14}
\end{aligned}$$

$$p_2(\alpha | \beta) \propto T(\alpha)^c \alpha^{n-1} \exp \left\{ -\alpha \left(\sum_{i=1}^{n_r} t_i^\beta + \sum_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta} + \sum_{i=n_{nr}+1}^n t_{ri}^{*\beta} \right) \right\}, \tag{2.15}$$

$$p_3(\lambda) \propto \lambda^{n_{nr}-1} \exp \left\{ -\lambda \left(\sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^* \right) \right\}. \tag{2.16}$$

The point and interval estimates under the Bayesian paradigm are obtained by generating samples from the conditional posterior distributions given in (2.14), (2.15) and (2.16). The samples on t_{li}^* , w_{li}^* and t_{ri}^* can be generated using expressions below

$$t_{li}^* = \left[-\frac{1}{\alpha} \ln \left\{ 1 - u_i \left(1 - \exp \left\{ -\alpha s_i^\beta \right\} \right) \right\} \right]^{1/\beta}; \quad i = n_r + 1, \dots, n_{nr} \tag{2.17}$$

$$w_{li}^* = -\frac{1}{\lambda} \ln \left[1 - u_i \left(1 - \exp \left\{ -\lambda (s_i - t_{li}^*) \right\} \right) \right]; \quad i = n_r + 1, \dots, n_{nr} \tag{2.18}$$

$$\text{and } t_{ri}^* = \left[-\frac{1}{\alpha} \ln \left\{ -u_i \exp\{-\alpha s_i^\beta\} \right\} \right]^{1/\beta}; \quad i = n_{nr} + 1, \dots, n, \quad (2.19)$$

where $u \sim \mathcal{U}(0, 1)$. These expressions are obtained through transformations based on conditional densities of these latent variables introduced in Section 2.2.

For the initial values of parameters and t_{li}^* , w_{li}^* and t_{ri}^* , the posterior samples can be generated using MCMC techniques as discussed below:

1. To obtain a sample for β from (2.14), we use M-H algorithm (Metropolis and Ulam 1949) by taking normal distribution as a proposal density such as the candidate point $\beta^t \sim \mathcal{N}(\hat{\beta}, Var(\hat{\beta}))$; where $\hat{\beta}$ and $Var(\hat{\beta})$ are ML estimate and variance of the ML estimate of α , respectively.
2. For reference prior $\pi_1(\Theta)$, c takes value 0, so that posterior samples for parameter α can be directly generated from $\mathcal{G}\left(n, \sum_{i=1}^{n_r} t_i^\beta + \sum_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta} + \sum_{i=n_{nr}+1}^n t_{ri}^{*\beta}\right)$. In case of prior $\pi_2(\Theta)$, c takes value $-1/2$. In this case, to generate random sample for α , we use M-H algorithm by considering Gamma distribution as the proposal density so that the candidate point $\alpha^t \sim \mathcal{G}\left(n, \sum_{i=1}^{n_r} t_i^\beta + \sum_{i=n_r+1}^{n_{nr}} t_{li}^{*\beta} + \sum_{i=n_{nr}+1}^n t_{ri}^{*\beta}\right)$.
3. In a similar way, we can directly generate posterior samples for λ from

$$\mathcal{G}\left(n_{nr}, \sum_{i=1}^{n_r} (s_i - t_i) + \sum_{i=n_r+1}^{n_{nr}} w_{li}^*\right).$$

After discarding the burn-in from generated chains the stationarity and convergence can be checked through cumsum plot, auto correlation function (*ACF*), and Gelman and Rubin's test statistics (Gelman and Rubin 1992). These samples from conditional posteriors can be used to calculate the point and interval estimates in the Bayesian framework.

As we know, the choice of loss function plays an important role in Bayesian estimation along with the prior elicitation. The loss function can be defined as the loss incurred in estimating the true value of the parameter. It is a function of the parameter and the estimator used to estimate it. The most commonly used loss function is the *SELF* which is a symmetric loss function that gives equal weight to underestimation and overestimation.

Suppose one wants to obtain the Bayesian estimator of some function of Θ , say, $g(\Theta)$ under SELF, then SELF is defined as

$$L_{SE}(\hat{g}(\Theta), g(\Theta)) = (\hat{g}(\Theta) - g(\Theta))^2,$$

where $\hat{g}(\Theta)$ is an estimator of $g(\Theta)$.

The Bayesian estimator of $g(\Theta)$ under SELF is given by

$$\hat{g}(\Theta)^{SE} = \frac{\int g(\Theta)\Pi(\Theta | \underline{t}) d\Theta}{\int \Pi(\Theta | \underline{t}) d\Theta},$$

where $\Pi(\Theta | \underline{t})$ is the posterior density of Θ .

2.4 Simulation Study

In this section, a simulation study is performed to illustrate the empirical behaviour of the Bayesian estimators obtained under reference priors derived in Section 2.3. Consider there are n individuals in the recall-based study. The time-to-event (T) is generated from the Weibull distribution. We took arbitrary value of the model parameters for the illustration purpose as $\beta = 1.15$ and $\alpha = 0.85$ and reported the results in the simulation table.

There may be several choices for the monitoring time distribution as per time/cost/labour constraints in a recall-based study. And accordingly, one can plan for monitoring time distribution as well as its domain. Here, we use uniform and exponential patterns to generate observations on monitoring time (S). Uniform monitoring time ensures consistency in monitoring a study or process at predetermined intervals. On the other hand, in some surveys, there may be more frequent visits at the beginning and then a decrease in frequency over time. To account for this, we use an exponential distribution to generate monitoring times (S).

Each individual in the study is monitored at a single time point, and the status of the event is recorded. The non-recall probability parameter is varied to maintain the proportion of non-recall and right-censored observations in the simulated data based on the sample size,

n . Additionally, we have maintained the proportions of non-recall and right-censored data as per the monitoring patterns considered, as described below:

- Under uniform monitoring, S is generated from $\mathcal{U}(0, 4.2)$. For $\lambda = 0.10$, the proportion of non-recall and right-censored observations come out to be approximately 0.12 and 0.26, respectively as per choice of n . Similarly, for $\lambda = 0.20$, the proportion of non-recall and right-censored observations come out to be approximately 0.20 and 0.26, respectively as per choice of n . The randomization in the proportion of non-recall observations diffuses any biases in the estimates.
- Under exponential monitoring pattern, S is generated from $\mathcal{E}(0.3)$. For $\lambda = 0.10$, the proportion of non-recall and right-censored observations come out to be approximately 0.19 and 0.25, respectively as per choice of n . Similarly, for $\lambda = 0.15$, the proportion of non-recall and right-censored observations come out to be approximately 0.25 and 0.26, respectively as per choice of n .

Under the above scenarios, the observations are generated for varying sample sizes $n = \{30, 50, 80\}$. Further, for each sample size, three categories of data are observed. For the simulated data, the Bayesian estimates are obtained under reference priors $\pi_1(\Theta)$ and $\pi_2(\Theta)$. To assess the performance of obtained point estimators, mean square error (MSE) and absolute bias (AB) are calculated. For the interval estimators, the average length (AL) and coverage probability (CP) are reported for 95% HPD interval.

Under uniform monitoring, the Bayes estimates of the parameters along with MSE and AB are reported in Table 2.2 for $\pi_1(\Theta)$ and $\pi_2(\Theta)$. Similarly, under exponential monitoring pattern, the results on same measures are reported in Table 2.3 for $\pi_1(\Theta)$ $\pi_2(\Theta)$. AL and CP are reported in Table 2.4 and Table 2.5 for uniform and exponential monitoring patterns, respectively.

From the simulated results, it can be seen that MSE , AB and AL decrease as the sample size increases. In other words, all the estimators perform consistently. Also, as the proportion of non-recall observation becomes large, the performance of estimators deteriorates in the form of their MSE , AB and AB irrespective of the monitoring patterns. Further, the CPs are found satisfactory at the nominal level 0.95.

Table 2.2: Bayes estimates of parameters along with their AB and MSE for uniform monitoring.

λ	n	$\pi_1(\Theta)$			$\pi_2(\Theta)$			
		Estimate	AB	MSE	Estimate	AB	MSE	
0.10	30	β	1.1872	0.0372	0.0170	1.1848	0.0348	0.0166
		α	0.8647	0.0147	0.0367	0.8763	0.0263	0.0366
		λ	0.1386	0.0386	0.0047	0.1386	0.0386	0.0047
	50	β	1.1773	0.0273	0.0082	1.1760	0.0260	0.0081
		α	0.8440	0.0060	0.0256	0.8513	0.0013	0.0253
		λ	0.1284	0.0284	0.0023	0.1284	0.0284	0.0023
	80	β	1.1701	0.0201	0.0048	1.1694	0.0194	0.0048
		α	0.8392	0.0108	0.0148	0.8437	0.0063	0.0146
		λ	0.1239	0.0239	0.0015	0.1239	0.0239	0.0015
0.20	30	β	1.2134	0.0634	0.0250	1.2100	0.0600	0.0242
		α	0.8253	0.0247	0.0434	0.8375	0.0125	0.0422
		λ	0.2354	0.0354	0.0102	0.2351	0.0351	0.0101
	50	β	1.1884	0.0384	0.0106	1.1866	0.0366	0.0104
		α	0.8082	0.0418	0.0223	0.8160	0.0340	0.0216
		λ	0.2169	0.0169	0.0056	0.2167	0.0167	0.0056
	80	β	1.1835	0.0335	0.0068	1.1824	0.0324	0.0067
		α	0.8053	0.0447	0.0153	0.8103	0.0397	0.0148
		λ	0.2195	0.0195	0.0041	0.2195	0.0195	0.0041

Table 2.3: Bayes estimates of parameters along with their *AB* and *MSE* for exponential monitoring.

λ	n	$\pi_1(\Theta)$			$\pi_2(\Theta)$			
		Estimate	AB	MSE	Estimate	AB	MSE	
0.10	30	β	1.1971	0.0471	0.0182	1.1944	0.0444	0.0178
		α	0.8432	0.0068	0.0483	0.8564	0.0064	0.0475
		λ	0.1162	0.0162	0.0026	0.1161	0.0161	0.0026
	50	β	1.1795	0.0295	0.0098	1.1780	0.0280	0.0097
		α	0.8360	0.0140	0.0257	0.8439	0.0061	0.0253
		λ	0.1084	0.0084	0.0013	0.1083	0.0083	0.0013
	80	β	1.1702	0.0202	0.0062	1.1694	0.0194	0.0061
		α	0.8241	0.0259	0.0162	0.8293	0.0207	0.0159
		λ	0.1062	0.0062	0.0009	0.1062	0.0062	0.0009
0.15	30	β	1.2053	0.0553	0.0220	1.2022	0.0522	0.0214
		α	0.8366	0.0134	0.0465	0.8505	0.0005	0.0454
		λ	0.1695	0.0195	0.0055	0.1693	0.0193	0.0055
	50	β	1.1796	0.0296	0.0103	1.1781	0.0281	0.0102
		α	0.8249	0.0251	0.0256	0.8335	0.0165	0.0249
		λ	0.1696	0.0196	0.0034	0.1694	0.0194	0.0034
	80	β	1.1718	0.0218	0.0053	1.1707	0.0207	0.0052
		α	0.8125	0.0375	0.0174	0.8181	0.0319	0.0170
		λ	0.1616	0.0116	0.0017	0.1615	0.0115	0.0017

Table 2.4: AL and CP of 95% HPD interval estimates for uniform monitoring.

λ	n	$\pi_1(\Theta)$		$\pi_2(\Theta)$		
		AL	CP	AL	CP	
0.10	30	β	0.6029	0.9813	0.6004	0.9859
		α	0.7480	0.9645	0.7467	0.9623
		λ	0.2381	0.9667	0.2373	0.9678
0.10	50	β	0.4595	0.9887	0.4575	0.9860
		α	0.5700	0.9253	0.5695	0.9323
		λ	0.1833	0.9832	0.1829	0.9659
0.10	80	β	0.3567	0.9889	0.3561	0.9859
		α	0.4489	0.9320	0.4477	0.9345
		λ	0.1437	0.9623	0.1435	0.9750
0.20	30	β	0.6544	0.9835	0.6517	0.9810
		α	0.7427	0.9225	0.7420	0.9310
		λ	0.3433	0.9520	0.3426	0.9535
0.20	50	β	0.4858	0.9860	0.4846	0.9875
		α	0.5642	0.9325	0.5634	0.9445
		λ	0.2565	0.9460	0.2560	0.9450
0.20	80	β	0.3802	0.9833	0.3790	0.9842
		α	0.4468	0.9045	0.4461	0.9145
		λ	0.2078	0.9267	0.2075	0.9250

Table 2.5: AL and CP of 95% HPD interval estimates for exponential monitoring.

λ	n	$\pi_1(\Theta)$		$\pi_2(\Theta)$		
		AL	CP	AL	CP	
0.10	30	β	0.6370	0.9850	0.6333	0.9883
		α	0.7814	0.9433	0.7805	0.9550
		λ	0.1781	0.9625	0.1776	0.9633
0.10	50	β	0.4754	0.9850	0.4739	0.9850
		α	0.5998	0.9383	0.5982	0.9433
		λ	0.1341	0.9513	0.1336	0.9533
0.10	80	β	0.3685	0.9850	0.3674	0.9800
		α	0.4690	0.9150	0.4679	0.9250
		λ	0.1055	0.9417	0.1053	0.9450
0.15	30	β	0.6642	0.9887	0.6611	0.9850
		α	0.7996	0.9335	0.7985	0.9420
		λ	0.2375	0.9433	0.2371	0.9450
0.15	50	β	0.4987	0.9817	0.4969	0.9833
		α	0.6162	0.9433	0.6155	0.9417
		λ	0.1866	0.9250	0.1863	0.9275
0.15	80	β	0.3867	0.9866	0.3851	0.9815
		α	0.4826	0.9260	0.4819	0.9310
		λ	0.1428	0.9433	0.1426	0.9450

2.5 Real Data Application

In this section, we performed real data analysis to illustrate the proposed methodologies. The data is taken from the fourth round of *NFHS* conducted in 2015–2016. *NFHS* is a large-scale, multi-dimensional survey being conducted on a sample of household representatives in India. It is managed by the International Institute for Population Sciences, Mumbai, India in collaboration with ICF, Calverton, Maryland, USA and the East-West Center, Honolulu, Hawaii, USA. In the *NFHS* survey, data related to population, health and nutrition mainly on children and women are collected.

The data used for the study is related to the duration of breastfeeding of the recent child of a mother. In *NFHS-IV*, the duration of breastfeeding is defined as the length of time a child is breastfed from birth until they are completely weaned up to 2 years or beyond. We considered births in last 60 months preceding the survey and the information was obtained on the duration of any breastfeeding for children by asking mothers a series of questions about their child's feeding practices, including whether the child was ever breastfed, when they were first breastfed, and whether they were currently being breastfed. Based on the responses to these questions, the duration of breastfeeding was calculated for each child in months

Before analysing the dataset, we check for the model compatibility (i.e. likelihood and the prior) based on the *PPV*.

2.5.1 Model Compatibility based on *PPV*

Model compatibility is an important concept that provides an assurance that the chosen model is right for the data at hand. Before drawing any inference from the data it becomes necessary to check the compatibility of the data with the chosen model. In case of poor resemblance with the data, an alternative model should be selected. There are some graphical as well as mathematical approaches for checking the compatibility of data with chosen models. Also, there are simulation-based approaches to assess the compatibility of the model. The classical approach uses tail area probabilities based on some defined statis-

tic. For details one may refer to one of them which is based on modified χ^2 and proposed by Hope (1968).

For obtaining the PPV under recall-based setting, we define the chi-square statistic as

$$\chi^2 = \sum_{i=1}^{n_r} \frac{(t_i - E(t_i|\Theta))^2}{Var(t_i|\Theta)} + \sum_{i=n_r+1}^{n_{nr}} \frac{(t_{li}^* - E(t_{li}^*|\Theta))^2}{Var(t_{li}^*|\Theta)} + \sum_{i=n_{nr}+1}^n \frac{(t_{ri}^* - E(t_{ri}^*|\Theta))^2}{Var(t_{ri}^*|\Theta)},$$

where the symbols $E(\cdot)$ and $Var(\cdot)$ are used for defining expectation and variance respectively. Since, non-recall and right-censored observations poses incompleteness in the real data, we first complete the data by replacing these observations with maximum of their censoring times and predictive mean times as suggested by Gelfand and Ghosh (1998).

The PPV based on the defined chi-square statistic is calculated as

$$PPV = \int_{\Theta} P[\chi_2^2 > \chi_1^2 | f, \Theta] \Pi(\Theta | \underline{t}) d\Theta,$$

where χ_1^2 and χ_2^2 are the calculated values of χ^2 for the observed and the predicted datasets respectively. The calculation of PPV consists of two steps described further. In the first step, we draw samples on Θ from $\Pi(\Theta | \underline{t})$ using algorithm in subsection 2.3.2 and calculate the value of χ_1^2 for the given data set. The samples on latent variables are obtained by using expressions (2.17) to (2.19). In the second step, we extract the predictive data set each of the same sizes as of given data from the model f using the simulated Θ and then calculate χ_2^2 using these predictive datasets. We then calculate $P[\chi_2^2 > \chi_1^2 | f, \Theta]$ as the number of times χ_2^2 exceeds χ_1^2 . These two steps are calculated with different simulated Θ and PPV is estimated as the posterior expectation of $P[\chi_2^2 > \chi_1^2 | f, \Theta]$.

2.5.2 Breastfeeding Data

We aim to estimate the duration of breastfeeding for recent children using this data. The mothers are visited at any particular age and information on variables related to their health, socio-economic characteristics and children are collected. The available variables of our interest are the birth of the child, the status of breastfeeding and the duration of breastfeeding. Based on the question of whether they are still breastfeeding or not, we

categorize them into censored and non-censored categories. Further, if they are under the non-censored category they are asked to recall the duration of breastfeeding. The mother's responses related to breastfeeding are categorized into six categories: (i) mothers who have breastfed at some point in the past but are not currently doing so, (ii) mothers who are still breastfeeding their child, (iii) mothers who have never breastfed their child, (iv) mothers who breastfed until the child died, (v) mothers who breastfed inconsistently (not-regularly), and finally (vi) mothers who don't know the duration.

We discard the category (iii) and (v) from the original sample. The category (vi) is considered non-recall. If the observation is recalled exactly, we get the exact duration of the breastfeeding. 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.

The *NFHS-IV* data consists of 178,291 birth observations that occurred before 60 months of the survey. For our analysis purpose, we have filtered the data on state level and considered 3,176 birth observations in Arunachal Pradesh state. Among these, there are 901 exact recall observations, 131 non-recall observations, and 2,144 right-censored observations. To analyse the data, we have scaled the observations by dividing them by 24. The *PPV*'s calculated using the method discussed in subsection 2.5.1 under priors $\pi_1(\Theta)$ and $\pi_2(\Theta)$ are 0.83 and 0.93, respectively. The estimated *PPV* reflects our model compatibility for the observed dataset. The MCMC samples from the conditional posterior distribution are generated using the proposed methodologies discussed in the subsection 2.3.2. Finally, the obtained estimates of parameters of interest are reported in Table 2.6. The estimates of the mean and median duration of any breastfeeding for recent children under $\pi_1(\Theta)$ and $\pi_2(\Theta)$ are obtained under SELF. The mean duration of any breastfeeding under $\pi_1(\Theta)$ and $\pi_2(\Theta)$ comes out to be 35.29 months and 35.27 months, respectively. Similarly, the median duration of any breastfeeding under $\pi_1(\Theta)$ and $\pi_2(\Theta)$ comes out to be 33.71 months and 33.69 months, respectively. We consider all possible grouping orders of the parameters of interest and provide the posterior inference for all considered groups except associated to which posterior comes out to be improper. One can go with any suitable choice of grouping

Table 2.6: Bayes estimates of parameters of Weibull model for breastfeeding data of Arunachal Pradesh under SELF.

	$\pi_1(\Theta)$			$\pi_2(\Theta)$		
	95% HPD			95% HPD		
	Est.	Lower	Upper	Est.	Lower	Upper
β	2.1936	2.1083	2.2948	2.1939	2.0927	2.2771
α	0.3293	0.3078	0.3522	0.3297	0.3075	0.3501
λ	0.1672	0.1398	0.1967	0.1670	0.1382	0.1949

set according to one's interest or the inferential importance of the parameter. For example, in the context of duration of breastfeeding data, if someone is mainly interested in the estimates related to the median time to breastfeeding of the recent child, he/she should proceed with $\pi_1(\Theta)$ which gives more weight to the shape and scale parameters equally. Also, if someone is mainly interested in the recall behaviour of the mothers, one should choose $\pi_1(\Theta)$ again following the same argument. In case, if the interest lies in the scale parameter (i.e., variability of the duration of breastfeeding) of the distribution concerning the data, one should proceed with $\pi_2(\Theta)$ in order to obtain the Bayesian estimates more efficiently.

2.6 Conclusion

In this chapter, objective Bayesian inference is developed for recall-based time-to-event data. The Bayesian inference is carried out for a number of ordered groups of parameters according to the inferential importance of the parameters after deriving the reference priors. Also, the propriety of posterior densities obtained under reference priors is established. The Bayesian estimates of Weibull parameters as well as the non-recall parameter are obtained under two reference priors for *SELF*. To illustrate the behaviour of obtained posterior estimators, Breastfeeding data from *NFHS-IV* is analysed. Also, the compatibility of the established model under the Bayesian paradigm is validated using a posterior predictive p-value. Using the established objective Bayesian method, researchers can choose their specific ordered group of parameters with respect to the inferential choice of the parameters and carry out

the Bayesian inference. Since one of the reference priors leads to an improper posterior which can not be used to draw valid inferences for that particular group of parameters so one has to look for alternative priors.

The current study assumes a Weibull distribution for time-to-event, which may not be appropriate in some cases. For instance, if the hazard function is not monotonic, alternative distributions such as the log-normal or generalized gamma distribution may be more suitable. Additionally, if the non-recall probability is assumed to have a form other than exponential, we suggest using piecewise constant or fuzzy functions under the objective Bayesian paradigm.

Chapter 3

Objective Bayesian Analysis of Cause-Dependent Competing Risks Data using Marshall-Olkin Bivariate Generalized Lifetime Distribution

3.1 Introduction

In multifarious circumstances, one is interested in the measurement of “risks”- a risk of breakdown, a risk of becoming ill, a risk of getting infected, and so forth. A specific problem arises when several risk factors are presented altogether. In such situations, the existence of one risk always entangles the inference of another. For an individual, when the several risks compete for occurrence of the event of interest, referred to as “competing risks” in literature. Competing risks set-up can be treated as a specific case of the multivariate model where causes or risks are considered as the multiple variables of an individual.

The information available in competing risks set-up is the time-to-event and the corresponding cause of failure for the individuals under observation. The collected information may be complete or sometimes may be incomplete with respect to time and/or cause. Sinha (1986) and Lawless (2003) described the conventional competing risks model to analyse complete data. In an experiment where the researcher is unable to obtain the complete information, then adopting a censoring procedure becomes an inescapable choice to analyse competing risks data. Assuming the lifetimes of the different causes as independent exponential distributions, Kundu et al. (2003) provided the estimates under progressive type-II

censored data. In many situations, the time to failure is known but the cause of failure is masked/hidden and for such observations Miyakawa (1984), Kuo and Yang (2000) and Tomer et al. (2014) extended the competing risks model accordingly.

In real-life scenarios, considering the competing risks as independent is somewhat not feasible. The assumption of dependent competing risks results in more reliable inferences about the problem at hand. However, in the literature, it can be noticed that limited work has been done on the cause-dependent competing risks modelling. Moeschberger (1974) defines that “the theoretical lifetime of an individual failing from one cause may be correlated with the theoretical lifetime of the same individual failing from a different cause.” A bivariate exponential lifetime model has been used by Wada et al. (1996) for the analysis of competing risks data involving two dependent risk components. Wang and Ghosh (2003) studied the dependent competing risks model in the presence of covariates, whereas Lindqvist and Skogsrud (2008) analysed if preventive maintenance has been taken for the components in between the inspection.

In competing risks problems, it is challenging to derive the distribution of marginal risks. Several times, if the marginals are completely known, then the joint distribution of individual failure times is not identifiable. For a competing risks problem, Tsiatis (1975) proved that the joint distribution of failure times is not identifiable by their minimum in the multivariate case. In literature, Crowder (1991) referred to this issue as the “identifiability crisis in competing risks analysis.”

Marshall and Olkin (1967) derived the bivariate exponential distribution and demonstrated that it has both an absolutely continuous and a singular part. Further, Proschan and Sullo (1973) obtained the likelihood function for the multivariate Marshall-Olkin bivariate exponential distribution (*MOBED*) parameters under the competing risks set-up. For the Marshall-Olkin bivariate model, Basu and Ghosh (1978) established that it is identifiable for analysing competing risks data with cause dependency. Kundu and Gupta (2013) developed the inferential procedure for the Bayesian estimation of the parameters of Marshall-Olkin bivariate Weibull distribution (*MOBWD*) under informative priors. It is inevitable to perform Bayesian estimation under non-informative priors when there is no prior infor-

mation available about the parameters. For the multi-parameter case, one of the popular non-informative priors is “reference prior” suggested by Berger and Bernardo (1992).

Cai et al. (2017) analysed *MOBWD* under type-I progressive hybrid censored condition for dependent competing risks incomplete data. The Bayesian analysis performed with non-informative prior is somewhat known as objective Bayesian analysis. Guan et al. (2013) considered the objective Bayesian analysis for *MOBED* for dependent causes of failure. Recently, Feizjavadian and Hashemi (2015), Xu and Zhou (2017), and Shen and Xu (2018) performed competing risks analysis using Marshall-Olkin set-up under different types of observations.

In this chapter, a bivariate probability distribution set-up is considered, which is capable of analysing the cause-dependent competing risks data. A generalized lifetime distribution (*GLD*) family, considered by Moore and Bilikam (1978) and Ajit Chaturvedi (2006), has been taken to capture the wide range of real data characteristics. We derive the Marshall-Olkin bivariate distribution for the generalized lifetime family, which accommodates various shapes of the hazard function. So we provide a general structure to obtain the estimators for different bivariate distributions’ parameters into one framework. This makes it easier for anyone to analyse the distributions available on one hand. We use the term *MOBGLD* (Marshall-Olkin bivariate generalized lifetime distribution) for the bivariate family of distributions obtained using the same process given by Marshall and Olkin (1967).

The rest of the chapter is laid out as follows. In Section 3.2, we first establish the *MOBGLD* and then derive its important properties to utilize in the cause-dependent competing risks analysis. The procedure to obtain the Fisher information matrix and variance-covariance matrix has been given for original parameters and re-parametrized cases as well in Section 3.3. In Section 3.4, we approach through the ordered grouping of parameters in reference priors under the Bayesian method. Section 3.5 is dedicated to simulation study for the derived theory. In Section 3.6, two real data sets, the prostate cancer and diabetic retinopathy study have been taken to illustration purpose. Finally, we conclude the chapter with a summary in Section 3.7.

3.2 Model Building for Dependent Competing Risks

In this section, our objective is to propose a model which is suitable for the cause-dependent competing risks set-up. All its important properties, required for the cause-dependent modeling, have been derived.

3.2.1 Generalized Lifetime Distribution

In literature, a *GLD* family is used to achieve a wide range of applicability in a single effort. Due to this fact, a *GLD* family, introduced by Moore and Bilikam (1978) has been considered. A random variable (rv), X , is said to follow the *GLD* if its probability density function(*pdf*) $f_{GLD}(x; \beta, \theta)$ and cumulative distribution function(*cdf*) $F_{GLD}(x; \beta, \theta)$, $x \in \Re^+$, respectively defined as

$$f_{GLD}(x; \beta, \theta) = \beta \theta g'(x) [g(x)]^{\beta-1} \exp\{-\theta[g(x)]^\beta\}, \quad (3.1)$$

and

$$F_{GLD}(x; \beta, \theta) = 1 - \exp\{-\theta[g(x)]^\beta\}, \quad (3.2)$$

where β and $\theta > 0$ are the parameters. The function $g(x)$ in (3.1) is a real-valued, differentiable and strictly increasing function of x . Here, it is assumed that the inverse of $g(x)$ exists where $g(0+) = 0$ and $g(\infty) = \infty$. For the specific choices of $g(x)$, the particular cases of *GLD* can be obtained as given in Table 3.1.

Table 3.1: Distributions for specific forms of $g(x)$ and parameters for *GLD*.

$g(x)$	(β, θ)	$f_{GLD}(x; \beta, \theta)$	Distribution
x	$(1, \theta)$	$f_E(x; \theta)$	Exponential Distribution
x	$(2, \theta)$	$f_R(x; \theta)$	Rayleigh Distribution
$\ln(x/a), a > 0$	$(1, \theta)$	$f_P(x; \theta)$	Pareto Distribution
$\ln(1 + x^b), b > 0$	$(1, \theta)$	$f_B(x; \theta)$	Burr Distribution
x	(β, θ)	$f_W(x; \beta, \theta)$	Weibull Distribution

For the inferential purpose, two cases, when θ and β both are unknown and when θ is unknown but β is known, are considered. The survival function $S_{GLD}(x; \beta, \theta)$ and hazard

rate $h_{GLD}(x; \beta, \theta)$ from (3.1), respectively, given by

$$S_{GLD}(x; \beta, \theta) = \exp\{-\theta[g(x)]^\beta\}, \quad (3.3)$$

and

$$h_{GLD}(x; \beta, \theta) = \beta\theta g'(x)[g(x)]^{\beta-1}. \quad (3.4)$$

It is to be noticeable from Table 3.1 that *GLD* covers a variety of hazard functions, such as constant, increasing as well as decreasing hazard rate.

3.2.2 Marshall-Olkin Bivariate Generalized Lifetime Distribution

In this section, first we will establish Marshall-Olkin bivariate distribution, proposed by Marshall and Olkin (1967), for the *GLD* family (3.1) and then its important properties will be derived. Let U_0, U_1 and U_2 are three independent rv following the same distribution from the *GLD* family (3.1) with parameters β and $\theta_i; i = 0, 1, 2.$, i.e., $U_0 \sim GLD(\beta, \theta_0)$, $U_1 \sim GLD(\beta, \theta_1)$ and $U_2 \sim GLD(\beta, \theta_2)$. After defining $X_1 = \min(U_0, U_1)$ and $X_2 = \min(U_0, U_2)$, the bivariate set-up of rv (X_1, X_2) follows *MOBGLD* with unknown parameters $(\beta, \Theta) = (\beta, \theta_0, \theta_1, \theta_2)$. This set-up can be utilized in competing risks modeling successfully. If we assume that X_1 and X_2 denotes the observations of two components attached to a series system, then $T = \min(X_1, X_2)$ will represent the system failure time. In this approach, the correlation parameter θ_0 plays an important role which facilitates *MOBGLD* to model onto a cause-dependent competing risks problems. As well as the above discussed competing system will reduce to *MOBGLD* with independent marginals if the correlation parameter θ_0 becomes zero. First, for the future purpose, let derive some important properties for *MOBGLD*. These properties will be helpful for further derivations.

Theorem 3.2.1. If $X_i; i = 1, 2$ is defined as $X_i = \min(U_0, U_i)$ where U_0, U'_i s are independent rv following $GLD(\beta, \theta_j) \forall j = 0, 1, 2$. Then for the bivariate rv $(X_1, X_2) \sim MOBGLD(\beta, \Theta)$ and $\Delta = \theta_0 + \theta_1 + \theta_2$, we have

- (i) $X_i \sim GLD(\beta, \theta_0 + \theta_i); i = 1, 2,$

(ii) $T = \min\{X_1, X_2\} \sim GLD(\beta, \Delta)$.

Proof. (i) The survival function of rv $X_i; i = 1, 2$ is defined as

$$S(x_i) = P(X_i > x_i) = P(\min(U_0, U_i) > x_i)$$

Since U_0 and U'_i s are independent, so

$$\begin{aligned} S(x_i) &= P(U_0 > x_i) P(U_i > x_i) \\ &= \exp\{-(\theta_0 + \theta_i)[g(x_i)]^\beta\}, \end{aligned}$$

which is the survival function of $GLD(\beta, \theta_0 + \theta_i)$ from (3.3) and hence $X_i \sim GLD(\beta, \theta_0 + \theta_i)$; $i = 1, 2$.

(ii) Let $T = \min(X_1, X_2)$ then the survival function of T is given by

$$\begin{aligned} S(t) &= P(\min(X_1, X_2) > t) \\ &= P(\min(U_0, U_1) > t, \min(U_0, U_2) > t) \end{aligned}$$

Since U_0, U_1 and U_2 are independent, so

$$\begin{aligned} S(t) &= P(U_0 > t) P(U_1 > t) P(U_2 > t) \\ &= \exp\{-\Delta[g(t)]^\beta\}. \end{aligned}$$

Which is the survival function of $GLD(\beta, \Delta)$ from (3.3) and hence $T \sim GLD(\beta, \Delta)$, where $\Delta = \theta_0 + \theta_1 + \theta_2$. \square

Theorem 3.2.2. If a bivariate rv (X_1, X_2) follows the $MOBGLD$ and $X_0 = \max(X_1, X_2)$, then the joint survival function of $MOBGLD$ is given by

$$S(x_1, x_2) = \prod_{i=0}^2 \exp\{-\theta_i[g(x_i)]^\beta\}. \quad (3.5)$$

Proof. By the definition of rv X_1 and X_2 , the joint survival function of (X_1, X_2) is given by

$$\begin{aligned} S(x_1, x_2) &= P(X_1 > x_1, X_2 > x_2) \\ &= P(U_0 > x_0, U_1 > x_1, U_2 > x_2), \quad \text{where } x_0 = \max \{x_1, x_2\} \\ &= P(U_0 > x_0)P(U_1 > x_1)P(U_2 > x_2) \\ &= \prod_{i=0}^2 \exp \left\{ -\theta_i [g(x_i)]^\beta \right\} \end{aligned}$$

Corollary 1 The joint survival function of (X_1, X_2) for *MOBGLD* can also be defined such as (see Marshall and Olkin 1967)

$$S(x_1, x_2) = \begin{cases} S_0(x_1, x_2) = S_{GLD}(x; \beta, \Delta) & \text{if } x_1 = x_2 = x, \\ S_1(x_1, x_2) = S_{GLD}(x_1; \beta, \theta_0 + \theta_1)S_{GLD}(x_2; \beta, \theta_2) & \text{if } x_1 > x_2, \\ S_2(x_1, x_2) = S_{GLD}(x_2; \beta, \theta_0 + \theta_2)S_{GLD}(x_1; \beta, \theta_1) & \text{if } x_2 > x_1. \end{cases} \quad (3.6)$$

□

Theorem 3.2.3. The joint pdf for (X_1, X_2) for *MOBGLD* is given by

$$f(x_1, x_2) = \begin{cases} f_0(x_1, x_2) = \frac{\theta_0}{\Delta} f_{GLD}(x; \beta, \Delta) & \text{if } x_1 = x_2 = x, \\ f_1(x_1, x_2) = f_{GLD}(x_1; \beta, \theta_0 + \theta_1)f_{GLD}(x_2; \beta, \theta_2) & \text{if } x_1 > x_2, \\ f_2(x_1, x_2) = f_{GLD}(x_2; \beta, \theta_0 + \theta_2)f_{GLD}(x_1; \beta, \theta_1) & \text{if } x_2 > x_1. \end{cases} \quad (3.7)$$

Proof. Let first consider the case when $x_1 > x_2$. The joint pdf $f_1(x_1, x_2)$ can be obtained by the second order partial differentiation of $S_1(x_1, x_2)$ from (3.6) as follows

$$\begin{aligned} f_1(x_1, x_2) &= \frac{\partial^2}{\partial x_1 \partial x_2} S_1(x_1, x_2) \\ &= \frac{\partial}{\partial x_1} S_{GLD}(x_1; \beta, \theta_0 + \theta_1) \frac{\partial}{\partial x_2} S_{GLD}(x_2; \beta, \theta_2) \\ &= f_{GLD}(x_1; \beta, \theta_0 + \theta_1) f_{GLD}(x_2; \beta, \theta_2) \end{aligned}$$

Similarly when $x_2 > x_1$, by the second order partial differentiation of $S_2(x_1, x_2)$ from (3.6),

we have

$$f_2(x_1, x_2) = f_{GLD}(x_2; \beta, \theta_0 + \theta_2) f_{GLD}(x_1; \beta, \theta_1).$$

The joint pdf $f_0(x_1, x_2)$, when $x_1 = x_2 = x$ can be obtained by using the fact that

$$\int_0^\infty \int_{x_2}^\infty f_1(x_1, x_2) dx_2 dx_1 + \int_0^\infty \int_{x_1}^\infty f_2(x_1, x_2) dx_1 dx_2 + \int_0^\infty f_0(x) dx = 1 \quad (3.8)$$

From the first integral expression of left hand side of equation (3.8), we have

$$\begin{aligned} \int_0^\infty \int_{x_2}^\infty f_1(x_1, x_2) dx_1 dx_2 &= \int_0^\infty \left[\int_{x_2}^\infty f_{GLD}(x_1; \beta, \theta_0 + \theta_1) f_{GLD}(x_2; \beta, \theta_2) dx_1 \right] dx_2 \\ &= \frac{\theta_2}{\Delta} \int_0^\infty f_{GLD}(x_2; \beta, \Delta) dx_2 \\ &= \frac{\theta_2}{\Delta}. \end{aligned}$$

Similarly from the second integral expression of equation (3.8), we get

$$\int_0^\infty \int_{x_1}^\infty f_2(x_1, x_2) dx_2 dx_1 = \frac{\theta_1}{\Delta}.$$

By replacing the above obtained values in equation (3.8) and as *rv* X follows $GLD(\beta, \theta)$ given in (3.1), it is easy to obtain that

$$f_0(x_1, x_2) = \frac{\theta_0}{\Delta} f_{GLD}(x; \beta, \Delta).$$

Corollary 2 The joint hazard function, $h(x_1, x_2)$, for bivariate *rv* (X_1, X_2) of MOBGLD is

$$h(x_1, x_2) = \begin{cases} h_0(x_1, x_2) = \beta \theta_0 g'(x) [g(x)]^{\beta-1} & \text{if } x_1 = x_2 = x, \\ h_1(x_1, x_2) = \beta^2 (\theta_0 + \theta_1) \theta_2 g'(x_1) g'(x_2) [g(x_1) g(x_2)]^{\beta-1} & \text{if } x_1 > x_2, \\ h_2(x_1, x_2) = \beta^2 (\theta_0 + \theta_2) \theta_1 g'(x_1) g'(x_2) [g(x_1) g(x_2)]^{\beta-1} & \text{if } x_2 > x_1. \end{cases} \quad (3.9)$$

Proof The hazard function of a bivariate random vector (X_1, X_2) can be defined as $h(x_1, x_2) =$

$f(x_1, x_2)/S(x_1, x_2)$. So let consider the case when $X_1 > X_2$ then the hazard function can be obtained as

$$\begin{aligned} h_1(x_1, x_2) &= \frac{f_{GLD}(x_1; \beta, \theta_0 + \theta_1)f_{GLD}(x_2; \beta, \theta_2)}{S_{GLD}(x_1; \beta, \theta_0 + \theta_1)S_{GLD}(x_2; \beta, \theta_2)} \\ &= h_{GLD}(x_1; \beta, \theta_0 + \theta_1)h_{GLD}(x_2; \beta, \theta_2) \\ &= \beta^2(\theta_0 + \theta_1)\theta_2 g'(x_1)g'(x_2)[g(x_1)g(x_2)]^{\beta-1}. \end{aligned}$$

Similarly for $X_2 > X_1$ and $X_1 = X_2 = X$, respectively, we get

$$h_2(x_1, x_2) = \beta^2(\theta_0 + \theta_2)\theta_1 g'(x_1)g'(x_2)[g(x_1)g(x_2)]^{\beta-1},$$

and

$$h_0(x_1, x_2) = \beta\theta_0 g'(x)[g(x)]^{\beta-1}. \quad \square$$

3.2.3 Ageing Properties for MOBGLD

If for an individual, multiple causes are present, then different causes may have non-identical ageing behavior with respect to their survival time. So it becomes necessary to inspect the ageing behavior of such systems. If the bivariate random vector $(X_1, X_2) \sim MOBGLD(\beta, \Theta)$ in such a way that $S(x_1 + t, x_2 + t)/S(x_1, x_2)$ monotonically decreases/increases in $x_1, x_2, t \in \mathbb{R}^+$, we say that (X_1, X_2) have multivariate increasing/decreasing failure rate (*MIFR/MDFR*) property (see Barlow and Proschan 1975). Obviously, constant hazard rate (*CFR*) is a special class of both. Based on this, a classification of *MOBGLD* has been presented in Table 3.2.

3.2.4 Dependent Competing Risks for MOBGLD

As the objective of this chapter is to inference the lifetime of cause-dependent competing risks, so let first match with the competing risks set-up for *MOBGLD*. Suppose that there are n independent series systems with two causes of failure in a lifetime experiment. Here, we assume that the causes are not independent. Let $T_j = \min \{X_{1j}, X_{2j}\}$ be failure time for the j^{th} series system; $j = 1, 2, \dots, n$; where X_{ij} denotes the failure time of j^{th} system

due to i^{th} component(cause), $i = 1, 2$. Further assume that (X_{1j}, X_{2j}) are the bivariate random vector such as $(X_{1j}, X_{2j}) \sim MOBGLD(\beta, \Theta)$. Let $(\delta_{1j}, \delta_{2j})$ be the indicator for the components causing failure for the system, i.e., when $T_j = X_{1j} = X_{2j}$, $T_j = X_{1j} < X_{2j}$ and $T_j = X_{2j} < X_{1j}$ then one get $(1, 1)$, $(1, 0)$ and $(0, 1)$, respectively. Here it is to be noted that T_j and $(\delta_{1j}, \delta_{2j})$ are independent. In this study, we are considering complete sample analysis, i.e., none of the observations are censored and no information is missing partially or completely.

Table 3.2: Ageing classification of $MOBGLD$ for specific distributions.

Model	Parameter	Survival function	Class
MOBE	$\beta = 1$	$\prod_{i=0}^2 \exp \{-\theta_i x_i\}$	MCFR
MOBR	$\beta = 2$	$\prod_{i=0}^2 \exp \{-\theta_i x_i^2\}$	MIFR
MOBP	$\beta = 1$	$\prod_{i=0}^2 \exp \{-\theta_i \log(x_i/a)\}$	MDFR
MOBB	$\beta = 1$	$\prod_{i=0}^2 \exp \{-\theta_i \log(1 + x_i^b)\}$	MDFR
MOBW	$\beta > 1 (\beta < 1)$	$\prod_{i=0}^2 \exp \{-\theta_i x_i^\beta\}$	MIFR(MDFR)

Note: $x_0 = \max(x_1, x_2)$

3.3 Inference for Competing Risks Model

In this section, two objectives have been fulfilled separately. In first, a simple Fisher information matrix has been calculated for original parameters by considering the parameter β known as well as unknown. However, sometimes one may also be interested in the inference of other characteristics; those can be acquired by the combination or transformation of parameters. Hence, in the second attempt, the Fisher information matrix has been calculated for re-parametrize set of parameters according to the statistic of keen interest.

3.3.1 Fisher Information Matrix

As discussed in Section 2.4, let n independent series systems with two causes of failure are considered in the experiment. The likelihood function can be formulated for mentioned

⁰ **MOBE:** Marshall-Olkin Bivariate Exponential; **MOBR:** Marshall-Olkin Bivariate Rayleigh; **MOBP:** Marshall-Olkin Bivariate Pareto; **MOBB:** Marshall-Olkin Bivariate Burr; **MOBW:** Marshall-Olkin Bivariate Weibull

bivariate competing risks observations $(T_j, \delta_{1j}, \delta_{2j})$; $j = 1, 2, \dots, n$, as follows (Lawless 2003)

$$L(\beta, \Theta | \underline{t}) = \prod_{j=1}^n \left[f(t_j, t_j) \right]^{\delta_{1j}\delta_{2j}} \left[-\frac{\partial S(x_{1j}, x_{2j})}{\partial x_1} \Big|_{(t_j, t_j)} \right]^{\delta_{1j}(1-\delta_{2j})} \left[-\frac{\partial S(x_{1j}, x_{2j})}{\partial x_2} \Big|_{(t_j, t_j)} \right]^{\delta_{2j}(1-\delta_{1j})}. \quad (3.10)$$

By using the appropriate expressions from (3.6) and (3.7), we have

$$\begin{aligned} L(\beta, \Theta | \underline{t}) &= \prod_{j=1}^{n_0} \left(\frac{\theta_0}{\Delta} f_{GLD}(t_j, \beta, \Delta) \right) \prod_{i=1}^{n_1} S_{GLD}(t_j; \beta, \theta_0 + \theta_2) f_{GLD}(t_j, \beta, \theta_1) \\ &\quad \prod_{j=1}^{n_2} S_{GLD}(t_j; \beta, \theta_0 + \theta_1) f_{GLD}(t_j, \beta, \theta_2) \\ &= \beta^{n_0} \theta_0^{n_0} \prod_{i=1}^{n_0} [g(t_j)]^{\beta-1} g'(t_j) \exp \left(-\Delta \sum_{i=1}^{n_0} g(t_j)^\beta \right) \\ &\quad \beta^{n_1} \theta_1^{n_1} \prod_{i=1}^{n_1} [g(t_j)]^{\beta-1} g'(t_j) \exp \left(-\theta_1 \sum_{i=1}^{n_1} g(t_j)^\beta \right) \exp \left(-(\theta_0 + \theta_2) \sum_{i=1}^{n_2} g(t_j)^\beta \right) \\ &\quad \beta^{n_2} \theta_2^{n_2} \prod_{i=1}^{n_2} [g(t_j)]^{\beta-1} g'(t_j) \exp \left(-\theta_2 \sum_{i=1}^{n_1} g(t_j)^\beta \right) \exp \left(-(\theta_0 + \theta_1) \sum_{i=1}^{n_2} g(t_j)^\beta \right). \end{aligned}$$

Then the likelihood equation (3.10) for $MOBGLD(\beta, \Theta)$ given data $(T_j, \delta_{1j}, \delta_{2j})$; $j = 1, 2, \dots, n$, comes out to be

$$L(\beta, \Theta | \underline{t}) = \beta^n \theta_0^{n_0} \theta_1^{n_1} \theta_2^{n_2} \prod_{j=1}^n \left\{ g'(t_j) [g(t_j)]^{\beta-1} \right\} \exp \left(-\Delta \sum_{j=1}^n [g(t_j)]^\beta \right), \quad (3.11)$$

where n_0, n_1 and n_2 are such as

$$\begin{aligned} \sum_{j=1}^n I(X_{1j} = X_{2j}) &= \sum_{j=1}^n \delta_{1j}\delta_{2j} = n_0, \\ \sum_{j=1}^n I(X_{1j} < X_{2j}) &= \sum_{j=1}^n \delta_{1j}(1 - \delta_{2j}) = n_1, \\ \sum_{j=1}^n I(X_{1j} > X_{2j}) &= \sum_{j=1}^n \delta_{2j}(1 - \delta_{1j}) = n_2. \end{aligned}$$

Here, $n = n_0 + n_1 + n_2$, as it is assumed that no information is censored or missing in the experiment. Taking the logarithm of (3.11), the log-likelihood function comes out to be

$$l(\beta, \Theta | \underline{t}) \propto n \ln \beta + n_0 \ln \theta_0 + n_1 \ln \theta_1 + n_2 \ln \theta_2 + (\beta - 1) \sum_{j=1}^n \ln g(t_j) - \Delta \sum_{j=1}^n [g(t_j)]^\beta. \quad (3.12)$$

Now taking first order partial derivative of (3.12) with respect to β and $\theta_i; i = 0, 1, 2$ and equating them to zero, one can obtain maximum likelihood (*ML*) estimator of parameters $(\beta, \theta_0, \theta_1, \theta_2)$.

In order to calculate the expectation of second-order partial derivatives of the log-likelihood function (3.12), if assume that $z = \Delta[g(t)]^\beta$ then z follows exponential distribution with mean 1. Now as $[g(t)]^\beta = z/\Delta$ and it implies that $g(t) = (z/\Delta)^{1/\beta}$ or $\ln g(t) = \frac{1}{\beta} \ln(z/\Delta)$. Hence the expectation term can be calculated such as

$$\begin{aligned} E[[g(t)]^\beta \ln g(t)] &= E\left[\frac{z}{\Delta} \cdot \frac{1}{\beta} \ln\left(\frac{z}{\Delta}\right)\right] \\ &= \frac{1}{\Delta\beta} [E[z \ln z] - \ln \Delta] \\ &= \frac{1}{\Delta\beta} \left[\int_0^\infty z \ln z \exp\{-z\} dz - \ln \Delta \right] \\ &= \frac{1}{\Delta\beta} \left[-|z \ln z \exp\{-z\}|_0^\infty + \int_0^\infty (1 + \ln z) \exp\{-z\} dz - \ln \Delta \right] \\ &= \frac{1}{\Delta\beta} [r_1 + (1 - \ln \Delta)], \end{aligned} \quad (3.13)$$

where $r_u = \int_0^\infty (\ln z)^u \exp\{-z\} dz; u \in \mathcal{N}$ and similarly

$$\begin{aligned} E[[g(t)]^\beta (\ln g(t))^2] &= E\left[\frac{z}{\Delta} \cdot \frac{1}{\beta^2} \ln\left(\frac{z}{\Delta}\right)^2\right] \\ &= \frac{1}{\Delta\beta^2} [E[z(\ln z)^2] + (\ln \Delta)^2 - 2(1 + r_1) \ln \Delta] \\ &= \frac{1}{\Delta\beta^2} \left[\int_0^\infty z(\ln z)^2 e^{-z} dz + (\ln \Delta)^2 - 2(1 + r_1) \ln \Delta \right] \\ &= \frac{1}{\Delta\beta^2} \left[2 \int_0^\infty \ln z e^{-z} dz + \int_0^\infty (\ln z)^2 e^{-z} dz + (\ln \Delta)^2 - 2(1 + r_1) \ln \Delta \right] \\ &= \frac{1}{\Delta\beta^2} [r_2 + 2r_1(1 - \ln \Delta) - \ln \Delta(2 - \ln \Delta)]. \end{aligned} \quad (3.14)$$

Using the calculated expectation terms above, the expectation terms of second-order partial

derivatives of the log-likelihood function (3.12) are as follows:

$$\begin{aligned} -E\left[\frac{\partial^2 l(\beta, \Theta | \underline{t})}{\partial \theta_i^2}\right] &= \frac{1}{\theta_i} \frac{n}{\Delta}; i = 0, 1, 2. \\ -E\left[\frac{\partial^2 l(\beta, \Theta | \underline{t})}{\partial \beta \partial \theta_i}\right] &= \frac{n}{\Delta \beta} \left(r_1 + (1 - \ln \Delta)\right); i = 0, 1, 2. \\ -E\left[\frac{\partial^2 l(\beta, \Theta | \underline{t})}{\partial \beta^2}\right] &= \frac{n}{\beta^2} \left(r_2 + 2r_1(1 - \ln \Delta) - \ln \Delta(2 - \ln \Delta) + 1\right); \\ -E\left[\frac{\partial^2 l(\beta, \Theta | \underline{t})}{\partial \theta_i \partial \theta_k}\right] &= 0; i \neq k = 0, 1, 2. \end{aligned}$$

It is noticeable that the *GLD* family can be bifurcated on the behalf of parameter β . If β is known, then we get exponential, Burr, Pareto and Rayleigh distribution, but for unknown β , Weibull distribution will be the particular case of *GLD*. So, accordingly, we can formulate the Fisher information matrix. Let consider two different cases:

Case (i): When β is known

The Fisher Information matrix for $\Theta = (\theta_0, \theta_1, \theta_2)$ is given by

$$\Sigma_1 = \begin{bmatrix} \frac{n}{\Delta} \frac{1}{\theta_0} & 0 & 0 \\ 0 & \frac{n}{\Delta} \frac{1}{\theta_1} & 0 \\ 0 & 0 & \frac{n}{\Delta} \frac{1}{\theta_2} \end{bmatrix}. \quad (3.15)$$

Case (ii): When β is unknown

The Fisher Information matrix for (β, Θ) is given by

$$\Sigma_2 = \begin{bmatrix} \frac{n}{\beta^2} k(\Delta) & \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] \\ \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & \frac{n}{\Delta} \frac{1}{\theta_0} & 0 & 0 \\ \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & 0 & \frac{n}{\Delta} \frac{1}{\theta_1} & 0 \\ \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & 0 & 0 & \frac{n}{\Delta} \frac{1}{\theta_2} \end{bmatrix}, \quad (3.16)$$

where $k(\Delta) = r_2 + 2r_1(1 - \ln \Delta) - \ln \Delta(2 - \ln \Delta) + 1$. Now a $100(1 - \alpha)\%$ asymptotic confidence interval for any function of parameter say, $h(\theta)$, can be obtained by

$$\left(h(\hat{\theta}) - Z_{\alpha/2} \sqrt{\text{Var}(h(\hat{\theta}))}, h(\hat{\theta}) + Z_{\alpha/2} \sqrt{\text{Var}(h(\hat{\theta}))}\right),$$

where $h(\hat{\theta})$ and $Var(h(\hat{\theta}))$ is the *ML* estimate and variance of $h(\hat{\theta})$, respectively. The $Z_{\alpha/2}$ is the $(\alpha/2)^{th}$ quantile of the standard normal distribution with $0 < \alpha < 1$.

3.3.2 Estimation under Re-parametrization

In competing risks analysis, always a keen interest would be in the cause which is more severe (performing poorly) in respect to others and subsequently in the study of its characteristics. So, if parameters are re-parametrized then much more inferential objectives can be achieved. So, a particular set after re-parametrization has been considered, such as, say,

$$\lambda_1 = \Delta = \theta_0 + \theta_1 + \theta_2, \quad \lambda_2 = \frac{\theta_1}{\Delta}, \quad \lambda_3 = \frac{\theta_2}{\Delta}.$$

The re-parametrized quantities $\lambda_2 = P(X_1 < X_2)$ and $\lambda_3 = P(X_2 < X_1)$ represent the relative risk due to cause 1 and cause 2, respectively. The transformation $(\beta, \Lambda) = (\beta, \lambda_1, \lambda_2, \lambda_3)$ from $(\beta, \Theta) = (\beta, \theta_0, \theta_1, \theta_2)$ is one-to-one with the inverse transformation where

$$\beta = \beta, \quad \theta_1 = \lambda_1 \lambda_2, \quad \theta_2 = \lambda_1 \lambda_3, \quad \theta_0 = \lambda_1(1 - \lambda_2 - \lambda_3).$$

Now the likelihood function under re-parametrization is given by

$$L(\beta, \Lambda | \underline{t}) = \beta^n \lambda_1^n \lambda_2^{n_1} \lambda_3^{n_2} (1 - \lambda_2 - \lambda_3)^{n_0} \prod_{j=1}^n \{g'(t_j)[g(t_j)]^{\beta-1}\} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\}. \quad (3.17)$$

Taking the logarithmic of equation (3.17) and differentiating with respect to unknown parameters and equating to zero, one can obtain the *ML* estimators of (β, Λ) . The inverse of Fisher information matrix can be obtained by utilizing matrix Σ and respective Jacobian transformations, such as

Case (i): When β is known

After re-parametrization, the Jacobian matrix of the transformation and Fisher information

matrix are follows:

$$J_1 = \begin{bmatrix} 1 - \lambda_2 - \lambda_3 & -\lambda_1 & -\lambda_1 \\ \lambda_2 & \lambda_1 & 0 \\ \lambda_3 & 0 & \lambda_1 \end{bmatrix}$$

and

$$\Sigma_{R_1} = J'_1 \Sigma_1 J_1 = n \begin{bmatrix} \frac{1}{\lambda_1^2} & 0 & 0 \\ 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}.$$

Hence, the inverse of the Fisher information matrix for Λ is given by

$$\Sigma_{R_1}^{-1} = \begin{bmatrix} \frac{\lambda_1^2}{n} & 0 & 0 \\ 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}. \quad (3.18)$$

Case (ii): When β is unknown

The Jacobian matrix of the transformation and Fisher information matrix are follows:

$$J_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \lambda_2 - \lambda_3 & -\lambda_1 & -\lambda_1 \\ 0 & \lambda_2 & \lambda_1 & 0 \\ 0 & \lambda_3 & 0 & \lambda_1 \end{bmatrix}$$

and

$$\Sigma_{R_2} = J'_2 \Sigma_2 J_2 = n \begin{bmatrix} \frac{k(\Lambda)}{\beta^2} & \frac{r_1+1-\ln\lambda_1}{\beta\lambda_1} & 0 & 0 \\ \frac{r_1+1-\ln\lambda_1}{\beta\lambda_1} & \frac{1}{\lambda_1^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}$$

So, the inverse of the Fisher information matrix for (β, Λ) is given by

$$\Sigma_{R_2}^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2 - r_1^2)} & -\frac{\beta\lambda_1(r_1+1-\ln\lambda_1)}{n(r_2 - r_1^2)} & 0 & 0 \\ -\frac{\beta\lambda_1(r_1+1-\ln\lambda_1)}{n(r_2 - r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2 - r_1^2)} & 0 & 0 \\ 0 & 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}. \quad (3.19)$$

The diagonal elements of the matrices (3.18) and (3.19) give the variances of the *ML* estimators of the parameter under consideration.

3.4 Bayesian Inference for Competing Risks Model

In Bayesian inference, the prior information regarding the parameter may be quite limited or not much credible sometimes. In such situations, it is expected that the observed data have enough freedom to express itself in the sense of information it contains. In such a case, considering a non-informative prior becomes a better choice for Bayesian inference. One of the popular non-informative priors is the reference prior, which is discussed with specific steps of the algorithm by Berger and Bernardo (1992). The reference prior is equivalent to Jeffrey's prior in the single parameter case and it overcomes the deficiencies of Jeffrey's prior in the multi parameter case. When some of the parameters have to be given more weight according to their inferential importance in the model, then it becomes inevitable to design the Bayesian inference accordingly. The reference prior is quite useful to fulfill this requirement.

As we discussed earlier, in competing risks analysis, we are interested in a cause wise inference. And many times we may have more concerns for a particular cause or a subset of causes presented in study. For example, in a clinical trial data is collected from cancer patients and the two possible causes of death of an individual may be cancer or any other factor. So in this study, we are focused on deaths due to cancer and related statistics. So parameters related to the cause distribution of cancer must be estimated with highest precision. Even though if all causes are of equal importance, then in competing risks analysis,

researchers are primarily interested towards those causes which are performing very poorly and lastly to those who are good enough. The reference prior based study provides such a facility to researchers. In reference priors set-up, we distribute the unknown parameters in different subsets. These sets are designed in an ordered manner by the expert according to the importance of sets(parameters) for the inferential purpose. So accordingly, parameters (i.e., causes) can be elected in different sets having more significance in respect to others (see Ghosh and Mukerjee 1992). Earlier Guan et al. (2013) discussed about the suitability of reference priors for the competing risks analysis.

In this chapter, some preferable choices of ordered grouping have been taken and respective reference priors are shown in Table 3.3 and Table 3.4 when the parameter β is considered known and unknown, respectively. These ordered groupings have been done on the basis of their inferential significance. In the following subsection, the reference prior π_{11} for known β and π_{21} and π_{24} for unknown β have been derived.

3.4.1 Reference Prior for MOBGLD

To obtain the reference prior for parameters according to our inferential interest, we proceed as follows. First of all, order a multi-dimensional parameter $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ and separate them into m groups of sizes g_1, g_2, \dots, g_m such as $\alpha_{(i)} = (\alpha_{N_{i-1}+1}, \alpha_{N_{i-1}+2}, \dots, \alpha_{N_i}) \dots$, where $N_j = \sum_{i=1}^j g_i$. Also let $\alpha_{[j]} = (\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(j)}) = (\alpha_1, \alpha_2, \dots, \alpha_{N_j})$ and $\alpha_{[\sim j]} = (\alpha_{N_j+1}, \dots, \alpha_k)$, then it is obvious that $\alpha_{[\sim 0]} = \alpha$ and $\alpha_{[0]}$ is vacuous. Let $\Sigma(\alpha)$ and $\Sigma^{-1}(\alpha)$ are Fisher information matrix and variance-covariance matrix respectively. Here Σ^{-1} can be defined as block partition matrix such as

$$\Sigma^{-1} = \begin{pmatrix} A_{11} & A_{21}^t & \dots & A_{m1}^t \\ A_{21} & A_{22} & \dots & A_{m2}^t \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{pmatrix},$$

where A_{ij} is a $n_j \times n_j$ matrix. Let Σ_j^{-1} = Upper Left ($N_j \times N_j$) corner of Σ^{-1} , with $\Sigma_m^{-1} = \Sigma^{-1}$, and $H_j = \Sigma_j$. Then the matrices h_j = Lower Right ($n_j \times n_j$) corner of H_j ; $j=1, 2, \dots, m$,

have an important role in deriving the reference priors. In particular, $h_1 = H_1 = A_{11}^{-1}$ and, if S is a block diagonal matrix, then $h_j = A_{jj}^{-1}$, $j=1,2,\dots,m$. Finally, the reference prior has been derived for possible orders using the lemma given by Berger and Bernardo (1992).

Lemma 3.4.1. *If $|h_j(\alpha)|$ depending only on $\alpha_{[j]}$ holds, for $j=1,2,\dots,m$, then the reference prior*

$$\pi(\alpha) = \lim_{k \rightarrow \infty} \frac{\pi^k(\alpha)}{\pi^k(\alpha^*)}, \quad \text{where } \pi^k(\alpha) = \left(\prod_{j=1}^m \frac{|h_j|^{1/2}}{\int_{\Theta(\alpha_{[j-1]})} |h_j|^{1/2} d\alpha_{(j)}} \right) I_{\Theta}(\alpha),$$

$|h_j(\alpha)|$ is the determinant of h_j , α^* is any fixed point in Θ , and Θ is a compact subset.

Here, few reference priors have been derived by considering the parameter β known as well as unknown.

Case (i): When β is known

Let consider an order grouping of parameters such as $\{\lambda_1, (\lambda_2, \lambda_3)\}$ to derive the reference prior, say $\pi_{11}(\Lambda)$. As here in this case, the number of groups is $m = 2$ with group sizes $g_1 = 1$ and $g_2 = 2$, given by $\alpha_{(1)} = \lambda_1$ and $\alpha_{(2)} = (\lambda_2, \lambda_3)$. Also $\alpha_{[1]} = \lambda_1$ and $\alpha_{[2]} = (\lambda_1, \lambda_2, \lambda_3)$ with $\alpha_{[\sim 0]} = (\lambda_1, \lambda_2, \lambda_3)$ and $\alpha_{[\sim 1]} = (\lambda_2, \lambda_3)$ $\alpha_{[0]}$ and $\alpha_{[\sim 2]}$ are vacuous. The variance-covariance matrix for $\{\lambda_1, (\lambda_2, \lambda_3)\}$ is

$$\Sigma^{-1} = \begin{bmatrix} \frac{\lambda_1^2}{n} & 0 & 0 \\ 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}.$$

Hence $\Sigma_1^{-1} = \frac{\lambda_1^2}{n}$, $H_1 = \frac{n}{\lambda_1^2}$, $h_1 = H_1$, $\Sigma_2^{-1} = \Sigma^{-1}$,

$$H_2 = \Sigma = \begin{bmatrix} \frac{n}{\lambda_1^2} & 0 & 0 \\ 0 & n\frac{1-\lambda_3}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{1-\lambda_2-\lambda_3} \\ 0 & \frac{n}{1-\lambda_2-\lambda_3} & n\frac{1-\lambda_2}{\lambda_3(1-\lambda_2-\lambda_3)} \end{bmatrix}, \quad \text{and } h_2 = n \begin{bmatrix} \frac{1-\lambda_3}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{1}{1-\lambda_2-\lambda_3} \\ \frac{1}{1-\lambda_2-\lambda_3} & \frac{1-\lambda_2}{\lambda_3(1-\lambda_2-\lambda_3)} \end{bmatrix}.$$

Now let choose the parametric space such as $\Theta_k = \{\lambda_1, (\lambda_2, \lambda_3)\} | a_{1k} < \lambda_1 < b_{1k}, a_{2k} < \lambda_2, a_{3k} < \lambda_3, \lambda_2 + \lambda_3 < d_k\}$, where $a_{1k}, a_{2k}, a_{3k} \rightarrow 0$, $b_{1k} \rightarrow \infty$, $d_k \rightarrow 1$. Since h_1 and h_2

satisfy Lemma 3.4.1, so the reference prior for $\{\lambda_1, (\lambda_2, \lambda_3)\}$ is given by

$$\pi_{11}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}. \quad (3.20)$$

Case (ii): When β is unknown

Let first derive the reference prior, say $\pi_{21}(\beta, \Lambda)$, for the ordered grouping $\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$. In this case, the number of groups is $m = 3$ given as $\alpha_{(1)} = (\lambda_2, \lambda_3)$, $\alpha_{(2)} = \beta$ and $\alpha_{(3)} = \lambda_1$. Also $\alpha_{[1]} = (\lambda_2, \lambda_3)$, $\alpha_{[2]} = (\lambda_2, \lambda_3, \beta)$, $\alpha_{[3]} = (\lambda_2, \lambda_3, \beta, \lambda_1)$, $\alpha_{[\sim 0]} = (\lambda_2, \lambda_3, \beta, \lambda_1)$; $\alpha_{[\sim 1]} = (\beta, \lambda_1)$ and $\alpha_{[\sim 2]} = \lambda_1$ and $\alpha_{[0]}$ and $\alpha_{[\sim 3]}$ as are vacuous. So the variance-covariance matrix for $\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$ is

$$\Sigma^{-1} = \begin{bmatrix} \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} & 0 & 0 \\ -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} & 0 & 0 \\ 0 & 0 & \frac{\beta^2}{n(r_2-r_1^2)} & -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} \\ 0 & 0 & -\frac{\beta\lambda_1(1+r_1-\ln\lambda)}{n(r_2-r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2-r_1^2)} \end{bmatrix}.$$

Then we obtain

$$\Sigma_1^{-1} = \begin{bmatrix} \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}, \quad H_1 = \begin{bmatrix} \frac{n(1-\lambda_3)}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{(1-\lambda_2-\lambda_3)} \\ \frac{n}{(1-\lambda_2-\lambda_3)} & \frac{n(1-\lambda_2)}{\lambda_3(1-\lambda_2-\lambda_3)} \end{bmatrix},$$

$$\Sigma_2^{-1} = \begin{bmatrix} \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} & 0 \\ -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} & 0 \\ 0 & 0 & \frac{\beta^2}{n(r_2-r_1^2)} \end{bmatrix}, \quad H_2 = \begin{bmatrix} \frac{n(1-\lambda_3)}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{(1-\lambda_2-\lambda_3)} & 0 \\ \frac{n}{(1-\lambda_2-\lambda_3)} & \frac{n(1-\lambda_2)}{\lambda_3(1-\lambda_2-\lambda_3)} & 0 \\ 0 & 0 & \frac{n(r_2-r_1^2)}{\beta^2} \end{bmatrix},$$

$$\Sigma_3^{-1} = \Sigma^{-1} \text{ and } H_3 = \Sigma = \begin{bmatrix} \frac{n(1-\lambda_3)}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{(1-\lambda_2-\lambda_3)} & 0 & 0 \\ \frac{n}{(1-\lambda_2-\lambda_3)} & \frac{n(1-\lambda_2)}{\lambda_3(1-\lambda_2-\lambda_3)} & 0 & 0 \\ 0 & 0 & \frac{nk(\lambda_1)}{\beta^2} & \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} \\ 0 & 0 & \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} & \frac{n}{\lambda_1^2} \end{bmatrix}.$$

Hence, we get $h_1 = H_1$, $h_2 = \frac{n(r_2 - r_1^2)}{\beta^2}$ and $h_3 = \frac{n}{\lambda_1^2}$. Now the parametric space $\Theta_k = \{(\lambda_2, \lambda_3), \beta, \lambda_1\} | a_{1k} < \beta < b_{1k}, a_{2k} < \lambda_1 < b_{2k}, a_{3k} < \lambda_2, a_{4k} < \lambda_3, \lambda_2 + \lambda_3 < d_k\}$; where $a_{1k}, a_{2k}, a_{3k}, a_{4k} \rightarrow 0$, $b_{1k}, b_{2k} \rightarrow \infty$, $d_k \rightarrow 1$. Since h_1, h_2 and h_3 satisfy the Lemma 3.4.1, so the reference prior for $\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$ is given by

$$\pi_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}. \quad (3.21)$$

Let attempt for one more reference prior, say, $\pi_{24}(\beta, \Lambda)$ for ordered grouping $\{\beta, \lambda_1, \lambda_2, \lambda_3\}$. For this case, the number of groups $m = 4$ which is given by $\alpha_{(1)} = \beta, \alpha_{(2)} = \lambda_1, \alpha_{(3)} = \lambda_2$ and $\alpha_{(4)} = \lambda_3$. So we have $\alpha_{[1]} = \beta, \alpha_{[2]} = (\beta, \lambda_1), \alpha_{[3]} = (\beta, \lambda_1, \lambda_2), \alpha_{[4]} = (\beta, \lambda_1, \lambda_2, \lambda_3), \alpha_{[\sim 0]} = (\beta, \lambda_1, \lambda_2, \lambda_3), \alpha_{[\sim 1]} = (\lambda_1, \lambda_2, \lambda_3), \alpha_{[\sim 2]} = (\lambda_2, \lambda_3), \alpha_{[\sim 3]} = \lambda_3$ whereas $\alpha_{[0]}$ and $\alpha_{[\sim 4]}$ are vacuous. The variance-covariance matrix for $\{\beta, \lambda_1, \lambda_2, \lambda_3\}$ is

$$\Sigma^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2 - r_1^2)} & -\frac{\beta \lambda_1 (1 + r_1 - \ln \lambda_1)}{n(r_2 - r_1^2)} & 0 & 0 \\ -\frac{\beta \lambda_1 (1 + r_1 - \ln \lambda_1)}{n(r_2 - r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2 - r_1^2)} & 0 & 0 \\ 0 & 0 & \frac{\lambda_2 (1 - \lambda_2)}{n} & -\frac{\lambda_2 \lambda_3}{n} \\ 0 & 0 & -\frac{\lambda_2 \lambda_3}{n} & \frac{\lambda_3 (1 - \lambda_3)}{n} \end{bmatrix}$$

Then we get $\Sigma_1^{-1} = \frac{\beta^2}{n(r_2 - r_1^2)}$, $H_1 = \frac{n(r_2 - r_1^2)}{\beta^2}$,

$$\Sigma_2^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2 - r_1^2)} & -\frac{\beta \lambda_1 (1 + r_1 - \ln \lambda_1)}{n(r_2 - r_1^2)} \\ -\frac{\beta \lambda_1 (1 + r_1 - \ln \lambda_1)}{n(r_2 - r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2 - r_1^2)} \end{bmatrix}, \quad H_2 = \begin{bmatrix} \frac{n k(\lambda_1)}{\beta^2} & \frac{n(1 + r_1 - \ln \lambda_1)}{\beta \lambda_1} \\ \frac{n(1 + r_1 - \ln \lambda_1)}{\beta \lambda_1} & \frac{n}{\lambda_1^2} \end{bmatrix}$$

$$\Sigma_3^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2 - r_1^2)} & -\frac{\beta \lambda_1 (1 + r_1 - \ln \lambda_1)}{n(r_2 - r_1^2)} & 0 \\ -\frac{\beta \lambda_1 (1 + r_1 - \ln \lambda_1)}{n(r_2 - r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2 - r_1^2)} & 0 \\ 0 & 0 & \frac{\lambda_2 (1 - \lambda_2)}{n} \end{bmatrix}, \quad H_3 = \begin{bmatrix} \frac{n k(\lambda_1)}{\beta^2} & \frac{n(1 + r_1 - \ln \lambda_1)}{\beta \lambda_1} & 0 \\ \frac{n(1 + r_1 - \ln \lambda_1)}{\beta \lambda_1} & \frac{n}{\lambda_1^2} & 0 \\ 0 & 0 & \frac{n}{\lambda_2 (1 - \lambda_2)} \end{bmatrix},$$

$\Sigma_4^{-1} = \Sigma^{-1}$ and $H_4 = \Sigma$. Hence we obtain

$$h_1 = \frac{n(r_2 - r_1^2)}{\beta^2}; \quad h_2 = \frac{n}{\lambda_1^2}; \quad h_3 = \frac{n}{\lambda_2 (1 - \lambda_2)} \text{ and } h_4 = \frac{1 - \lambda_2}{\lambda_3 (1 - \lambda_2 - \lambda_3)}.$$

Now, for the parametric space $\Theta_k = \{(\beta, \lambda_1, \lambda_2, \lambda_3) | a_{1k} < \beta < b_{1k}, a_{2k} < \lambda_1 < b_{2k}, a_{3k} <$

$\lambda_2, a_{4k} < \lambda_3, \lambda_2 + \lambda_3 < d_k\}$; where $a_{1k}, a_{2k}, a_{3k}, a_{4k} \rightarrow 0, b_{1k}, b_{2k} \rightarrow \infty, d_k \rightarrow 1$. Now we have to compute the following integral $\int_{\Theta(\alpha_{[j-1]})} |h_j|^{1/2} d\alpha_{(j)}$ for $j = 1, 2, 3$ and 4. So we obtain

$$\begin{aligned}\int_{\Theta(\alpha_{[0]})} |h_1|^{1/2} d\beta &= \int_{a_{1k}}^{b_{1k}} |h_1|^{1/2} d\beta = \{n(r_2 - r_1^2)\}^{1/2} (\log b_{1k} - \log a_{1k}), \\ \int_{\Theta(\alpha_{[1]})} |h_2|^{1/2} d\lambda_1 &= \int_{a_{2k}}^{b_{2k}} |h_2|^{1/2} d\lambda_1 = n^{1/2} (\log b_{2k} - \log a_{2k}).\end{aligned}$$

For some λ_3° , we get

$$\begin{aligned}\int_{\Theta(\alpha_{[2]})} |h_3|^{1/2} d\lambda_2 &= \int_{a_{3k}}^{d_k - \lambda_3^\circ} |h_3|^{1/2} d\lambda_2 = n^{1/2} [\sin^{-1}(1 - 2a_{3k}) - \sin^{-1}(1 - 2(d_k - \lambda_3^\circ))], \\ \int_{\Theta(\alpha_{[3]})} |h_4|^{1/2} d\lambda_3 &= \int_{a_{4k}}^{d_k - \lambda_2} |h_4|^{1/2} d\lambda_3 \\ &= \sqrt{1 - \lambda_2} \left[\sin^{-1} \left(\frac{1 - \lambda_2 - 2a_{4k}}{1 - \lambda_2} \right) - \sin^{-1} \left(\frac{1 - \lambda_2 - 2(d_k - \lambda_2)}{1 - \lambda_2} \right) \right].\end{aligned}$$

Using Lemma 3.4.1, we obtain the reference prior for $\{\beta, \lambda_1, \lambda_2, \lambda_3\}$, given by

$$\pi_{24}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)(1 - \lambda_2 - \lambda_3)]^{-1/2}. \quad (3.22)$$

Following the same procedure, one can also obtain other reference priors. A triplet indicator (l_1, l_2, l_3) has been used to represent all the considered reference priors and their respective posterior densities. On the basis of this triplet, a generalized form of reference prior as well as posterior density is given as follows:

Table 3.3: Reference priors for parameters, Λ , of $MOBGLD(\Lambda)$ and for parameters, Θ , of $MOBGLD(\Theta)$.

Ordered grouping	Reference prior for Λ	Reference prior for Θ
$\{\lambda_1, (\lambda_2, \lambda_3)\}$	$\pi_{11}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{11}(\Theta)$
$\{(\lambda_2, \lambda_3), \lambda_1\}$	$\pi_{11}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{11}(\Theta)$
$\{\lambda_1, \lambda_2, \lambda_3\}$	$\pi_{12}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{11}(\Theta)$
$\{\lambda_1, \lambda_3, \lambda_2\}$	$\pi_{12}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{12}(\Theta)$
$\{\lambda_3, \lambda_2, \lambda_1\}$	$\pi_{13}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_3)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{13}(\Theta)$
$\{\lambda_2, \lambda_3, \lambda_1\}$	$\pi_{13}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_3)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{13}(\Theta)$

Case (i): When β is known

Table 3.4: Reference priors for parameters, (β, Λ) , of $MOBGLD(\beta, \Lambda)$ and for parameters, (β, Θ) , of $MOBGLD(\beta, \Theta)$.

Ordered grouping	Reference Prior for (β, Λ)	Reference Prior for (β, Θ)
$\{\beta, (\lambda_1, \lambda_2, \lambda_3)\}$	$\pi_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{21}(\beta, \Theta)$
$\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$	$\pi_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{21}(\beta, \Theta)$
$\{(\beta, \lambda_1), (\lambda_2, \lambda_3)\}$	$\pi_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{21}(\beta, \Theta)$
$\{(\lambda_2, \lambda_3), \lambda_1, \beta\}$	$\pi_{22}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3) k(\lambda_1)]^{-1/2}$	$\omega_{22}(\beta, \Theta)$
$\{\lambda_3, \lambda_2, \lambda_1, \beta\}$	$\pi_{23}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_3) (1 - \lambda_2 - \lambda_3) k(\lambda_1)]^{-1/2}$	$\omega_{23}(\beta, \Theta)$
$\{(\beta, \lambda_1, \lambda_2, \lambda_3)\}$	$\pi_{24}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2) (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\omega_{24}(\beta, \Theta)$

The generalized form of reference prior for parameter Λ as derived in (3.20) is given by

$$\xi(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)^{l_1} (1 - \lambda_3)^{l_2} (1 - \lambda_2 - \lambda_3) k(\lambda_1)^{l_3}]^{-1/2}. \quad (3.23)$$

Hence for the expression of reference prior given in (3.23), we get the posterior density as follows:

$$\begin{aligned} p(\Lambda | \underline{t}) &\propto \xi(\Lambda) \cdot L(\Lambda | \underline{t}) \\ &\propto \lambda_1^{n-1} \lambda_2^{n_1 - \frac{1}{2}} \lambda_3^{n_2 - \frac{1}{2}} (1 - \lambda_2)^{-\frac{l_1}{2}} (1 - \lambda_3)^{-\frac{l_2}{2}} k(\lambda_1)^{-\frac{l_3}{2}} (1 - \lambda_2 - \lambda_3)^{n_0 - \frac{1}{2}} \\ &\exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^{\beta^*} \right\}, \end{aligned} \quad (3.24)$$

where β^* is a known value of β .

Case (ii): When β is unknown

The generalized form of reference prior for (β, Λ) as derived in (3.21) and (3.22) is given by

$$\xi(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)^{l_1} (1 - \lambda_3)^{l_2} (1 - \lambda_2 - \lambda_3) k(\lambda_1)^{l_3}]^{-1/2}. \quad (3.25)$$

In respect to reference prior (3.25), the posterior density will be as follows:

$$\begin{aligned} p(\beta, \Lambda | \underline{t}) &\propto \xi(\beta, \Lambda) \cdot L(\beta, \Lambda | \underline{t}) \\ &\propto \beta^{n-1} \lambda_1^{n-1} \lambda_2^{n_1 - \frac{1}{2}} \lambda_3^{n_2 - \frac{1}{2}} (1 - \lambda_2)^{-\frac{l_1}{2}} (1 - \lambda_3)^{-\frac{l_2}{2}} k(\lambda_1)^{-\frac{l_3}{2}} (1 - \lambda_2 - \lambda_3)^{n_0 - \frac{1}{2}} \end{aligned}$$

$$\prod_{j=1}^n \{[g(t_j)]^{\beta-1} g'(t_j)\} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\}. \quad (3.26)$$

The reference prior ‘ $\xi(\cdot)$ ’ defined in (3.23) and (3.25) can take different form of reference priors according to triplet (l_1, l_2, l_3) where, $l_i = 0$ or 1 ; $i = 1, 2, 3$. All considered reference priors with corresponding triplet indicators are given in Table 3.5 for both the case when the parameter β is known and when it is unknown.

Table 3.5: All reference priors for the parameter Λ when β is known and for (β, Λ) when β is unknown with corresponding triplets (l_1, l_2, l_3) .

β known			β unknown		
Triplet (l_1, l_2, l_3)	Reference Prior (Λ)	Reference Prior (Θ)	Triplet (l_1, l_2, l_3)	Reference Prior (β, Λ)	Reference Prior (β, Θ)
(0, 0, 0)	$\pi_{11}(\Lambda)$	$\omega_{11}(\Theta)$	(0, 0, 0)	$\pi_{21}(\beta, \Lambda)$	$\omega_{21}(\beta, \Theta)$
(1, 0, 0)	$\pi_{12}(\Lambda)$	$\omega_{12}(\Theta)$	(0, 0, 1)	$\pi_{22}(\beta, \Lambda)$	$\omega_{22}(\beta, \Theta)$
(0, 1, 0)	$\pi_{13}(\Lambda)$	$\omega_{13}(\Theta)$	(0, 1, 1)	$\pi_{23}(\beta, \Lambda)$	$\omega_{23}(\beta, \Theta)$
			(1, 0, 0)	$\pi_{24}(\beta, \Lambda)$	$\omega_{24}(\beta, \Theta)$

To obtain reference prior for original parameters (β, Θ) , we define the Jacobian matrix of the inverse transformation from (β, Λ) to (β, Θ) as

$$H = \frac{\partial(\beta, \Lambda)}{\partial(\beta, \Theta)} = \begin{bmatrix} \frac{\partial\beta}{\partial\beta} & \frac{\partial\beta}{\partial\theta_1} & \frac{\partial\beta}{\partial\theta_2} & \frac{\partial\beta}{\partial\theta_3} \\ \frac{\partial\lambda_1}{\partial\beta} & \frac{\partial\lambda_1}{\partial\theta_1} & \frac{\partial\lambda_1}{\partial\theta_2} & \frac{\partial\lambda_1}{\partial\theta_3} \\ \frac{\partial\lambda_2}{\partial\beta} & \frac{\partial\lambda_2}{\partial\theta_1} & \frac{\partial\lambda_2}{\partial\theta_2} & \frac{\partial\lambda_2}{\partial\theta_3} \\ \frac{\partial\lambda_3}{\partial\beta} & \frac{\partial\lambda_3}{\partial\theta_1} & \frac{\partial\lambda_3}{\partial\theta_2} & \frac{\partial\lambda_3}{\partial\theta_3} \end{bmatrix}.$$

The generalized form of reference prior for (β, Θ) can be obtained as

$$\Psi(\beta, \Theta) = \xi(\beta(\beta), \Lambda(\Theta)) |H|,$$

where $|\cdot|$ is determinant.

Theorem 3.4.2. The posterior density $p(\Lambda | \underline{t})$ under the reference prior $\pi(\Lambda)$ is proper for $n_0 > 0, n_1 > 0, n_2 > 0$ with at least two failures.

Proof. For a reference prior π , we have posterior density, such as

$$p(\Lambda \mid \underline{t}) \propto \pi(\Lambda) \cdot L(\Lambda \mid \underline{t})$$

To prove the posterior density proper, we have to show that it gives a finite value after integration, i.e., $\int p(\Lambda \mid \underline{t}) d\Lambda < \infty$. Let first prove for the posterior density $p_{11}(\Lambda \mid \underline{t})$ under reference prior $\pi_{11}(\Lambda)$. To find posterior density $p_{11}(\Lambda \mid \underline{t})$ consider the triplet $(0, 0, 0)$ in (3.24). Now we have to calculate

$$\begin{aligned} I_{11} = \int_0^\infty & \left[\iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1 - \frac{1}{2}} \lambda_3^{n_2 - \frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0 - \frac{1}{2}} d\lambda_2 d\lambda_3 \right] \\ & \lambda_1^{n-1} \left(\prod_{i=1}^n g'(t_j) \right) e^{-\lambda_1 \sum_{i=1}^n g(t_j)} d\lambda_1. \end{aligned}$$

If $\mathcal{B}(.,.)$ denotes the beta function, then we have

$$\mathcal{B}(a+1, b+c+2) \cdot \mathcal{B}(b+1, c+1) = \int_{0 < x+y < 1} x^a y^b (1-x-y)^c dx dy.$$

Now using this result, the above integral can be written, such as

$$\begin{aligned} I_{11} &= \mathcal{B}\left(n_1 + \frac{1}{2}, n_0 + n_2 + 1\right) \mathcal{B}\left(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}\right) \left(\prod_{i=1}^n g'(t_j) \right) \int_0^\infty \lambda_1^{n-1} e^{-\lambda_1 \sum_{i=1}^n g(t_j)} d\lambda_1 \\ &= \mathcal{B}\left(n_1 + \frac{1}{2}, n_0 + n_2 + 1\right) \mathcal{B}\left(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}\right) \left(\prod_{i=1}^n g'(t_j) \right) \frac{\Gamma(n)}{\left(\sum_{i=1}^n g(t_j) \right)^n} < \infty. \end{aligned}$$

Hence, the posterior density $p_{11}(\Lambda \mid \underline{t})$ is proper. Similar procedure can be used for other cases to prove them proper as well. \square

3.4.2 Posterior Estimation

The posterior analysis of the parameters can be done by using the conditional posterior distributions for different reference priors based on values of triplet (l_1, l_2, l_3) . We use Gibbs sampling method, a *MCMC* technique Chen et al. (2000), to obtain the Bayesian estimator

of the parameters. A step-by-step procedure is defined for obtaining the samples (estimates) for all parameters in both the cases and for all considered priors in Table 3.5.

Case (i): When β is known

The full conditional posterior in presence of known β are as follows:

$$p(\lambda_1 | \beta^*, data) \propto \lambda_1^{n-1} [k(\lambda_1)]^{-\frac{l_3}{2}} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^{\beta^*} \right\}, \quad (3.27)$$

$$p(\lambda_2, \lambda_3 | data) \propto \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1-\lambda_2)^{-\frac{l_1}{2}} (1-\lambda_2)^{-\frac{l_2}{2}} (1-\lambda_2-\lambda_3)^{n_0-\frac{1}{2}}. \quad (3.28)$$

- As in Table 3.5, we have three reference priors, π_{11} , π_{12} and π_{13} with $l_3 = 0$. So, in all three priors, we get exact gamma distribution to generate a sample for λ_1 , i.e., $\lambda_1 \sim \mathcal{G}\left(n, \sum_{j=1}^n [g(t_j)]^{\beta^*}\right)$.
- It is easy to see that all three reference priors π_{11} , π_{12} and π_{13} in Table 3.5 have different set of l_1 and l_2 . So, in respect to each reference prior a unique procedure needed to apply.
 - In case of reference prior π_{11} , the joint posterior density follows Dirichlet distribution, i.e., $(\lambda_2, \lambda_3) \sim \text{Dir}(n_1 + 1/2, n_2 + 1/2, n_0 + 1/2)$. Now if a rv $Y_i \sim \mathcal{G}(\alpha_i, 1); i = 1, 2, \dots, K$ and assume that $X_i = Y_i / \sum_{i=1}^K Y_i$, then $(X_1, X_2, \dots, X_{K-1}) \sim \text{Dir}(\alpha_1, \alpha_2, \dots, \alpha_K)$. Then it is obvious to write that $\lambda_2 = \frac{y_1}{y_1+y_2+y_3}; \lambda_3 = \frac{y_2}{y_1+y_2+y_3}$, where $y_1 \sim \mathcal{G}(n_1 + 1/2, 1)$, $y_2 \sim \mathcal{G}(n_2 + 1/2, 1)$ and $y_3 \sim \mathcal{G}(n_0 + 1/2, 1)$. It is quite easy to generate random number from y_1, y_2 and y_3 and so for λ_2 and λ_3 .
 - In case of the reference priors π_{12} , another transformation procedure can be applied. Let assume that $z_1 = \lambda_2$ and $z_2 = \lambda_3 / (1 - \lambda_2)$. Then it can be directly obtained from (3.24) that z_1 and z_2 are independent as well as both are following beta distribution, i.e., $Z_1 \sim \mathcal{B}(n_1 + 1/2, n_0 + n_2 + 0.5)$ and $Z_2 \sim \mathcal{B}(n_2 + 0.5, n_0 + n_2 + 0.5)$. So a sample can be easily generated from beta distribution, so as from λ_2 and λ_3 .
 - In case of π_{13} , assume that $z_1 = \lambda_3$ and $z_2 = \lambda_2 / (1 - \lambda_3)$, then by the similar

procedure as for π_{12} , samples can be obtained for λ_2 and λ_3 .

Case (ii): When β is unknown

The full conditional of the parameters can be obtained from (3.26) as follows:

$$p(\beta \mid \lambda_1, data) \propto \beta^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\} \prod_{j=1}^n [g(t_j)]^{\beta-1}, \quad (3.29)$$

$$p(\lambda_1 \mid \beta, data) \propto \lambda_1^{n-1} [k(\lambda_1)]^{-\frac{l_3}{2}} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\}, \quad (3.30)$$

$$p(\lambda_2, \lambda_3 \mid data) \propto \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1-\lambda_2)^{-\frac{l_1}{2}} (1-\lambda_2)^{-\frac{l_2}{2}} (1-\lambda_2-\lambda_3)^{n_0-\frac{1}{2}}. \quad (3.31)$$

- To obtain a posterior sample for β from (3.29), we use the M-H algorithm by considering normal distribution as a proposal density, such as the candidate point $\beta^c \sim \mathcal{N}(\hat{\beta}, Var(\hat{\beta}))$.
- As in Table 3.5, we consider four different reference priors, $\pi_{21}, \pi_{22}, \pi_{23}$ and π_{24} .

As for reference priors π_{21} and π_{24} , we have $l_3 = 0$ and hence a sample can be generated directly for λ_1 from $\mathcal{G}\left(n, \sum_{j=1}^n [g(t_j)]^\beta\right)$. But for the reference prior π_{22} and π_{23} , where $l_3 = 1$, we use M-H algorithm using gamma distribution as a proposal density such as candidate point $\lambda_1^c \sim \mathcal{G}\left(n, \sum_{j=1}^n [g(t_j)]^\beta\right)$. Here, it is to be noted that (β, λ_1) be a vector of parameters for which MCMC induced by the Gibbs sampler. First, we choose an arbitrary starting point (β^0, λ_1^0) and then proceed accordingly. So, as we collect the sample from both, it is easy to obtain the Bayes estimates.

- As no expression of β is presented in the joint posterior distribution of (λ_2, λ_3) hence the same steps will be followed in *case (ii)* as discussed earlier in *case (i)*.

The procedure defined above provides a general form of full conditional posteriors with respect to *MOBGLD*. For any particular member of the family, the same can be drawn by setting the value of $g(t)$ and the parameters accordingly.

3.5 Simulation Study

In this section, simulation studies have been performed for the Bayesian estimators under different reference priors of the unknown parameters. The performance of an estimator has been shown using *AB*, root mean square error (*RMSE*) and *CP*.

Under simulation study, we take sample size $n = 20, 30, 40$ and 50 to evaluate and compare the efficiency of the estimators derived in Section 3.3.2 and Section 3.4. All the simulation studies are repeated 1000 times to achieve consistency in the results. We have applied the given set-up for all five particular cases of *MOBGLD*. The results are shown for *MOBED*(0.87, 1.30, 1.50), *MOBRD*(0.33, 0.66, 0.76) and *MOBPD*(1.15, 1.65, 1.71) when $a = 0.50$, *MOBBD*(1.25, 1.64, 1.72) when $b = 1.50$ and *MOBWD*(0.75, 0.55, 1.25, 1.35) in Table 3.6, Table 3.7, Table 3.8, Table 3.9 and Table 3.10, respectively.

The important establishments from the simulated results are as follows:

- All the estimators are performing well for all the members of *MOBGLD*. Also, as the sample size is increasing, the *AB* and *RMSE* of estimators are gradually decreasing.
- Here, it is to be noted that no single reference prior is best for all the parameters of interest. It is because the formulation of reference priors is solely based on the inferential interest of the researcher. For example, it can be seen from the simulation table of *MOBED*, π_{13} performs better for λ_1 for larger n and π_{11} , π_{12} and π_{13} all perform well for λ_2 and λ_3 .
- For all the members of *MOBGLD*, the coverage probability of the Bayesian estimators is close to the nominal level.

The simulation study can be concluded that all estimators are performing good under the considered reference priors.

Table 3.6: AB, RMSE and CP of the Bayes estimators for MOBED(0.87, 1.30, 1.50).

n	π_{11}			π_{12}			π_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.7128	0.9432	0.9390	0.7101	0.9403	0.9410	0.7108	0.9420	0.9410
	λ_2	0.0779	0.0974	0.9420	0.0794	0.1001	0.9450	0.0789	0.0974	0.9420
	λ_3	0.0790	0.0973	0.9420	0.0800	0.0980	0.9510	0.0795	0.0994	0.9530
30	λ_1	0.5773	0.7428	0.9440	0.5777	0.7431	0.9460	0.5776	0.7439	0.9450
	λ_2	0.0684	0.0857	0.9550	0.0695	0.0873	0.9370	0.0683	0.0858	0.9370
	λ_3	0.0695	0.0867	0.9360	0.0700	0.0871	0.9450	0.0700	0.0883	0.9390
40	λ_1	0.4643	0.6056	0.9540	0.4651	0.6048	0.9540	0.4646	0.6051	0.9570
	λ_2	0.0602	0.0755	0.9420	0.0608	0.0766	0.9380	0.0605	0.0755	0.9400
	λ_3	0.0638	0.0789	0.9320	0.0639	0.0791	0.9320	0.0646	0.0800	0.9320
50	λ_1	0.4161	0.5358	0.9470	0.4156	0.5354	0.9520	0.4157	0.5357	0.9520
	λ_2	0.0540	0.0675	0.9370	0.0546	0.0682	0.9370	0.0539	0.0675	0.9370
	λ_3	0.0547	0.0681	0.9570	0.0547	0.0682	0.9430	0.0552	0.0687	0.9520

Table 3.7: AB, RMSE and CP of the Bayes estimators for MOBRD(0.33, 0.66, 0.76).

n	π_{11}			π_{12}			π_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.3257	0.4273	0.9520	0.3256	0.4276	0.9540	0.3255	0.4279	0.9500
	λ_2	0.0794	0.0989	0.9660	0.0817	0.1018	0.9470	0.0792	0.0992	0.9570
	λ_3	0.0824	0.1020	0.9640	0.0829	0.1031	0.9580	0.0845	0.1041	0.9550
30	λ_1	0.2623	0.3404	0.9470	0.2618	0.3402	0.9460	0.2624	0.3401	0.9460
	λ_2	0.0687	0.0859	0.9380	0.0693	0.0873	0.9390	0.0693	0.0864	0.9460
	λ_3	0.0704	0.0881	0.9480	0.0713	0.0887	0.9330	0.0710	0.0894	0.9440
40	λ_1	0.2201	0.2840	0.9530	0.2196	0.2832	0.9550	0.2199	0.2838	0.9520
	λ_2	0.0567	0.0707	0.9610	0.0574	0.0718	0.9600	0.0570	0.0707	0.9600
	λ_3	0.0593	0.0729	0.9640	0.0595	0.0733	0.9640	0.0597	0.0737	0.9630
50	λ_1	0.1970	0.2551	0.9530	0.1975	0.2557	0.9520	0.1979	0.2559	0.9560
	λ_2	0.0535	0.0666	0.9380	0.0544	0.0674	0.9410	0.0536	0.0666	0.9440
	λ_3	0.0573	0.0711	0.9280	0.0573	0.0713	0.9330	0.0581	0.0718	0.9410

3.6 Real Data Analysis

In this section, two data sets are analysed to illustrate the usefulness of the models we proposed in the earlier section. The first data is of prostate cancer, whereas the second data is of diabetic retinopathy study. Assuming the cause-dependent competing risks set-up is

Table 3.8: AB, RMSE and CP of the Bayes estimators for $MOBPD(1.15, 1.65, 1.71)$ when $a = 0.5$.

n	π_{11}			π_{12}			π_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.8690	1.1702	0.9460	0.8703	1.1702	0.9450	0.8686	1.1694	0.9430
	λ_2	0.0848	0.1033	0.9340	0.0859	0.1057	0.9350	0.0853	0.1035	0.9470
	λ_3	0.0796	0.0983	0.9650	0.0793	0.0985	0.9530	0.0817	0.1010	0.9480
30	λ_1	0.6966	0.9216	0.9510	0.6985	0.9249	0.9500	0.6981	0.9240	0.9510
	λ_2	0.0632	0.0796	0.9490	0.0647	0.0808	0.9570	0.0636	0.0796	0.9510
	λ_3	0.0653	0.0808	0.9530	0.0653	0.0808	0.9620	0.0662	0.0823	0.9550
40	λ_1	0.5999	0.7851	0.9510	0.6011	0.7854	0.9510	0.6008	0.7858	0.9460
	λ_2	0.0590	0.0725	0.9540	0.0597	0.0734	0.9580	0.0589	0.0727	0.9470
	λ_3	0.0590	0.0745	0.9500	0.0592	0.0744	0.9490	0.0594	0.0757	0.9490
50	λ_1	0.5219	0.6836	0.9430	0.5245	0.6852	0.9490	0.5231	0.6844	0.9440
	λ_2	0.0521	0.0653	0.9580	0.0523	0.0659	0.9450	0.0523	0.0655	0.9470
	λ_3	0.0519	0.0660	0.9410	0.0520	0.0658	0.9410	0.0529	0.0669	0.9390

Table 3.9: AB, RMSE and CP of the Bayes estimators for $MOBBD(1.25, 1.64, 1.72)$ when $b = 1.5$.

n	π_{11}			π_{12}			π_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.8751	1.2163	0.9470	0.8737	1.2165	0.9470	0.8741	1.2143	0.9470
	λ_2	0.0781	0.0983	0.9450	0.0793	0.1008	0.9480	0.0790	0.0983	0.9460
	λ_3	0.0793	0.0974	0.9640	0.0793	0.0978	0.9670	0.0810	0.1001	0.9580
30	λ_1	0.7103	0.9181	0.9410	0.7101	0.9164	0.9350	0.7097	0.9156	0.9390
	λ_2	0.0658	0.0827	0.9630	0.0671	0.0842	0.9510	0.0656	0.0827	0.9480
	λ_3	0.0679	0.0843	0.9420	0.0682	0.0844	0.9460	0.0686	0.0858	0.9440
40	λ_1	0.5927	0.7589	0.9450	0.5921	0.7579	0.9440	0.5929	0.7597	0.9440
	λ_2	0.0589	0.0720	0.9560	0.0593	0.0729	0.9540	0.0590	0.0721	0.9510
	λ_3	0.0575	0.0728	0.9490	0.0578	0.0728	0.9470	0.0588	0.0739	0.9440
50	λ_1	0.5312	0.7015	0.9490	0.5323	0.7016	0.9460	0.5319	0.7003	0.9450
	λ_2	0.0510	0.0628	0.9510	0.0516	0.0645	0.9500	0.0510	0.0640	0.9550
	λ_3	0.0546	0.0681	0.9320	0.0544	0.0681	0.9350	0.0553	0.0689	0.9360

suitable, the analysis of both data sets has been done using $MOBGLD$.

3.6.1 Prostate Cancer Data

The prostate cancer data set originally consists of 502 patients, reported by Byar and Green (1980). We only use the major causes of death related to patients: prostate cancer (127

Table 3.10: AB, RMSE and CP of the Bayes estimators for MOBWD(0.75, 0.55, 1.25, 1.35).

n	π_{21}			π_{22}			π_{23}			π_{24}		
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP
20	β	0.1174	0.1612	0.9080	0.1174	0.1613	0.9170	0.1175	0.1615	0.9170	0.1175	0.1613
	λ_1	0.9774	1.9253	0.9380	0.9898	2.0136	0.9310	0.9779	1.9728	0.9420	0.9726	1.8892
	λ_2	0.0828	0.1047	0.9910	0.0831	0.1051	0.9920	0.0794	0.1006	0.9460	0.0798	0.1015
	λ_3	0.0845	0.1069	0.9910	0.0845	0.1065	0.9930	0.0824	0.1035	0.9490	0.0812	0.1031
30	β	0.0935	0.1250	0.9210	0.0935	0.1248	0.9190	0.0936	0.1248	0.9160	0.0943	0.1260
	λ_1	0.6478	1.0392	0.9440	0.6476	1.0817	0.9440	0.6495	1.0458	0.9410	0.6554	1.0553
	λ_2	0.0677	0.0842	0.9990	0.0682	0.0071	0.9980	0.0654	0.0813	0.9640	0.0660	0.0822
	λ_3	0.0679	0.0850	0.9930	0.0674	0.0071	0.9930	0.0660	0.0829	0.9610	0.0653	0.0821
40	β	0.0746	0.0941	0.9270	0.0745	0.0940	0.9330	0.0745	0.0941	0.9270	0.0748	0.0946
	λ_1	0.4922	0.6631	0.9510	0.4902	0.6619	0.9550	0.4901	0.6608	0.9560	0.4943	0.6669
	λ_2	0.0579	0.0742	0.9920	0.0578	0.0741	0.9920	0.0560	0.0723	0.9430	0.0567	0.0729
	λ_3	0.0586	0.0744	0.9950	0.0587	0.0747	0.9950	0.0571	0.0727	0.9430	0.0569	0.0722
50	β	0.0688	0.0880	0.9250	0.0689	0.0882	0.9180	0.0686	0.0879	0.9230	0.0689	0.0883
	λ_1	0.4334	0.5806	0.9510	0.4346	0.5816	0.9510	0.4350	0.5820	0.9470	0.4370	0.5845
	λ_2	0.0546	0.0671	0.9950	0.0543	0.0668	0.9950	0.0529	0.0654	0.9600	0.0535	0.0657
	λ_3	0.0553	0.0690	0.9950	0.0549	0.0687	0.9960	0.0545	0.0678	0.9510	0.0541	0.0675

patients), cerebrovascular accident (31 patients) and other causes (24 patient), a sum of 182 patients (see Table 3.11). To make the data suitable to our model derived in Section 3.3, we initially grouped the patients as per their cause of death. The first group consists of those patients whose death is caused by prostate cancer or other causes; and the second group consists of those patients whose death is caused by cerebrovascular or other causes. One of the advantages of grouping the patients in such a way is that one can compare the risk of death from all causes except cerebrovascular and with all causes except prostate cancer.

For j^{th} patient, let X_{1j} and X_{2j} be the latent time of death (in months) due to group 1 and group 2 of causes, respectively. So that $T_j = \min(X_{1j}, X_{2j})$ be the observed death time of the j^{th} patient. Hence, if the patient is died of prostate cancer, then only X_{1j} is observed, if the patient is died of cerebrovascular accident, then only X_{2j} is observed and if the patient is died of other causes, then X_{1j} and X_{2j} both are observed simultaneously. In this data set,

Table 3.11: Patients summary statistics of prostate cancer data.

Cause of Death	n	Min	Q_1	Q_2	Mean	Q_3	Max	SD
Prostate cancer	127	2.00	12.00	24.00	26.26	37.00	71.00	17.35
Cerebrovascular	31	3.00	21.00	33.00	31.32	42.00	60.00	15.14
Other causes	24	1.00	8.25	20.00	25.75	35.50	74.00	22.79

time of death for patients is recorded in months. We analyse the data with failure times in years by dividing it by 12. Now, before performing any inferential analysis, we first check the nature of its hazard rate by using the total time on test (*TTT*) plot. It has been shown by Aarset (1987) that the scaled *TTT* transform is convex (concave) if the hazard rate is decreasing (increasing). The resulting *TTT* Figure 3.1 suggests to fit an increasing hazard rate distribution for this data. So we fit *MOBWD* and *MOBRD* distributions for the data and obtained the *AIC* and *BIC* values viz. 926.87 and 939.69 for *MOBWD* and 950.22 and 959.84 for *MOBRD*, respectively. Hence, we fit *MOBWD* as its *AIC* and *BIC* are minimum and estimates of the parameters of *MOBWD* are given in Table 3.12.

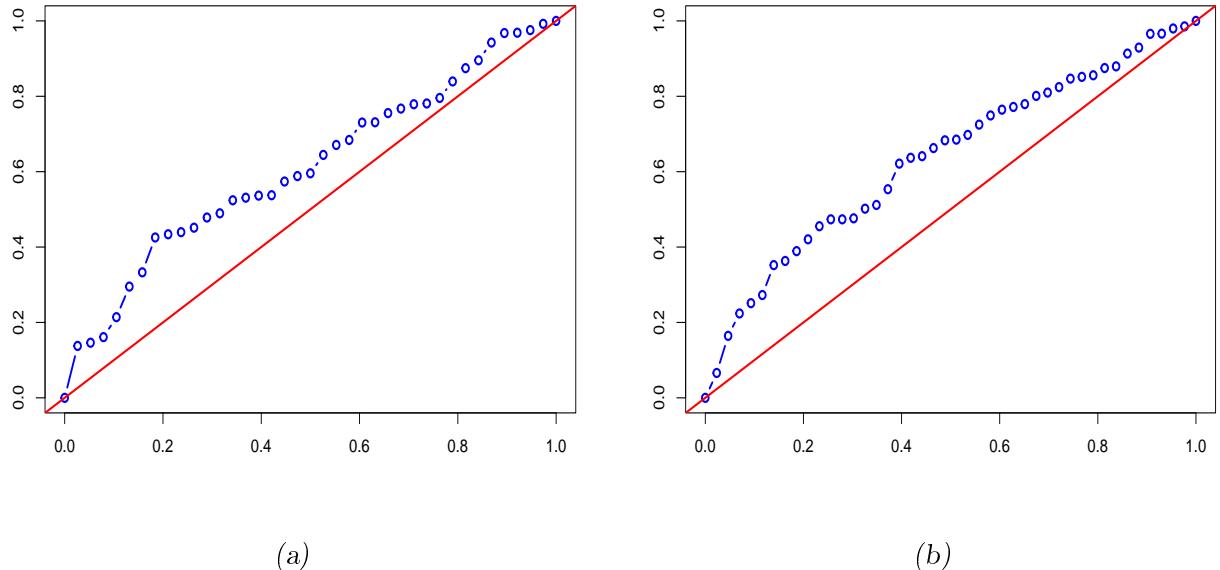


Figure 3.1: TTT plot for (a) X_1 (time to death of patients in group 1) and (b) X_2 (time to death of patients in group 2)

3.6.2 Diabetic Retinopathy Study Data

Diabetic Retinopathy study data belong to a case study conducted by the National Eye Institute from 1972 to 1978. Originally, the study was conducted on 1758 patients to assess the effect of laser treatment in delaying the onset of blindness in patients with diabetic retinopathy. For analysis purpose, we have considered only the white male patients who received the argon laser treatment. For more details about this data set given in Table

Table 3.12: Bayes estimates (95% HPD Interval) of parameters for *MOBWD* of prostate cancer data.

Method		β	λ_1	λ_2	λ_3
Bayes	Prior				
	π_{21}	1.5062 (1.3708, 1.6484)	0.2533 (0.1907, 0.3204)	0.6964 (0.5526, 0.8603)	0.1728 (0.1068, 0.2413)
	π_{22}	1.5015 (1.3558, 1.6309)	0.2551 (0.1963, 0.3193)	0.6961 (0.5457, 0.8700)	0.1730 (0.1145, 0.2398)
	π_{23}	1.5087 (1.3707, 1.6408)	0.2523 (0.1877, 0.3229)	0.6952 (0.6243, 0.7558)	0.1720 (0.1185, 0.2400)
	π_{24}	1.5096 (1.3710, 1.6350)	0.2527 (0.2015, 0.3090)	0.6944 (0.6254, 0.7658)	0.1718 (0.1195, 0.2234)

3.13, one can see Xu and Zhou (2017). The data set contains observation on X_1 , time to blindness of the eye under laser treatment and X_2 , time to blindness of the eye with no treatment. That is, we have minimum of time to blindness $T = \min(X_1, X_2)$ whereas there is a possibility that $X_1 = X_2$. In the data set, time of failure is given in days. We

Table 3.13: Time to blindness in days and causes for 71 white male diabetic patients.

Cause	Time
Eye under Laser Treatment	266, 583, 79, 93, 805, 344, 306, 415, 178, 1484, 315, 1252, 642, 407, 356, 699, 667, 126, 350, 84, 392, 901, 276, 520, 503, 584, 355, 1302
Eye without laser Treatment	91, 154, 547, 707, 469, 1313, 790, 125, 777, 307, 637, 577, 517, 287, 717, 141, 427, 36, 588, 350, 567, 1140, 448, 904, 485, 248, 423, 285, 315, 727, 210, 409, 227
Both Eyes	285, 622, 272, 1137, 1653, 471, 663, 966, 203, 1247

analyse the data with failure times in years by dividing it by 365. Now before performing any inferential analysis, we first check the nature of its hazard using the *TTT* plot. The resulting *TTT* Figure 3.2 suggest to fit a increasing hazard rate distribution. So, we fit *MOBWD* and *MOBRD* distributions to this data and obtained *AIC* and *BIC* values viz. 326.23 and 335.29 for *MOBWD* and 332.85 and 339.64 for *MOBRD*, respectively. Hence, we fit *MOBWD* as its *AIC* and *BIC* are minimum and estimates of the parameters of *MOBWD* are given in Table 3.14.

As given by Table 3.14, the probability of failure of the eye receiving laser treatment

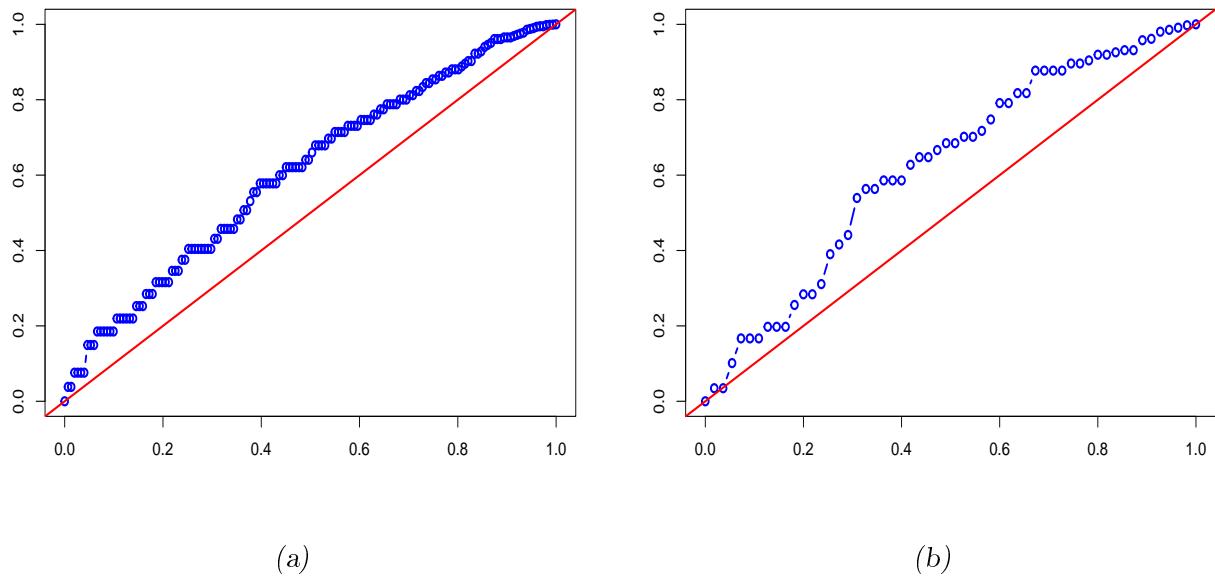


Figure 3.2: TTT plot for (a) time to blindness of eye under laser treatment observations, X_1 , and (b) time to blindness of the eye with no treatment observations, X_2

Table 3.14: Bayes estimates (95% HPD interval) of parameters for MOBWD of Diabetic Retinopathy study data.

Method	β	λ_1	λ_2	λ_3
Bayes	Prior			
π_{21}	1.5532 (1.3226, 1.8107)	0.4741 (0.3368, 0.6536)	0.3980 (0.2483, 0.5819)	0.4640 (0.3054, 0.6810)
π_{22}	1.5514 (1.2973, 1.8246)	0.4677 (0.3268, 0.6213)	0.4010 (0.2452, 0.5820)	0.4719 (0.3055, 0.7138)
π_{23}	1.5527 (1.2964, 1.8011)	0.4778 (0.3416, 0.6492)	0.3891 (0.2871, 0.4948)	0.4622 (0.3519, 0.5635)
π_{24}	1.5549 (1.3608, 1.7551)	0.4653 (0.3415, 0.6086)	0.3950 (0.2849, 0.5033)	0.4611 (0.3467, 0.5674)

is less than that of the eye with no treatment, i.e., $\lambda_2 < \lambda_3$, we may consider that laser treatment is quite helpful in delaying the onset of blindness. Moreover, from Figure 3.3, it is clear that the survival probability of the eye receiving laser treatment is greater than that of the eye with no treatment throughout the time. Hence, the survival is more favourable to the eye receiving laser treatment. Since one may be interested in the time to blindness in the eye receiving laser treatment, i.e., in the parameter λ_3 , so for Bayesian inference of the unknown parameters, one can choose the reference prior π_{23} , which gives more weight

to the parameter λ_3 .

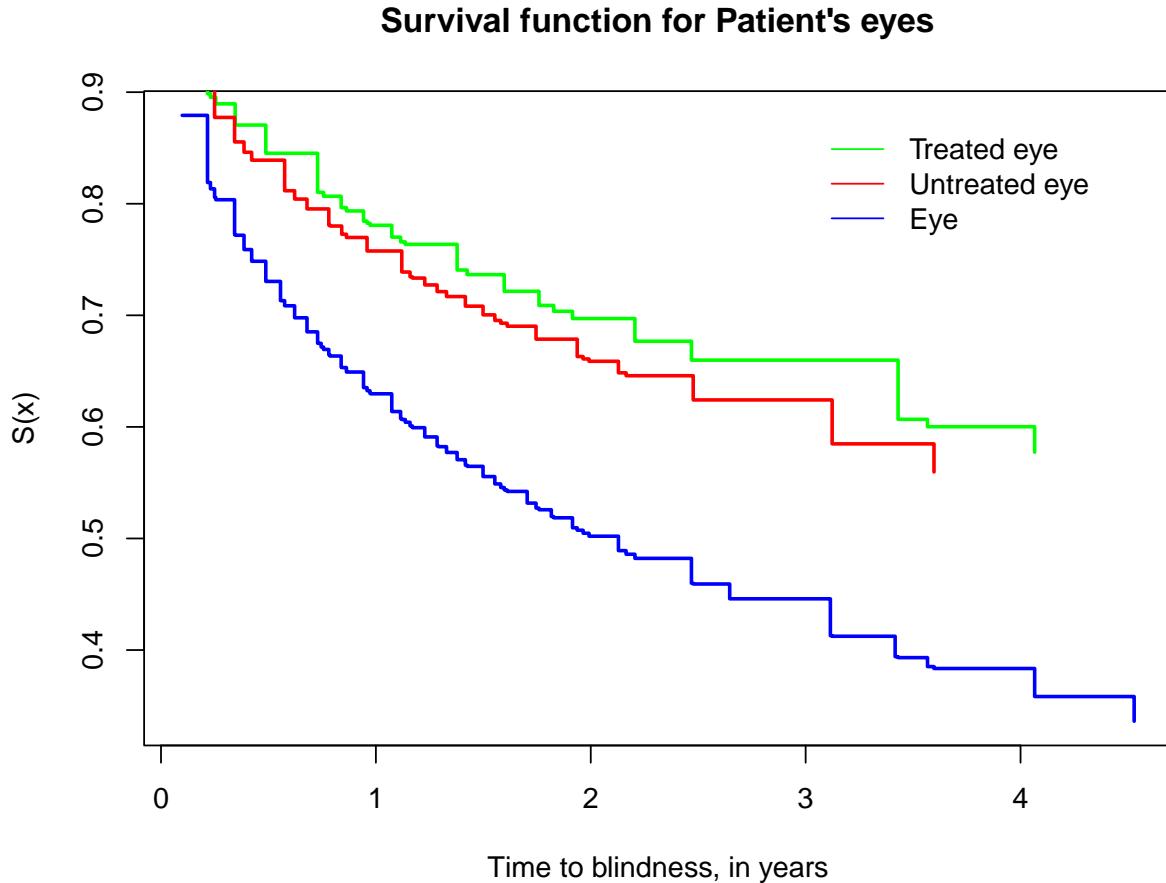


Figure 3.3: Survival function estimate for DRS data

3.7 Conclusion

In this chapter, we have proposed a novel bivariate distribution based on a generalized family of lifetime distribution using the Marshall-Olkin set-up. The proposed distribution is utilized to analyse the competing risks data with cause dependency. The dependency between causes originated due to the non-zero probability of simultaneous occurrence of the causes. We also obtained the bivariate version of the Marshall-Olkin distribution for each member of the generalized family of lifetime distribution. The estimation is carried out for the transformed set of parameters. Under the Bayesian paradigm, different sets of ordered parameters are considered and correspondingly the reference prior for each set of ordered

parameters is derived. Further, the propriety of obtained posteriors is determined under each reference prior. The Bayesian estimation was done using reference priors and classical estimators are compared with their Bayesian counterparts in terms of AB , $RMSE$ and CP . Based on these statistical measures, we can choose a suitable reference prior to our set parameters based on the inferential interest. To demonstrate the performance of estimators, a simulation study has been conducted for each member of the $MOBGLD$ with competing risks data set-up. In the end, two real datasets are utilized to illustrate the applicability of the established model.

Although in this chapter, the work is done for $MOBGLD$ competing risks data in the case of a complete sample, it can be further extended for competing risks data under different censoring schemes, such as type-I censoring, type-II censoring, random censoring, progressively hybrid censoring, etc. and also for incomplete data problem.

Chapter 4

Objective Bayesian Analysis of Cause-Dependent Competing Risks Data with Partial Information

4.1 Introduction

In life-testing experiments, the main concern lies in the unknown parameters of distribution assumed for the variable under study; e.g., in testing the light bulbs, researchers are interested in the mean life length, rate of failure of the bulbs or shape of the distribution, and so forth. Hence, the unknown parameters of the model are always keen of interest. Thus it becomes inevitable to perform statistical inferences to elicit more information about these parameters. The inferential procedure can be carried out using two major phenomena in the statistical literature- classical inference and Bayesian inference. The latter phenomenon considers the unknown parameter as a random variable, unlike in classical inference, which considers it as an unknown but fixed quantity. In the Bayesian approach, prior beliefs about the unknown parameter can be formulated with a specific probability distribution whenever some prior information is available. After observing the data, the prior distribution is nested with the likelihood to give the posterior distribution of the parameter. Hence, the prior belief about the parameter is updated in light of observed data. Moreover, this posterior information can be utilized as prior information for future data. Based on the prior information available with the unknown parameters, the Bayesian analysis is done either with informative priors or non-informative priors. If the prior information about the

parameter of interest is sufficiently available, the Bayesian analysis is preferred under informative priors such as conjugate priors. However, when the information available is vague or inadequate, non-informative priors are used. For a detailed discussion on the selection of priors under the objective Bayesian paradigm, one can see Ghosh and Mukerjee (1992), Kass and Wasserman (1996) and Ghosh (2011).

In addition, reference priors are one of the non-informative priors, which maximizes the divergence between prior and posterior density asymptotically. Reference prior works in such a way that the impact of the prior diffuses, and posterior inference becomes data-driven. Bernardo (1979) derived the reference priors and later Berger and Bernardo (1992) extended it giving an algorithm for the derivation of reference priors for the parameters of a number of distributions. Apart from all the above, sometimes a situation arises where adequate information is available for some parameters and the information is missing or vague for other parameters; e.g., for a $\mathcal{N}(\mu, \sigma^2)$ population, it is possible that one has prior information about the unknown parameter σ^2 , but there is no information about the mean parameter μ . In such a case of partial prior information available with some of the parameters, the marginal or conditional reference prior becomes indispensable to be derived for Bayesian analysis as suggested by Sun and Berger (1998).

In time-to-event analysis, the event of interest may be the death of an individual, relapse of a disease after remission, failure of a system, etc. It is often the case that the event of interest may occur with multiple endpoints. This phenomenon is known as competing risks in the statistical literature. While many authors considered competing risks to be independent, resulting in mathematical treatment much easier. However, in most situations, the independence assumption is quite susceptible. One of the available methods to model the dependency among competing risks is a multivariate probability distribution. Marshall and Olkin (1967) introduced one of the multivariate probability distribution models when there is a non-zero probability of simultaneous occurrence of two causes of failure. Shen and Xu (2018) analysed the competing risks data for dependent causes considering Marshall-Olkin bivariate Weibull distribution under three different methods. While Zhang et al. (2022) obtained the system reliability under the Bayesian paradigm for a multicomponent stress-

strength model assuming *MOBWD*. Recently, Barnwal and Panwar (2022) considered the *MOBGLD* to model the dependency between competing risks and performed the classical and Bayesian inference with reference priors derived for different ordered grouping of the parameters of interest. Also, Kundu and Gupta (2013) performed the Bayes estimation of the parameters of Marshall-Olkin bivariate Weibull distribution considering the Gamma prior for scale parameters. Objective Bayesian analysis with partial information has been done by several authors for parameters of univariate distributions in different scenarios. Xu and Tang (2011) performed Bayesian analysis for Birnbaum–Saunders’s distribution with partial information for a type-II censored sample and proved the propriety of the obtained posterior. Further, Seo and Kim (2017) analysed the upper record value data under the Bayesian paradigm using partial information for unknown parameters of Rayleigh Distribution. Moreover, Seo (2020) performed the same for Weibull distribution with partial information for generalized type-II progressive hybrid censored data.

The novelty of the work is that objective Bayesian inference is performed using reference prior with partial information for competing risks data using *MOBWD*. The purpose of this chapter is twofold: firstly, we are concerned with the changes in the behaviour of Bayesian estimators when information is available on some of the parameters i.e., in case of partial prior information; and secondly to perform the testing of the hypothesis under the Bayesian paradigm using Bayes factor for testing whether the two risks are identical or not. In Section 4.2, we have introduced the well-known *MOBWD* along with its joint survival and joint density function. The likelihood formulation in the classical framework for the unknown parameters has been given in Section 4.3. In Section 4.4, the reference priors for unknown parameters of *MOBWD* have been derived using partial prior information available with parameters and propriety of the posterior has been proved. The simulation study has been performed to illustrate the performance of estimators derived in Section 4.5. Finally, two real data sets have been analysed in Section 4.6.

4.2 Preliminaries

Let a lifetime random variable X follow the Weibull distribution with scale parameter α and shape parameter β . Consider the *pdf* and survival function of $\mathcal{W}(\beta, \alpha)$ as follows

$$f(x; \beta, \alpha) = \alpha\beta x^{\beta-1} \exp(-\alpha x^\beta) \quad (4.1)$$

$$S(x; \beta, \alpha) = \exp(-\alpha x^\beta); \quad x, \beta, \alpha > 0. \quad (4.2)$$

Assume three independent random variables U, V and W each following Weibull distribution with the same shape parameter β and scale parameters α_i ; $i = 0, 1, 2$. After defining random variables $X_1 = \min(U, W)$ and $X_2 = \min(V, W)$, the bivariate random vector (X_1, X_2) is said to follow the well-known *MOBWD*.

If $(X_1, X_2) \sim MOBWD(\beta, \alpha_0, \alpha_1, \alpha_2)$, then the following results are easy to establish

(i) $X_1 \sim \mathcal{W}(\beta, \alpha_0 + \alpha_1)$ and $X_2 \sim \mathcal{W}(\beta, \alpha_0 + \alpha_2)$;

(ii) $P(X_1 = X_2) = \frac{\alpha_0}{\alpha}$, $P(X_1 > X_2) = \frac{\alpha_2}{\alpha}$, $P(X_1 < X_2) = \frac{\alpha_1}{\alpha}$ and

(iii) $\min\{X_1, X_2\} \sim \mathcal{W}(\beta, \alpha)$, where $\alpha = \alpha_0 + \alpha_1 + \alpha_2$.

The joint survival function of (X_1, X_2) for *MOBWD* can also be defined such as (Kundu and Dey 2009)

$$S(x_1, x_2) = \begin{cases} S_1(x_1, x_2) = S(x_1; \beta, \alpha_0 + \alpha_1)S(x_2; \beta, \alpha_2) & \text{if } x_1 > x_2, \\ S_2(x_1, x_2) = S(x_2; \beta, \alpha_0 + \alpha_2)S(x_1; \beta, \alpha_1) & \text{if } x_2 > x_1, \\ S_0(x_1, x_2) = S(x; \beta, \alpha) & \text{if } x_1 = x_2 = x. \end{cases} \quad (4.3)$$

Similarly, the joint pdf for (X_1, X_2) for *MOBWD* is given by (Kundu and Dey 2009)

$$f(x_1, x_2) = \begin{cases} f_1(x_1, x_2) = f(x_1; \beta, \alpha_0 + \alpha_1)f(x_2; \beta, \alpha_2) & \text{if } x_1 > x_2, \\ f_2(x_1, x_2) = f(x_2; \beta, \alpha_0 + \alpha_2)f(x_1; \beta, \alpha_1) & \text{if } x_2 > x_1, \\ f_0(x_1, x_2) = \frac{\alpha_0}{\alpha} f(x; \beta, \alpha) & \text{if } x_1 = x_2 = x. \end{cases} \quad (4.4)$$

4.3 Likelihood Construction for Dependent Competing Risks Model

In a life-testing experiment, assume that n independent series systems with two components are being tested. The notion of component independence is untenable in this case. Let $T_j = \min\{X_{1j}, X_{2j}\}$ be failure time for the j^{th} series system ; $j = 1, 2, \dots, n$; where X_{ij} ; ($i = 1, 2$) denotes the failure of j^{th} system due to i^{th} component and components are assumed to be dependent. The dependence arises from the simultaneous failure of both components. The observed data is of the form $\{T_j, \delta_{1j}, \delta_{2j}\}$; where δ_{ij} is an indicator function which takes the value 1 if j^{th} system's failure is caused by i^{th} component and value 0 if system's failure is caused by both the components. Hence, failure of the system may be caused by simultaneous failure of both components with a non-zero probability. Now further assume that (X_{1j}, X_{2j}) follows *MOBWD*, i.e. $(X_{1j}, X_{2j}) \sim MOBWD(\beta, \alpha_0, \alpha_1, \alpha_2)$.

To carry out inference for the unknown parameters, it is always an interest to infer about the characteristics which are more important than the parameters themselves. In competing risks analysis, the researcher is more interested in the cause of failure, which is more severe than other causes. So, a particular set after re-parametrization has been considered such as, say,

$$\lambda_1 = \alpha = \alpha_0 + \alpha_1 + \alpha_2, \quad \lambda_2 = \frac{\alpha_1}{\alpha}, \quad \lambda_3 = \frac{\alpha_2}{\alpha}.$$

The re-parametrized quantities $\lambda_2 = P(X_1 < X_2)$ and $\lambda_3 = P(X_2 < X_1)$ represent the relative risk due to cause 1 and 2 respectively. The transformation $(\beta, \Lambda) = (\beta, \lambda_1, \lambda_2, \lambda_3)$ from $(\beta, \alpha_0, \alpha_1, \alpha_2)$ is one-to-one with the inverse transformation where

$$\beta = \beta, \quad \alpha_1 = \lambda_1 \lambda_2, \quad \alpha_2 = \lambda_1 \lambda_3, \quad \alpha_0 = \lambda_1(1 - \lambda_2 - \lambda_3).$$

Now the likelihood function under re-parametrization is given by

$$L(\beta, \Lambda | \underline{t}) = \beta^n \lambda_1^n \lambda_2^{n_1} \lambda_3^{n_2} (1 - \lambda_2 - \lambda_3)^{n_0} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}, \quad (4.5)$$

where n_0, n_1 and n_2 are the numbers of systems which have failed due to the failure of both components, the first component and the second component respectively. Symbolically, $\sum_{j=1}^n I(X_{1j} = X_{2j}) = \sum_{j=1}^n \delta_{1j}\delta_{2j} = n_0$, $\sum_{j=1}^n I(X_{1j} < X_{2j}) = \sum_{j=1}^n \delta_{1j}(1 - \delta_{2j}) = n_1$ and $\sum_{j=1}^n I(X_{1j} > X_{2j}) = \sum_{j=1}^n \delta_{2j}(1 - \delta_{1j}) = n_2$. Here $n = n_0 + n_1 + n_2$, as it is assumed that no information is censored or missing in the experiment.

The *ML* estimators of the unknown parameters (β, Λ) can be obtained by taking the partial derivative of the natural logarithm of the likelihood function (4.5) with respect to (β, Λ) and equating them to zero. In a similar way, the Fisher information matrix, say Σ can be obtained as

$$\Sigma = -E \left[\frac{\partial^2 \log L(\beta, \Lambda | \underline{t})}{\partial(\beta, \Lambda)^2} \Big| (\beta, \Lambda) \right].$$

Hence, the Fisher information matrix has the following form:

$$\Sigma = n \begin{bmatrix} \frac{k(\lambda_1)}{\beta^2} & \frac{r_1+1-\log \lambda_1}{\beta \lambda_1} & 0 & 0 \\ \frac{r_1+1-\log \lambda_1}{\beta \lambda_1} & \frac{1}{\lambda_1^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}, \quad (4.6)$$

where $k(\lambda_1) = r_2 + 2r_1(1 - \log \lambda_1) - \log \lambda_1(2 - \log \lambda_1) + 1$, and

$$r_u = \int_0^\infty (\log z)^u \exp\{-z\} dz; \quad u = 1, 2.$$

The present form of the likelihood function (4.5) is convenient for the posterior analysis. Hence, we utilize the likelihood function in the next section to establish the posterior density and the Bayesian estimates of the unknown parameters (β, Λ) of the dependent competing risks model will be obtained accordingly.

4.4 Bayesian Estimation

The Bayesian inference for unknown parameters has become inevitable in life-testing experiments. To update the information about parameters in light of the observed data, the prior information is nested with the likelihood function. If one has only partial prior information

on some of the parameters and the information available to other parameters is vague, then the Bayesian estimation is carried out with partial prior information available as suggested by Sun and Berger (1998). In case when the available information about the parameter is vague, one of the most popular priors is the reference prior proposed by Berger and Bernardo (1992). The idea behind reference prior is that it maximizes the Kullback-Leibler divergence between the prior and posterior density asymptotically. Consequently, the posterior inference becomes data-dominated and the effect of the prior is diffused. We derive reference prior for the unknown parameter of interest using the partial prior information available with the other parameter as proposed by Sun and Berger (1998).

Suppose one is interested in the ordered group (θ_1, θ_2) of parameters, where parameters θ_1 and θ_2 are of dimensions d_1 and d_2 respectively.

- (i) Given the marginal subjective prior $\pi^s(\theta_2)$, the conditional reference prior for θ_1 given θ_2 , say $\pi^r(\theta_1|\theta_2)$ can be derived.
- (ii) Given the subjective conditional prior $\pi^s(\theta_2|\theta_1)$, the marginal reference prior for θ_1 is given by Theorem 1 of Sun and Berger (1998) as follows

$$\pi^r(\theta_1) \propto \left(\frac{\det \Sigma}{\det \Sigma_{22}} \right)^{1/2}, \quad (4.7)$$

where Σ is the Fisher Information matrix for (θ_1, θ_2) and Σ_{22} is Fisher information matrix for θ_2 for fixed θ_1 .

4.4.1 Reference Prior with Partial Information

In this subsection, we derive reference prior in the case when partial information is available to some of the unknown parameters. Also, we establish that the derived reference prior leads to the proper posterior density.

- (i) *When partial prior information is available for parameter λ_1 only*

Assume that parameter λ_1 has a Gamma prior, say $\mathcal{G}(a, b)$ given by

$$\pi_1^s(\lambda_1 | a, b) \propto \lambda_1^{a-1} e^{-b\lambda_1}; \quad \lambda_1 > 0; \quad a, b > 0. \quad (4.8)$$

Here a and b are the hyper-parameters specifying the prior distribution of λ_1 . The choice of hyper-parameters is very important to carry out the Bayesian inference.

Now to obtain the joint prior for (β, Λ) , rearranging the rows and columns of matrix Σ given in (4.6), we get

$$\Sigma_1 = n \begin{bmatrix} \frac{1}{\lambda_1^2} & \frac{r_1+1-\log \lambda_1}{\beta \lambda_1} & 0 & 0 \\ \frac{r_1+1-\log \lambda_1}{\beta \lambda_1} & \frac{k(\lambda_1)}{\beta^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}. \quad (4.9)$$

The determinant of matrix Σ_1 in (4.9) is

$$\det \Sigma_1 = \frac{n}{\beta^2} \frac{k(\lambda_1)}{\lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)},$$

which can be written as

$$\det \Sigma_1 \propto g_1(\theta_1) \cdot g_2(\theta_2), \quad (4.10)$$

where

$$g_1(\theta_1) = k(\lambda_1) \text{ and } g_2(\theta_2) = \frac{1}{\beta^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)}.$$

Then according to Theorem 2 given in Sun and Berger (1998), the conditional reference prior for $\theta_2 = (\beta, \lambda_2, \lambda_3)$ given $\theta_1 = \lambda_1$ is

$$\begin{aligned} \pi_1^r(\theta_2 | \theta_1) &\propto \sqrt{g_2(\theta_2)} \\ &\propto \beta^{-1} (\lambda_2 \lambda_3)^{-1/2} (1 - \lambda_2 - \lambda_3)^{-1/2}. \end{aligned} \quad (4.11)$$

The joint prior for (β, Λ) is given by

$$\begin{aligned} \xi_1(\beta, \Lambda) &\propto \pi_1^r(\beta, \lambda_2, \lambda_3 | \lambda_1) \cdot \pi_1^s(\lambda_1 | a, b) \\ &\propto \beta^{-1} \lambda_1^{a-1} e^{-b\lambda_1} (\lambda_2 \lambda_3)^{-1/2} (1 - \lambda_2 - \lambda_3)^{-1/2}. \end{aligned} \quad (4.12)$$

In respect to the prior (4.12), the posterior density will be as follows:

$$\begin{aligned}
 p_1(\beta, \Lambda | \underline{t}) &\propto \xi_1(\beta, \Lambda) \cdot L(\beta, \Lambda | \underline{t}) \\
 &\propto \beta^{n-1} \lambda_1^{n+a-1} \lambda_2^{(n_1-\frac{1}{2})} \lambda_3^{(n_2-\frac{1}{2})} (1 - \lambda_2 - \lambda_3)^{(n_0-\frac{1}{2})} \\
 &\cdot \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\}.
 \end{aligned} \tag{4.13}$$

(ii) When partial prior information is available for parameters (λ_2, λ_3) only

Given (β, λ_1) , the conditional subjective prior for (λ_2, λ_3) has a Dirichlet prior, say $Dir(a_0, a_1, a_2)$ given by

$$\pi_2^s(\lambda_2, \lambda_3 | \lambda_1, a_0, a_1, a_2) \propto \lambda_2^{a_1-1} \lambda_3^{a_2-1} (1 - \lambda_2 - \lambda_3)^{a_0-1}; \quad a_i > 0; \forall i = 0, 1, 2. \tag{4.14}$$

The partition of matrix Σ is given by

$$\Sigma_{11} = n \begin{bmatrix} \frac{k(\lambda_1)}{\beta^2} & \frac{r_1+1-\log \lambda_1}{\beta \lambda_1} \\ \frac{r_1+1-\log \lambda_1}{\beta \lambda_1} & \frac{1}{\lambda_1^2} \end{bmatrix}, \tag{4.15}$$

and

$$\Sigma_{22} = n \begin{bmatrix} \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix} \tag{4.16}$$

The determinant of Σ_{11} is

$$\det \Sigma_{11} = \frac{n(r_2 - r_1^2)}{\beta^2 \lambda_1^2}.$$

The determinant of Σ_{22} is

$$\det \Sigma_{22} = \frac{n}{\lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)}.$$

The determinant of Σ is

$$\det \Sigma = \det \Sigma_{11} \cdot \det \Sigma_{22}$$

Since the quantity $\frac{\det \Sigma}{\det \Sigma_{22}}$ does not depend on (λ_2, λ_3) , the marginal reference prior for

(β, λ_1) is given by using (4.7)

$$\pi_2^r(\beta, \lambda_1) \propto \frac{1}{\beta \lambda_1}. \quad (4.17)$$

The joint prior for (β, Λ) is given by

$$\begin{aligned} \xi_2(\beta, \Lambda) &\propto \pi_2^r(\beta, \lambda_1) \cdot \pi_2^s(\lambda_2, \lambda_3 \mid \lambda_1, a_0, a_1, a_2) \\ &\propto (\beta \lambda_1)^{-1} \lambda_2^{a_1-1} \lambda_3^{a_2-1} (1 - \lambda_2 - \lambda_3)^{a_0-1}. \end{aligned} \quad (4.18)$$

In respect to the prior (4.18), the posterior density will be as follows:

$$\begin{aligned} p_2(\beta, \Lambda \mid \underline{t}) &\propto \xi_2(\beta, \Lambda) \cdot L(\beta, \Lambda \mid \underline{t}) \\ &\propto \beta^{n-1} \lambda_1^{n-1} \lambda_2^{(n_1+a_1-1)} \lambda_3^{(n_2+a_2-1)} (1 - \lambda_2 - \lambda_3)^{(n_0+a_0-1)} \\ &\cdot \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}. \end{aligned} \quad (4.19)$$

4.4.2 Posterior Propriety

In this subsection, the propriety of posterior densities $p_1(\beta, \Lambda \mid \underline{t})$ and $p_2(\beta, \Lambda \mid \underline{t})$ obtained under the joint priors $\xi_1(\beta, \Lambda)$ and $\xi_2(\beta, \Lambda)$ respectively, has been established.

Theorem 4.4.1. The posterior density $p_1(\beta, \Lambda \mid \underline{t})$ under the joint prior $\xi_1(\beta, \Lambda)$ is proper for $n_0 > 0, n_1 > 0, n_2 > 0$ with at least two failures.

Proof. The posterior density $p_1(\beta, \Lambda \mid \underline{t})$ is given by (4.13). To prove the posterior density proper, we have to show that it gives a finite value after integration i.e.

$$\int p_1(\beta, \Lambda \mid \underline{t}) d\beta d\Lambda < \infty.$$

We have to calculate the following integral

$$\begin{aligned} I^1 &= \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}} d\lambda_2 d\lambda_3 \right\} \right. \\ &\quad \left. \lambda_1^{n+a-1} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta. \end{aligned}$$

Now using the beta function, the above integral can be written such as

$$\begin{aligned} I^1 &= A_1 \int_0^\infty \left[\int_0^\infty \lambda_1^{n+a-1} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ &= C_1 \int_0^\infty \beta^{n-1} \frac{\prod_{j=1}^n t_j^\beta}{\left(b + \sum_{i=1}^n t_i^\beta \right)^{n+a}} d\beta \leq C_1 \int_0^\infty \beta^{n-1} \prod_{i=1}^n \left(\frac{t_i}{b + t_{(n)}} \right)^\beta d\beta < \infty, \end{aligned}$$

where $t_{(n)} = \max_{1 \leq i \leq n} t_i$, $A_1 = \mathcal{B}(n_1 + \frac{1}{2}, n_0 + n_2 + 1) \cdot \mathcal{B}(n_2 + \frac{1}{2}, n_0 + \frac{1}{2})$ and $C_1 = A_1 \Gamma(n+a) \left(\prod_{j=1}^n t_j \right)^{-1}$. The last inequality holds because $\prod_{i=1}^n \frac{t_i}{b+t_{(n)}} < 1$. Hence the posterior density $p_1(\beta, \Lambda | \underline{t})$ is proper. \square

Theorem 4.4.2. The posterior density $p_2(\beta, \Lambda | \underline{t})$ under the joint prior $\xi_2(\beta, \Lambda)$ is proper for $n_0 > 0, n_1 > 0, n_2 > 0$ with at least two failures.

Proof. The posterior density $p_2(\beta, \Lambda | \underline{t})$ is given by (4.19). To prove the posterior density proper, we have to show that it gives a finite value after integration i.e. $\int p_2(\beta, \Lambda | \underline{t}) d\beta d\Lambda < \infty$.

We have to calculate the following integral

$$\begin{aligned} I^* &= \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1+a_1-1} \lambda_3^{n_2+a_2-1} (1 - \lambda_2 - \lambda_3)^{n_0+a_0-1} d\lambda_2 d\lambda_3 \right\} \right. \\ &\quad \left. \lambda_1^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta. \end{aligned}$$

Now using the beta function, the above integral can be written such as

$$\begin{aligned} I &= A_2 \int_0^\infty \left[\int_0^\infty \lambda_1^{n-1} e^{-\lambda_1 \sum_{i=1}^n t_i^\beta} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ &= C_2 \int_0^\infty \beta^{n-1} \frac{\prod_{j=1}^n t_j^\beta}{\left(\sum_{i=1}^n t_i^\beta \right)^n} d\beta \leq C_2 \int_0^\infty \beta^{n-1} \prod_{i=1}^n \left(\frac{t_i}{t_{(n)}} \right)^\beta d\beta < \infty, \end{aligned}$$

where $t_{(n)} = \max_{1 \leq i \leq n} t_i$, $A_2 = \mathcal{B}(n_1+a_1, n_0+n_2+a_0+a_2) \cdot \mathcal{B}(n_2+a_2, n_0+a_0)$ and $C_2 = A_2 \Gamma(n) \left(\prod_{j=1}^n t_j \right)^{-1}$. The last inequality holds because $\prod_{i=1}^n \frac{t_i}{t_{(n)}} < 1$. Hence the posterior density $p_2(\beta, \Lambda | \underline{t})$ is proper. \square

4.4.3 Posterior Estimation

The posterior analysis of the unknown parameters has been done using conditional distributions. Since the obtained distributions are not in standard form, we use the Gibbs sampling method, a MCMC technique (Chen et al. 2000).

The full conditionals of the parameters can be obtained using (4.13) as follows:

$$p_3(\beta | \lambda_1, \text{data}) \propto \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\}, \quad (4.20)$$

$$p_4(\lambda_1 | \beta, \text{data}) \propto \lambda_1^{n+a-1} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\}, \quad (4.21)$$

$$p_5(\lambda_2, \lambda_3 | \text{data}) \propto \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}}. \quad (4.22)$$

Also, the full conditionals of the parameters using (4.19) are:

$$p_6(\beta | \lambda_1, \text{data}) \propto \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}, \quad (4.23)$$

$$p_7(\lambda_1 | \beta, \text{data}) \propto \lambda_1^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}, \quad (4.24)$$

$$p_8(\lambda_2, \lambda_3 | \text{data}) \propto \lambda_2^{n_1+a_1-1} \lambda_3^{n_2+a_2-1} (1 - \lambda_2 - \lambda_3)^{n_0+a_0-1}. \quad (4.25)$$

To obtain the posterior samples on the parameters from the posteriors given in (4.13) and (4.19), we follow the procedure given below:

- The posterior sample for β from (4.20) and (4.23) can be obtained using M-H algorithm by considering normal distribution as a proposal density such as the candidate point $\beta^c \sim \mathcal{N}(\hat{\beta}, \text{Var}(\hat{\beta}))$.
- From (4.21) and (4.24), it is clear that a sample can be generated directly for λ_1 from $\mathcal{G}\left(n, \sum_{j=1}^n t_j^\beta\right)$ and $\mathcal{G}\left(n+a, (b + \sum_{j=1}^n t_j^\beta)\right)$. The values of hyper-parameters a and b are chosen arbitrarily. Here it is to be noted that (β, λ_1) is a vector of parameters for which MCMC is induced by the Gibbs sampler. First, we choose an arbitrary starting point (β^0, λ_1^0) and then proceed accordingly. So, as we collect the sample from both it

is easy to obtain the Bayes estimates.

- The joint posterior density follows Dirichlet distribution i.e. $(\lambda_2, \lambda_3) \sim \text{Dir}(n_1 + a_1, n_2 + a_2, n_0 + a_0)$. Now if a rv $Y_i \sim \mathcal{G}(a_i, 1); i = 1, 2, \dots, K$ and assume that $X_i = Y_i / \sum_{i=1}^K Y_i$, then $(X_1, X_2, \dots, X_{K-1}) \sim \text{Dir}(a_1, a_2, \dots, a_K)$. Then it is obvious to write that $\lambda_2 = \frac{y_1}{y_1+y_2+y_3}$; $\lambda_3 = \frac{y_2}{y_1+y_2+y_3}$ where $y_1 \sim \mathcal{G}(n_1 + a_1, 1)$, $y_2 \sim \mathcal{G}(n_2 + a_2, 1)$ and $y_3 \sim \mathcal{G}(n_0 + a_0, 1)$ for suitable choice of hyper-parameters (a_0, a_1, a_2) . It is quite easy to generate random numbers from y_1, y_2 and y_3 and so for λ_2 and λ_3 .

After removing the burn-in from the generated chains, the stationarity and convergence can be checked using the cumsum plot, *ACF*, and Gelman and Rubin's test statistics Gelman and Rubin 1992. Along with the selection of priors, the choice of the loss function is also important to draw the required inference related to unknown parameters of interest. The Bayesian estimator of unknown parameters under the commonly used *SELF* is the mean of the posterior distribution. Suppose one wants to obtain the Bayesian estimator of some function of (β, Λ) , say, $g(\beta, \Lambda)$ under *SELF*, then the *SELF* is defined as

$$L_{SE}(\hat{g}(\beta, \Lambda), g(\beta, \Lambda)) = (\hat{g}(\beta, \Lambda) - g(\beta, \Lambda))^2,$$

where $\hat{g}(\beta, \Lambda)$ is an estimator of $g(\beta, \Lambda)$.

The Bayesian estimator of $g(\beta, \Lambda)$ under *SELF* is given by

$$\hat{g}(\beta, \Lambda)^{SE} = \frac{\int g(\beta, \Lambda)p(\beta, \Lambda | \underline{t}) d\beta d\Lambda}{\int p(\beta, \Lambda | \underline{t}) d\beta d\Lambda},$$

where $p(\beta, \Lambda | \underline{t})$ is the posterior density of (β, Λ) .

4.4.4 Testing under Bayesian Paradigm

In the Bayesian framework, testing of hypotheses can be carried out using posterior odds. Consider the dependent model established in Section 4.3. Assume that one wants to test the hypothesis $H_0 : \lambda_2 = \lambda_3$ against $H_1 : \lambda_2 \neq \lambda_3$. For doing this, the posterior odds are

given by

$$\text{posterior odds} = \frac{p(\beta, \Lambda \mid \underline{t}, H_1)}{p(\beta^*, \alpha^* \mid \underline{t}, H_0)}$$

which further can be written as

$$\text{posterior odds} = \frac{p(\underline{t} \mid H_1)}{p(\underline{t} \mid H_0)} \frac{\pi(H_1)}{\pi(H_0)} \quad (4.26)$$

The first term on the right-hand side of (4.26) is known as the Bayes Factor (*BF*). We use the *BF* to test the hypothesis stated earlier. Since under H_0 , the two competing risks are identical, hence suppose that the observed data (t_1, t_2, \dots, t_n) is coming from a Weibull distribution with shape and scale parameters β^* and α^* respectively. It should be noted that the shape parameter for the Weibull distribution is the same under null and alternative hypotheses, i.e. $\beta = \beta^*$. The likelihood function for the unknown parameters β^* and α^* given the data is

$$L(\beta^*, \alpha^* \mid \underline{t}) = \beta^{*n} \alpha^{*n} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \exp \left(-\alpha^* \sum_{j=1}^n t_j^{\beta^*} \right). \quad (4.27)$$

The log-likelihood function is

$$\log L(\beta^*, \alpha^* \mid \underline{t}) = n \log(\beta^*) + n \log(\alpha^*) + (\beta^* - 1) \sum_{j=1}^n t_j - \alpha^* \sum_{j=1}^n t_j^{\beta^*}. \quad (4.28)$$

In this case, a reference prior with partial prior information for the unknown parameters β^* and α^* is derived. From the log-likelihood equation (4.28), the negative of second partial derivative for α^* is

$$-\frac{\partial^2}{\partial \alpha^{*2}} \log L(\beta^*, \alpha^* \mid \underline{t}) \propto \frac{1}{\alpha^{*2}}.$$

According to Sun and Berger (1998), the conditional reference prior for α^* is

$$\pi(\alpha^* \mid \beta^*) \propto \frac{1}{\alpha^*}.$$

A non-informative prior for β^* can be taken as

$$\pi(\beta^*) \propto \frac{1}{\beta^{*\nu}} ; -\infty < \nu < \infty.$$

Hence the joint prior for the unknown parameters β^* and α^* is given by

$$\pi(\beta^*, \alpha^*) \propto \frac{1}{\alpha^* \beta^{*\nu}}. \quad (4.29)$$

In respect to prior given in (4.29), the posterior density will be as follows:

$$\begin{aligned} p(\beta^*, \alpha^* | \underline{t}) &\propto L(\beta^*, \alpha^* | \underline{t}).\pi(\beta^*, \alpha^*) \\ &\propto \beta^{*n-\nu} \alpha^{*n-1} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \exp \left(-\alpha^* \sum_{j=1}^n t_j^{\beta^*} \right). \end{aligned} \quad (4.30)$$

The marginal distribution of t under H_0 is given by

$$\begin{aligned} p(\underline{t} | H_0) &= \int_0^\infty \int_0^\infty p(\beta^*, \alpha^* | \underline{t}) d\alpha^* d\beta^* \\ &= \int_0^\infty \int_0^\infty \beta^{*n-\nu} \alpha^{*n-1} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \exp \left(-\alpha^* \sum_{j=1}^n t_j^{\beta^*} \right) d\alpha^* d\beta^* \\ &= \Gamma(n) I_1, \end{aligned} \quad (4.31)$$

where

$$I_1 = \int_0^\infty \beta^{*n-\nu} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \left(\sum_{j=1}^n t_j^{\beta^*} \right)^{-n} d\beta^*.$$

Now, we calculate the BF using posterior densities $p_1(\beta, \Lambda | \underline{t})$ and $p_2(\beta, \Lambda | \underline{t})$ given by (4.13) and (4.19). So, using $p_1(\beta, \Lambda | \underline{t})$ the marginal distribution of t under H_1 is given by

$$\begin{aligned} p_1(\underline{t} | H_1) &= \int p_1(\beta, \Lambda | \underline{t}) d\beta d\Lambda \\ &= \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1 - \frac{1}{2}} \lambda_3^{n_2 - \frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0 - \frac{1}{2}} d\lambda_2 d\lambda_3 \right\} \right. \\ &\quad \left. \lambda_1^{n+a-1} \exp \left\{ -\lambda_1(b + \sum_{j=1}^n t_j^\beta) \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \end{aligned}$$

$$= \mathcal{B}(n_1 + \frac{1}{2}, n_0 + n_2 + 1) \cdot \mathcal{B}(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}) \Gamma(n + a) I_2, \quad (4.32)$$

where

$$I_2 = \int_0^\infty \beta^{n-1} \left(\prod_{j=1}^n t_j^{\beta-1} \right) \left(b + \sum_{j=1}^n t_j^\beta \right)^{-(n+a)} d\beta.$$

Hence, using (4.31) and (4.32), the Bayes Factor is given by

$$BF_1 = \frac{p_1(\underline{t} | H_1)}{p(\underline{t} | H_0)} = \mathcal{B}(n_1 + \frac{1}{2}, n_0 + n_2 + 1) \cdot \mathcal{B}(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}) \frac{\Gamma(n + a)}{\Gamma(n)} \frac{I_2}{I_1}. \quad (4.33)$$

Similarly, using $p_2(\beta, \Lambda | \underline{t})$, the marginal distribution of \underline{t} under H_1 is given by

$$\begin{aligned} p_2(\underline{t} | H_1) &= \int p_2(\beta, \Lambda | \underline{t}) d\beta d\Lambda \\ &= \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1 + a_1 - 1} \lambda_3^{n_2 + a_2 - 1} (1 - \lambda_2 - \lambda_3)^{n_0 + a_0 - 1} d\lambda_2 d\lambda_3 \right\} \right. \\ &\quad \left. \lambda_1^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ &= \mathcal{B}(n_1 + a_1, n_0 + n_2 + a_0 + a_2) \cdot \mathcal{B}(n_2 + a_2, n_0 + a_0) \Gamma(n) I_3, \end{aligned} \quad (4.34)$$

where

$$I_3 = \int_0^\infty \beta^{n-1} \left(\prod_{j=1}^n t_j^{\beta-1} \right) \left(\sum_{j=1}^n t_j^\beta \right)^{-n} d\beta.$$

Similarly, using (4.31) and (4.34), the BF is given by

$$BF_2 = \frac{p_2(\underline{t} | H_1)}{p(\underline{t} | H_0)} = \mathcal{B}(n_1 + a_1, n_0 + n_2 + a_0 + a_2) \cdot \mathcal{B}(n_2 + a_2, n_0 + a_0) \frac{I_3}{I_1}. \quad (4.35)$$

The decision rule for Bayesian testing is that one would reject H_0 in favour of H_1 if the value of the BF is very high. Hence in the context of Bayesian testing, if the null hypothesis would be rejected then the two dependent competing risks would not be considered identical.

4.4.5 Posterior Predictive Compatibility of Model

It is indispensable to perform the compatibility of the model (i.e., likelihood and prior) established in the previous section before applying it to real data. Model compatibility is a necessary step to show that the established model is rightly used for the data. The initial and simplest approach considers the predictive capability of the model with regard to some of its important characteristics such as hazard rate, survival rate and mean time to failure etc., using graphical tools. In this section, the model compatibility for the observed data in the Bayesian framework is performed through *PPV* given by Guttman (1967). To obtain the *PPV*, the chi-square statistic, a measure of discrepancy (Upadhyay et al. 2001), is given below

$$\chi^2 = \sum_{i=1}^n \frac{(t_i - E(T_i^* | \beta, \Lambda))^2}{V(T_i^* | \beta, \Lambda)}. \quad (4.36)$$

Hence, the *PPV* based on the discrepancy measure defined in (4.36), can be obtained as

$$PPV = \int P[\chi_2^2 > \chi_1^2 | \beta, \Lambda] p(\beta, \Lambda | \underline{t}) d\beta d\Lambda,$$

where χ_1^2 and χ_2^2 are the calculated values of χ^2 for the observed data and predictive data sets, respectively. The *PPV* is defined as the expected value of the classical p-value, where the expectation is taken with respect to posterior density. Since we have considered the case of a complete sample, so to calculate *PPV*, the procedure described by Gelman et al. (1996) is followed. First of all, we draw posterior samples on (β, Λ) from $p(\beta, \Lambda | \underline{t})$ and then obtain the value of χ_1^2 for the given data set. After that, the predictive data sets are generated each of the same size as that of the observed data. The value of $P[\chi_2^2 > \chi_1^2 | \beta, \Lambda]$ is calculated as the number of times χ_2^2 exceeds χ_1^2 . The procedure defined above is replicated a number of times with extracted values of parameters and thus *PPV* is obtained as the posterior expectation of $P[\chi_2^2 > \chi_1^2 | \beta, \Lambda]$.

4.5 Simulation Study

In this section, we have carried out a simulation study to illustrate the performance of derived estimators in Section 4.4. The simulation study has been performed to obtain the Bayesian estimates of the unknown parameters under reference priors with partial prior information (*RPPI*), namely, *RPPI-I* and *RPPI-II* derived in Subsection 4.4.1. The performance of estimators is assessed by absolute relative bias (*ARB*), *MSE* and *CP*. We have taken sample sizes $n = 20, 40, 60$ and 100 to evaluate and compare the efficiency of estimators as the sample size increases. The simulation study is based on 5000 samples to remove undesired sampling fluctuations and to achieve consistency in the outputs. The output of the simulation performed is shown in Table 4.1 to Table 4.4 for two different values of shape parameter β as from $MOBWD(1.55, 0.65, 0.35, 0.45)$ and $MOBWD(0.75, 0.65, 0.35, 0.45)$, respectively. For Bayesian inference, hyper-parameter values are chosen arbitrarily. For posterior sample generation and computational purposes, **R** software is used. After eliciting the priors, the Bayes estimates of the parameters are obtained using the M-H algorithm. The Markov chain of length 10^5 is generated and the first 2×10^4 observations are removed as a burn-in period. After that remaining observations were chosen at a gap of 15 to remove any serial correlation among them.

The *AL* of *HPD* and *CP* of interval estimates with varying sample sizes are reported in Table 4.2 and Table 4.4. The Bayes estimates of the unknown parameters are obtained under *SELF* for $\gamma = (-1.5, 1.5)$ to capture the behaviour of estimators in case of underestimation and overestimation. The important findings from the simulated results can be summarized. From simulated results, we can see that all the Bayes estimators perform efficiently with the increase in sample size. *ARB* and *MSE* of the estimators decrease as sample size increases. This is true under each reference prior with partial information. The *CP* for parameters (β, Λ) tends near to the nominal level of 0.95 as the sample size increases. The *HPD* length for the parameter λ_1 under *RPPI-I* is shorter than the *HPD* length for the parameter λ_1 under *RPPI-II* and the pattern remains the same with an increase in sample size. The *HPD* length for the parameters λ_2 and λ_3 under *RPPI-II* is shorter than the *HPD*

length for the parameter λ_2 and λ_3 under *RPPI-I* and the pattern remains the same with an increase in sample size.

Table 4.1: ARB and MSE of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$ under *SELF*.

n	RPPI-I		RPPI-II		
	ARB	MSE	ARB	MSE	
20	β	0.1219	0.0577	0.1233	0.0592
	λ_1	0.1816	0.1244	0.2019	0.1626
	λ_2	0.2954	0.0079	0.2664	0.0065
	λ_3	0.2560	0.0096	0.2272	0.0077
40	β	0.0950	0.0360	0.0953	0.0363
	λ_1	0.1276	0.0585	0.1337	0.0645
	λ_2	0.2138	0.0041	0.2012	0.0037
	λ_3	0.1805	0.0049	0.1690	0.0043
60	β	0.0784	0.0240	0.0785	0.0241
	λ_1	0.1034	0.0368	0.1061	0.0390
	λ_2	0.1786	0.0030	0.1709	0.0027
	λ_3	0.1510	0.0034	0.1445	0.0031
100	β	0.0611	0.0144	0.0611	0.0143
	λ_1	0.0798	0.0219	0.0812	0.0227
	λ_2	0.1388	0.0018	0.1358	0.0017
	λ_3	0.1181	0.0021	0.1155	0.0020

Table 4.2: *AL* of 95% *HPD* interval and coverage probability of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$.

n	RPPI-I		RPPI-II		
	AL	CP	AL	CP	
20	β	0.9978	0.9726	1.0025	0.9738
	λ_1	1.2708	0.9584	1.3431	0.9458
	λ_2	0.3308	0.9404	0.3226	0.9452
	λ_3	0.3604	0.9338	0.3490	0.9456
40	β	0.7294	0.9540	0.7310	0.9522
	λ_1	0.8980	0.9498	0.9201	0.9402
	λ_2	0.2481	0.9274	0.2441	0.9574
	λ_3	0.2697	0.9376	0.2645	0.9512
60	β	0.5998	0.9528	0.6005	0.9532
	λ_1	0.7334	0.9538	0.7450	0.9508
	λ_2	0.2070	0.9474	0.2046	0.9538
	λ_3	0.2239	0.9364	0.2209	0.9474
100	β	0.4688	0.9548	0.4690	0.9552
	λ_1	0.5678	0.9500	0.5731	0.9448
	λ_2	0.1627	0.9442	0.1617	0.9458
	λ_3	0.1762	0.9446	0.1749	0.9490

Table 4.3: ARB and MSE of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$ under *SELF*.

n		RPPI-I		RPPI-II	
		ARB	MSE	ARB	MSE
20	β	0.1480	0.0210	0.1500	0.0216
	λ_1	0.1796	0.1182	0.1995	0.1521
	λ_2	0.2948	0.0078	0.2666	0.0064
	λ_3	0.2455	0.0090	0.2181	0.0072
40	β	0.1011	0.0095	0.1014	0.0096
	λ_1	0.1308	0.0610	0.1370	0.0674
	λ_2	0.2198	0.0044	0.2070	0.0039
	λ_3	0.1795	0.0048	0.1678	0.0043
60	β	0.0823	0.0062	0.0825	0.0062
	λ_1	0.1048	0.0379	0.1080	0.0403
	λ_2	0.1794	0.0029	0.1718	0.0027
	λ_3	0.1487	0.0033	0.1424	0.0031
100	β	0.0644	0.0037	0.0644	0.0037
	λ_1	0.0803	0.0220	0.0817	0.0227
	λ_2	0.1393	0.0018	0.1367	0.0017
	λ_3	0.1163	0.0021	0.1137	0.0020

Table 4.4: AL of 95% HPD interval and coverage probability of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$.

n		RPPI-I		RPPI-II	
		AL	CP	AL	CP
20	β	0.5317	0.9540	0.5350	0.9542
	λ_1	1.2688	0.9656	1.3397	0.9498
	λ_2	0.3321	0.9460	0.3236	0.9534
	λ_3	0.3616	0.9428	0.3497	0.9532
40	β	0.3685	0.9548	0.3691	0.9522
	λ_1	0.9014	0.9522	0.9239	0.9452
	λ_2	0.2484	0.9250	0.2445	0.9558
	λ_3	0.2697	0.9410	0.2646	0.9532
60	β	0.2991	0.9482	0.2997	0.9490
	λ_1	0.7330	0.9494	0.7443	0.9456
	λ_2	0.2068	0.9428	0.2045	0.9500
	λ_3	0.2238	0.9390	0.2210	0.9504
100	β	0.2301	0.9470	0.2304	0.9454
	λ_1	0.5674	0.9496	0.5725	0.9476
	λ_2	0.1627	0.9464	0.1618	0.9468
	λ_3	0.1762	0.9440	0.1749	0.9464

4.6 Real Data Application

In this section, to elucidate the performance of estimators established in the previous section, two real data sets are utilized. The first data set is from a diabetic retinopathy study, while the second data set is from prostate cancer. These two data sets are analysed using MOBWD under competing risks set up.

4.6.1 Prostate Cancer Data

Here, we have used the prostate cancer data set, which was first reported in Byar and Green (1980). The data set originally consists of 502 patients. We consider only 182 patients for our purpose corresponding to the major causes of death related to patients. The major causes of death are distributed for these 182 patients as 127 patients died of prostatic cancer; 31 patients died of cerebrovascular accident, and 24 patients died of causes other than these two. The main purpose of analysing this dataset is that we can compare the risk of death under all causes except cerebrovascular accident with all causes except prostatic cancer. Before analysing the dataset, we first form the groups of patients as per their cause of death to make the data suitable for our dependent competing risks model. The first group includes those patients who died of prostatic cancer or other causes, and the second group includes those who died of a cerebrovascular accident or other causes. The summary statistics of the patients according to their cause of death are given in Table 3.11.

For j^{th} patient, let X_{1j} and X_{2j} be the latent time of death (in months) due to group 1 and group 2 of causes, respectively. So that $T_j = \min(X_{1j}, X_{2j})$ be the observed death time of the j^{th} patient. Hence, if the patient dies of prostatic cancer, then only X_{1j} is observed; if the patient dies of cerebrovascular accident, then only X_{2j} is observed; and if the patient dies of other causes, then X_{1j} and X_{2j} both are observed simultaneously. We first go for the model predictive capability using some graphical tool based on survival rate. For this, we first generate 50 posterior samples of the same size as the size of prostate cancer data corresponding to the dependent competing risks model using the distribution of time to death for patients. For each generated predictive sample, the survival rate is calculated. Figure 4.1

shows that the survival rate of patients is greatly superimposed by the survival rates of the 50 predictive samples. Before performing an estimation procedure for unknown parameters for the prostate cancer data, we check the model compatibility in the form of PPV discussed in Subsection 4.4.5. To calculate the PPV , we generate 200 posterior samples of the same size as the size of the prostate cancer data corresponding to the dependent competing risks model. Further, we obtain the value of χ_1^2 for the data for each generated posterior sample on parameters (β, Λ) . In the next step, we generated 1000 predictive samples from the model, each of the same sizes as of the data, for each value of the set of parameters (β, Λ) . Based on these 1000 predictive samples, we calculated χ_2^2 . On comparing the value of χ_2^2 with χ_1^2 , we obtain $P[\chi_2^2 > \chi_1^2]$. Finally, the above procedure is repeated a sufficient number of times to calculate PPV . We found the estimated value of PPV equals 0.619. Thus, with a greater value of PPV , we can conclude that the considered model (i.e. likelihood and prior) in the Bayesian paradigm is compatible with the considered dataset. Now, we are in a position to

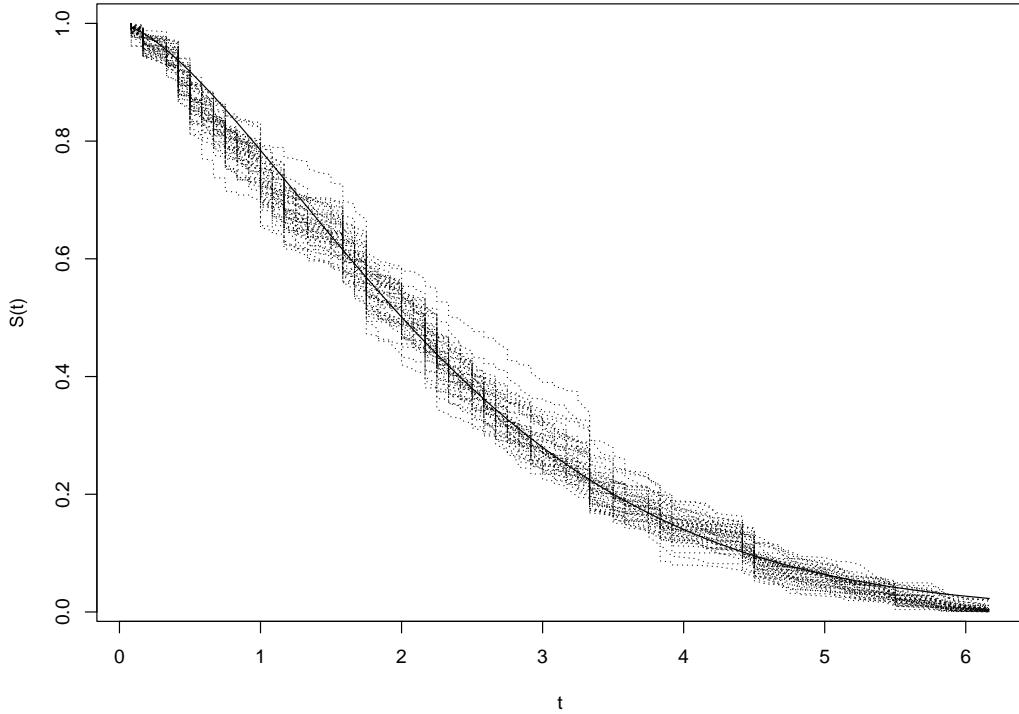


Figure 4.1: Estimated survival plots for observed prostate cancer data (solid line) and predictive datasets.

perform the Bayesian analysis of prostate cancer data. The Bayesian estimation procedure is carried out under reference priors *RPPI-I* and *RPPI-II* derived in Subsection 4.4.1. Since, for the real data, we do not have any prior information, the value of hyper-parameters is taken to be 0.001. The *MCMC* algorithm is applied to generate 10^5 posterior samples on the parameters and the first 2×10^4 samples are discarded to eliminate the effect of the initial value used for iteration. After that for removing the serial correlation of sampled draws, one value is chosen in every 10 iterations. Thus, the Bayesian estimates are based on 8,000 posterior samples. The values of hyperparameters are chosen arbitrary. Using the full conditionals (4.20) to (4.25), sample values are generated for parameters (β, Λ) following the *MCMC* procedure defined in Subsection 4.4.3. The Bayesian estimates of the parameters (β, Λ) along with their 95% *HPD* interval and *HPD* length under RPPI-I and RPPI-II are presented in Table 4.5, respectively.

Table 4.5: Bayes estimates of parameters of *MOBWD* for prostate cancer data under *SELF*.

RPPI-I				RPPI-II				
	SELF	95% HPD			SELF	95% HPD		
		Lower	Upper	Length		Lower	Upper	Length
β	1.5062	1.3810	1.6201	0.2391	1.5068	1.3814	1.6235	0.2421
λ_1	0.2531	0.2025	0.3047	0.1021	0.2527	0.2025	0.3037	0.1012
λ_2	0.6980	0.6281	0.7632	0.1352	0.6971	0.6303	0.7618	0.1315
λ_3	0.1701	0.1172	0.2259	0.1086	0.1707	0.1201	0.2274	0.1073

4.6.2 Diabetic Retinopathy Study Data

A diabetic retinopathy study (*DRS*) was conducted by National Eye Institute to evaluate the effect of laser treatment on delaying the onset of blindness in patients diagnosed with diabetic retinopathy. Laser treatment was administered to one eye at random, while the other was left untreated. The time to blindness for the patients was recorded for each eye. It was observed that, for ten individuals (approx 14%), the time to blindness was determined to be the same in both eyes. Considering this data set as competing risks data, the event

of interest is blindness of the eye system and the two risks of blindness are treatment and absence of treatment. As a result, the two risks are not independent and are being analysed using *MOBWD*. Let us denote X_1 as the time to the blindness of the eye under laser treatment and X_2 being the time to the blindness of the eye in the absence of treatment. To perform the model predictive capability using some graphical tools, we first generate 50 posterior samples of the same size as the size of *DRS* study data corresponding to the dependent competing risks model using the distribution of time to the blindness of eyes of diabetic patients. For each generated predictive sample, the survival rate is calculated. Figure 4.2 shows that the survival rate of the blindness of eyes for diabetic patients is greatly superimposed by the survival rates of the 50 predictive samples. Now we perform the model compatibility in the form of *PPV* discussed in Subsection 4.4.5. To calculate *PPV*, we adopted the same procedure as in the case of prostate cancer data. We found the estimated *PPV* equals 0.383. Thus, with a greater *PPV*, we can conclude that the considered model (i.e. likelihood and prior) is compatible with the considered dataset. The Bayesian analysis of the considered dataset is carried out under reference priors *RPPI-I* and *RPPI-II* derived in Subsection 4.4.1. The values of hyperparameters are chosen arbitrary. Since we do not have any prior information for the real data, we have taken all the hyper-parameters to be 0.001. Using the full conditionals (4.20) to (4.25), sample values are generated for parameters (β, Λ) following the *MCMC* procedure defined in Subsection 4.4.3. The *MCMC* algorithm is applied to generate 10^5 posterior samples on the parameters and the first 2×10^4 samples are discarded to eliminate the effect of the initial value used for iteration. After that for removing the serial correlation of sampled draws, one value is chosen in every 10 iterations. Thus, the Bayesian estimates are based on 8,000 posterior samples. The Bayesian estimates of the parameters (β, Λ) along with their 95% *HPD* interval and *HPD* length under *RPPI-I* and *RPPI-II* are presented in Table 4.6, respectively. We further perform the Bayesian testing of hypotheses to test whether the two competing risks are identical or not. To do so, we calculate the *BF* in (4.33) and (4.35) obtained under *RPPI-I* and *RPPI-II*, respectively. The calculated *BF* comes out to be very low, so the null hypothesis of the two risks being identical can not be rejected in the Bayesian context; hence, due to the lack of proper

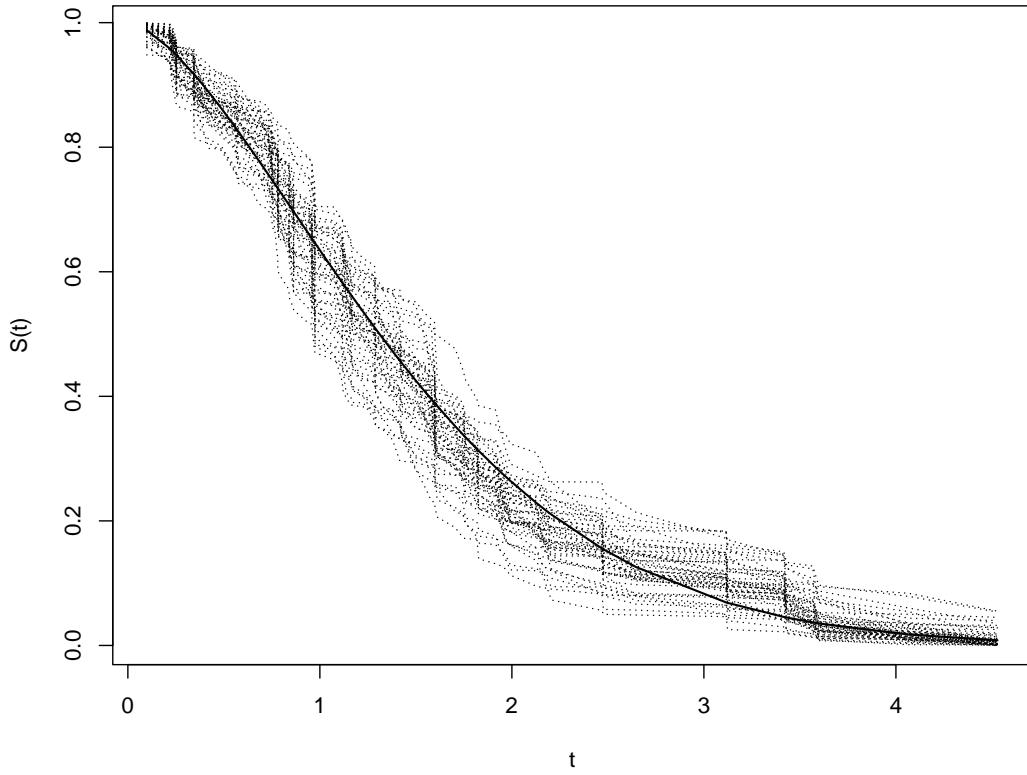


Figure 4.2: Estimated survival plots for observed DRS data (solid line) and predictive datasets.

evidence, we can conclude that the laser treatment is not effective in delaying the blindness of eyes in diabetic patients.

Table 4.6: Bayes estimates of parameters of *MOBWD* for DRS data under *SELF*.

	RPPI-I				RPPI-II			
	SELF		95% HPD		SELF		95% HPD	
			Lower	Upper	Length		Lower	Upper
β	1.5548	1.3479	1.7557	0.4078	1.5559	1.3431	1.7546	0.4115
λ_1	0.4713	0.3414	0.5991	0.2578	0.4710	0.3371	0.6031	0.2660
λ_2	0.3949	0.2857	0.5057	0.2200	0.3943	0.2854	0.5077	0.2224
λ_3	0.4646	0.3495	0.5777	0.2283	0.4642	0.3531	0.5828	0.2297

4.7 Conclusion

In this chapter, we have provided the objective Bayesian method for analysing the dependent competing risks data arising from *MOBWD*. The objective Bayesian approach is carried out in cases when one has prior information for part of the parameters and for the rest of the parameters, the prior information is vague or not available. So, we have derived the reference priors for the unknown parameters of *MOBWD* when partial prior information is available to some of the parameters only. The reference prior is derived for two cases when the prior information is available with parameters (β, λ_1) only and when the same is available with (λ_2, λ_3) only. We also established that the derived reference priors, yet improper, lead to a suitable posterior distribution for the parameters. The Bayesian estimates of the unknown parameters are obtained under the two reference priors *RPPI-I* and *RPPI-II* using *SELF*. As a result, if some of the parameters have no prior information, the reference prior with partial information obtained is an alternative approach to carrying out the Bayesian estimation. The compatibility of the proposed model under the Bayesian paradigm is performed using the *PPV*. It is shown that the established model is compatible with the considered datasets on the basis of *PPV*. In the end, the practical applicability of the established model under the Bayesian paradigm is illustrated with the help of the two datasets used in the previous chapter.

Although this chapter focused on the estimation of unknown parameters of the *MOBWD*, the proposed method can be applied to other bivariate distributions with partial prior information. The proposed method can also be extended for competing risks data under different censoring schemes, such as type-I censoring, type-II censoring, progressively hybrid censoring, etc, and also for incomplete data problems.

Chapter 5

Bayesian Joint Modelling of Longitudinal and Competing Risks Data with Cause-Dependent Masking

5.1 Introduction

In many clinical trials or follow-up studies, a response variable is measured repeatedly over different time points during the follow-up period. The repeated measurements of the response variable are known as longitudinal data in the literature. In such studies, along with the repeated measurements, time-to-event of interest is also recorded for each individual in the study generating longitudinal and survival data jointly. The time-to-event of interest may induce non-ignorable missing data for the longitudinal response. An interesting example can be found in AIDS clinical trial in which the CD4 lymphocyte cell counts (longitudinal outcome), a potential biomarker for diagnosing the AIDS disease, are being measured at different time points for each patient in the study and the time to diagnosis (survival outcome) of AIDS which may censor the repeated observations for the patients, is also being recorded. It is always possible that the longitudinal biomarker and the survival outcome are correlated as in the case of the AIDS clinical study example mentioned above and hence, the joint modelling of both outcomes/processes is indispensable. Many times, an individual is subject to more than one risk simultaneously and hence, along with the time-to-event, the cause of the event is also observed. In literature, such type of data is known as competing risks data. In clinical trial studies, the longitudinal data are often

ignored while modelling the competing risks data and thus, does not take into account of the relationship between the two outcomes. If the two processes are correlated in some way then the separate modelling of the two outcomes may substantially produce biased estimates and consequently, the efficiency of the statistical inference is relatively not high. In such a case, joint modelling of both longitudinal and competing risks survival data is necessary. The joint modelling of both outcomes produces relatively unbiased estimates and increases the efficiency of the statistical inference.

An essential feature of joint modelling is to define an appropriate association structure between the two outcomes as it depicts the effect of the longitudinal outcome on the time-to-event outcome and vice-versa. In the joint modelling framework, different kinds of association structures have been proposed so far. The problem at hand is choosing the most appropriate one for the joint model under consideration. The choice can be made based on the clinical background of the study of interest which may not be accessible. Elashoff et al. (2008) and Huang et al. (2011a) analysed the sub-models for longitudinal and competing risks data with correlated random effects association structure whereas Andrinopoulou et al. (2017) considered that the random effect corresponding to competing risks sub-model is the cumulative total of the current value of longitudinal response and thus both sub-models are known to be linked. One can go through Alsefri et al. (2020) for a decent review of literature on association structure in a joint modelling framework. Whenever the two outcomes would be uncorrelated, the joint modelling will simply reduce to two independent models. The longitudinal outcome and competing risks survival outcome can be linked with latent structure or shared random effects. Also, the dependency between the longitudinal and survival outcome is accounted for by common covariates of the two outcomes. For analysing the joint model, conditional independence is assumed between the two outcomes given the unobserved random effects and common covariates.

In a competing risks scenario, a study becomes complex if the true cause of failure is not known. This situation is observed due to lack of some proper diagnostic tools, time and cost constraints on the procedure etc. The unknown cause of failure data is referred to as masked data in the literature. In biomedical studies, it is common that the cause of death is missing

for the individuals who were at risk of death due to more than one cause. For instance, in the case of examining the cause of death for a patient who is at risk of dying due to cancer, it is possible that his/her cause of death may be masked with causes other than cancer. Another interesting example can be found in Kodell and Chen (1987) in animal bioassays, where the death of an animal due to a tumour may be masked with the other causes with the tumour present. Many authors have analysed the masked data with the symmetry assumption that the probability of masking does not depend on the true cause of failure and assumed the equality of masking probabilities for all the considered causes under consideration. Kuo and Yang (2000) and Rai et al. (2021) relaxed the symmetry assumptions and assumed that the masking probabilities are different for each cause of failure for the masked data. In the cause-specific sub-model for competing risks data, we are considering the cause-dependent masking scenario for the joint modelling of the two processes.

The estimation of the longitudinal and competing risks model parameters jointly can be carried out with frequentist as well as Bayesian approach. The *ML* method of estimation is very popular in the classical approach whereas the Bayesian estimation is frequently carried out with *MCMC* method. Our focus will be on the estimation of parameters under the Bayesian paradigm as it is computationally more tractable for high dimensional random effects and against the model complexity. The Bayesian method is quite helpful in the estimation of the parameters when the dimensionality of the data is high with respect to random effects and covariates. Moreover, it incorporates the prior information available with respect to the unknown parameters which further helps in improving the estimation accuracy. Hu et al. (2009) analysed the joint model linked with correlated random effects under the Bayesian paradigm. Similarly, Rue et al. (2017) developed a joint model under the Bayesian paradigm and showed its applicability to study patient-ventilator asynchronies in critical care patients. Recently, Sheikh et al. (2021) proposed a Bayesian joint model for prostate cancer data with missing grades and analysed the data with independent masking assumption.

In this chapter, we develop a flexible joint model for longitudinal data and competing risks data incorporating cause-dependent masking under Bayesian paradigm. Particularly,

the longitudinal observations are analysed through *LMM* while considering the random effects and errors as normally distributed. Further, the cause-specific proportional hazard model is deployed to study the competing risks data. The two outcomes are linked through shared random effect method. Additionally, the model is able to deal with the cause dependent masking scenario in competing risks data where we assume that the probability of masking is not independent of the causes.

The rest of the chapter is laid out as follows. In Section 5.2, we discuss the PLCO data which motivated the work being done. The joint model along with the sub-models for longitudinal data and competing risks data with masked causes of failure are structured in Section 5.3. Under the Bayesian paradigm, the priors and corresponding posteriors were derived for the parameters of the joint model in Section 5.4. In Section 5.5 the simulation study has been performed to establish the empirical behaviour of the joint model proposed. In the end, the chapter is concluded with the results and their discussion in Section 5.6.

5.2 Motivation of the Work

The Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial is a population-based randomised and controlled trial sponsored by the national cancer institute. It was conducted to assess the screening exams' effectiveness in reducing the mortality related to PLCO cancers along with their diagnosis. In the trial, a total 76678 number of study participants were recruited between the years 1993 and 2001 for the diagnosis of prostate cancer among men. The screening was ended in year 2006. In this study, race, family history of cancer and smoking status were collected at baseline and weight (lbs), height (inches) and body mass index (BMI) were collected longitudinally over the study period. The prostate-specific antigens (PSAs), a clinically accepted biomarker for prostate cancer, were collected from the study participants periodically throughout the follow-up period. The time to diagnosis of prostate cancer due to low grade or high grade is the variable of interest in competing risks process. Whereas, the PSA is the variable of interest in longitudinal data. The grade of the prostate cancer is defined on the basis of Gleason score. A participant is

said to be diagnosed of low grade cancer if the Gleason score is less than 7 and a high grade cancer if the Gleason score is 7 to 10. After considering the participants who had at least two follow-up visits, which led to a total of 32650 participants available. A total of 3562 males were diagnosed with prostate cancer during the study period, among whom 2026 males were diagnosed with low grade, 1495 were diagnosed with high grade and 41 males were diagnosed with missing grade prostate cancer. The grade of the cancer is generally an indication of how quickly the cancer cells are likely to grow and spread to other parts of the body. Hence, the grades are useful in helping the doctors to plan the treatment and determine the prognosis according to the severity of the condition of the patient. The prostate cancer data from PLCO motivates us to establish a flexible joint modelling framework to analyse the longitudinal data and competing risks data simultaneously while incorporating cause-dependent masking.

5.3 Joint Modelling

The inference based on longitudinal data and time-to-event data in their respective field of study may produce biased results as both types of observations are recorded on the same individuals showing inherent dependency between them. Thus, a joint inference of both outcomes can address statistical issues that can not be handled in separate modelling of both outcomes. Hence, joint modelling is needed to deal with this environment and also one can incorporate additional information such as informative dropout of the individuals etc. from the study. Joint modelling can reduce the bias in parameter estimation effectively and increases the efficiency of statistical inference over separate modelling of both outcomes (see Chen et al. 2011; Ibrahim et al. 2001). The joint modelling of longitudinal data and time-to-event data has three components mainly- a sub-model for longitudinal data, a sub-model for competing risks data and an association structure which specifies the correlation between them precisely. To correlate the longitudinal outcome and time-to-event data, one can use a latent variable approach, and multivariate Gaussian process (see Elashoff et al. 2008) which accommodate the linear association between them and a multivariate copula

which defines a non-linear relationship between the two outcomes (see Zhang et al. 2022).

5.3.1 Sub-model for Longitudinal Data

A longitudinal data set refers to repeated observations being taken on a response variable for an individual over time in a follow-up study. Suppose n individuals are being followed up in a study. For i^{th} individual, let m_i longitudinal measurements are taken at observation points $\tau \in \{\tau_{i1}, \tau_{i2}, \dots, \tau_{im_i}\}$; where $0 \leq \tau_{i1} < \tau_{i2} < \dots < \tau_{im_i}$ and $m_i \geq 0$, $\forall i \in I^+ = \{1, \dots, n\}$. It may be possible for each individual that the repeated measurements on longitudinal response are taken at different time points. This is the case of an imbalanced longitudinal data study (Colosimo et al. 2012). In case observation points are common for all individuals under study and consequently $m_i = m; \forall i \in I^+$, it reduces to balanced longitudinal data study. In this study, a *LMM* (Laird and Ware 1982) is considered to analyse the longitudinal data. Let $Y_i(\tau)$ denotes the longitudinal measurement on response variable for i^{th} individual over time τ . Since the observations are taken over time and it is associated with the time-to-event of interest, hence it is inevitable to track the evolution of longitudinal outcome over time for an individual and its variation between the individuals. So, to incorporate the within and between individuals' variation over time, we consider the *LMM* given by

$$Y_i(\tau) = Y_i^*(\tau) + \epsilon_i(\tau), \quad (5.1)$$

where the observed longitudinal response variable $Y_i(\tau)$ is defined as sum of $Y_i^*(\tau)$, true value of the outcome and measurement error $\epsilon_i(\tau)$. The model (5.1) is a *LMM* which comprises random effects as well as fixed effects, so it can be further defined as

$$Y_i(\tau) = \mathcal{A}_i(\tau)' \vartheta_i + \mathbf{X}'_i \gamma + \epsilon_i(\tau), \quad (5.2)$$

where $\mathcal{A}_i(\tau)$ is a $(p+1)$ -dimensional vector of function of τ , ϑ_i is $(p+1)$ -dimensional vector of random effects, $\mathbf{X}_i = (x_{i1}, \dots, x_{ir_1})'$ is r_1 -dimensional vector of explanatory variables which may or may not be consist of time-varying covariates, $\gamma = (\gamma_1, \dots, \gamma_{r_1})'$ is the r_1 -dimensional vector of fixed-effects corresponding to \mathbf{X}_i .

In the model (5.2), $\epsilon_i(\tau)$ is a measurement error which is accountable for the variation in Y_i which cannot be explained by the model itself. It is assumed to be normally distributed with mean zero and variance σ^2 , i.e., $\epsilon_i \sim \mathcal{N}(0, \sigma^2 I_{m_i})$ which is salutary for model implementation. Although the distribution of ϵ_i is subject to the behaviour of the longitudinal response variable Y_i . Notably, if the longitudinal response Y_i is skewed or multimodal, then assuming the normal distribution for it, is not preferred. In that case, one has to opt for a skewed distribution such as skewed normal or skewed t distribution (see Huang et al. 2011b; Lu 2017). Now, we define random effects as $\vartheta_i = \vartheta + \vartheta_i^*$, where ϑ and ϑ_i^* are $(p+1)$ -dimensional vector of overall effect and $(p+1)$ -dimensional vector of individual-specific random effects, respectively. We further assume that random effects and measurement errors are independent and $\vartheta_i^* \sim \mathcal{N}(0, \Sigma)$. The variance-covariance matrix Σ is $(p+1) \times (p+1)$ symmetric and positive-definite matrix. Further, to overcome the issue of slow convergence in *MCMC* sampling, we use Cholesky decomposition of Σ as suggested by Pinheiro and Bates (1996), and write $\Sigma = \Gamma\Gamma'$, so that re-parametrized quantity is given by $\vartheta_i^* = \Gamma\vartheta_i^R$; where $\vartheta_i^R \sim \mathcal{N}(0, I_{p+1})$ and Γ is a lower triangular matrix of order $(p+1) \times (p+1)$ and can be written as

$$\Gamma = \begin{bmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & 0 & \cdots & 0 \\ b_{31} & b_{32} & b_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{p+1,1} & b_{p+1,2} & b_{p+1,3} & \cdots & b_{p+1,p+1} \end{bmatrix};$$

where diagonal elements b_{ll} ; $l = 1, \dots, p+1$ are positive and there is no such restriction on the non-diagonal elements.

The model (5.2) can be rewritten as

$$Y_i(\tau) = \mathcal{A}_i(\tau)'(\vartheta + \Gamma\vartheta_i^R) + \mathbf{X}'_i\gamma + \epsilon_i(\tau). \quad (5.3)$$

It is to be noted that the evolution of trajectory for i^{th} individual is given by \mathcal{A}_i over

observation time τ . In mixed-effect model (5.3), the trajectory function will take linear form if $p = 1$ and quadratic form if $p = 2$ and so forth. Let the parameter vector of longitudinal sub-model is denoted by $\Psi_1 = (\Gamma, \gamma, \vartheta, \sigma^2)$. It is to be worth mentioning that fixed effect parameters are responsible for the inter-individual differences whereas the random effects are capable of capturing the intra-individual differences over time. Also, random effects are the factors varying across individuals.

Thus, the *LMM* is appropriate to analyse the longitudinal response as it captures the individual level variation through random effects as well as between individuals variation through the fixed effects (Elashoff et al. 2007, 2008; Henderson et al. 2000). To study individual characteristics, it is desired to trace the trajectory of individuals over time. Generally, for individuals, the trajectory of the longitudinal response is non-identical due to heterogeneity among them. So, it is needed to capture those variations over the individuals with time. Thus, random effects play a vital role in it as it is accountable for the unobserved heterogeneity within individuals.

5.3.2 Sub-model for Masked Competing Risks Data

In survival studies, time-to-event data refers to the time to death of an individual, diagnosis of a disease, hospital discharge and so forth. It may be a case where the event can be occurred due to one of the possible causes acting on the individual simultaneously. This is the scenario of competing risks data where individuals are at risk of more than one cause at the same time and only one of them appears to be the cause of the event. Now we define the data structure in competing risks scenario.

Assuming K different causes of failure, let $T_{i1}^*, \dots, T_{iK}^*$ be the true failure times for each cause and C_i be the non-informative censoring time for i^{th} individual along with an indicator function δ_i , which takes value 0 if the individual is censored otherwise 1. The observed event time either failure or censoring is given by $T_i = \min(T_{i1}^*, \dots, T_{iK}^*, C_i)$ for i^{th} individual. It may be possible that for some individuals, the true cause of failure is unknown due to the time-consuming process of identification, cost issues, the complexity of the failure mechanism or inefficient procedure to diagnose the cause etc. In literature, such incompleteness in the

competing risks data related to cause is known as masked data. Masked data can be seen in the industrial fields where the true cause of failure (component) for a system is not known and it may be ascertained to the subset of components.

Guess et al. (1991) defined a minimum random subset which contains the true cause of failure for systems under masked data problem. Also, Rai et al. (2021) analysed the masked data using Lindley distribution on the similar ground. Now, on the same pattern, we define a set S_i for i^{th} individual whose failure is observed due to one of the causes k ; $k=1, \dots, K$. Thus, in case of uncensored observation, a minimum random subset (MRS), i.e., $S_i \subseteq \{1, \dots, K\}$ will be observed.

For i^{th} individual,

$$S_i = \begin{cases} S_k & ; \text{if cause of failure is } k=1, \dots, K, \\ S_m & ; \text{if cause of failure is masked;} \end{cases} \quad (5.4)$$

where S_k is the set of individuals failed with observed cause k and $S_m = \{1, 2, \dots, K\}$ is the set of all causes and associated with completely masked individuals.

Now, let n_k be the number of individuals whose failure is due to k^{th} cause and n_m be the number of individuals whose cause of failure is masked and thus, $(n - n_m - \sum_{k=1}^K n_k)$ individuals are censored. Further, along with the time-to-event of interest, observation on some covariates Z_i is also recorded, where Z_i may or may not share common covariates with \mathbf{X}_i for i^{th} individual. Thus the observed competing risks survival data for i^{th} individual can be defined as $d_i = (t_i, s_i, z_i, \delta_i); i = 1, \dots, n$.

In this article, a cause-specific hazard model is considered for time-to-event data to analyse the effect of covariates on the occurrence of the event. Further, to associate the sub-model for competing risks data with the sub-model for longitudinal data, random effects ϑ^R are taken as covariates along with Z . For k^{th} observed cause of failure, the sub-model is defined as

$$h_k(t_i | z_i) = h_{0k}(t_i) \exp(\alpha'_k \vartheta^R + z'_i \beta_k); \quad k = 1, \dots, K. \quad (5.5)$$

where h_{0k} is non-negative baseline hazard function for cause k , β_k is r_2 -dimensional vector of regression coefficients with respect to r_2 -dimensional vector of covariates and α_k is the association parameter which is responsible for the joint inference of competing risks data and longitudinal data. In case, $\alpha_k = 0$, both the sub-models will be independent and hence they can be analysed separately. The baseline hazard function h_{0k} for cause k can be left unspecified as it will be a nuisance parameter while estimating the parameters of the cause-specific hazard model (5.5).

The baseline hazard function defines the hazard of the event in absence of covariates and hence plays an important role in comparing the effect of them on the time-to-event of interest. We specify a flexible parametric form for h_{0k} taking Weibull baseline hazard function defined as $h_{0k}(t_i) = \lambda_k t_i^{\lambda_k - 1}$ and hence, the cause-specific hazard model (5.5) can be written as

$$h_k(t_i | z_i) = \lambda_k t_i^{\lambda_k - 1} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k); \quad k = 1, 2, \dots, K. \quad (5.6)$$

In other words, the time to failure due to cause k follows a Weibull distribution with shape parameter λ_k and scale parameter $\exp\left(-\frac{\alpha'_k \vartheta_i^R + z'_i \beta_k}{\lambda_k}\right)$. It is clear that the scale parameter is a function of covariates as well as random effects whereas the shape parameter is free from those quantities. Although the parametric assumption for baseline hazard is quite straightforward, a more general form such as piecewise constant function, B-spline or some non-parametric form for baseline hazard can be assumed (see Zhang et al. 2021, 2022).

In the PLCO cancer prevention trial, there are some cases for which the grade of prostate cancer for diagnosed patients is missing. As the grade of a cancer is indication of the working mechanism (e.g., growth and spreading) of the cancer cells, it plays a vital role in determining the treatment protocol and prognosis for the patients. The assumption of cause-dependent masking is assumed here for analysing those missing grades which in turn will be helpful for the grade classification. In the case of cause-dependent masking with complete missingness, we define

$$q_k = P(S_i = \{k\} | z_i, K_i = k) \text{ and } 1 - q_k = P(S_i = \{1, \dots, K\} | z_i, K_i = k), \quad (5.7)$$

where the identification probability q_k is the probability of observing the k^{th} true cause inducing failure and the masking probability $1 - q_k$ is the probability of getting masked observation for i^{th} individual whose failure is actually induced by k^{th} cause. Here, we assume that q_k is independent of time but depends on the true cause of failure.

The likelihood contribution of i^{th} individual for competing risks data can be defined as

$$L_{ci}(\Psi_2 \mid \delta_i, \vartheta_i^R, \mathbf{z}_i) = \begin{cases} \sum_{k=1}^K (1 - q_k) h_k(t_i \mid z_i) S(t_i \mid z_i) & \text{if } \delta_i = 1, |S_i| > 1, \\ q_k h_k(t_i \mid z_i) S(t_i \mid z_i) & \text{if } \delta_i = 1, |S_i| = 1, \\ S(t_i \mid z_i) & \text{otherwise.} \end{cases}$$

where $S(t_i \mid z_i) = \exp \left\{ - \sum_{k=1}^K \int_0^{t_i} h_k(v \mid z) dv \right\}$.

Now, the likelihood of competing risks data can be written as

$$\begin{aligned} L_c(\Psi_2 \mid \delta, \vartheta^R, \mathbf{z}) &= \prod_{i=1}^n L_{ci}(\Psi_2 \mid \delta_i, \vartheta_i^R, \mathbf{z}_i) \\ &= \prod_{i=1}^n \left\{ \sum_{k \in S_i} h_k(t_i \mid z_i) P(S_i \mid z_i, K_i = k) \right\}^{\delta_i} \exp \left\{ - \sum_{k=1}^K \int_0^{t_i} h_k(v \mid z) dv \right\}, \end{aligned} \tag{5.8}$$

where K_i denotes the true cause of failure for i^{th} individual and $\Psi_2 = (\lambda_k, \alpha_k, \beta_k, q_k)$ be the vector of model parameters for competing risks data.

It can be observed that the likelihood function (5.8) for competing risks data is not in a convenient form for joint modelling with longitudinal data. So it becomes necessary that first, we deal with the terms related to masked observations in the likelihood. Hence, to represent the likelihood expression (5.8) in an analytically convenient form, we introduce a latent random variable U_i for all $i \in S_m$, following multinomial distribution. For i^{th} individual in S_m , U_i follows $\mathcal{MD}(1; p_{i1}, \dots, p_{iK})$ and the probability of successes can be

defined as

$$p_{ik} = \frac{(1 - q_k)h_k(t_i | z_i)}{\sum_{k=1}^K (1 - q_k)h_k(t_i | z_i)}; \quad i \in S_m \text{ and } k = 1, \dots, K. \quad (5.9)$$

Now by using the posterior probabilities defined in (5.9) and the corresponding value of latent variable U_i , one can assign a cause k to each masked observations in S_m . So, the likelihood function (5.8) can be written explicitly by using U_i and respectively the probability of successes p_{ik} . Hence, the likelihood for competing risks sub-model can be rewritten as

$$L_c(\Psi_2 | \delta, \vartheta^R, \mathbf{z}) = \prod_{i=1}^n \prod_{k=1}^K \left\{ \prod_{i \in S_k} \{q_k h_k(t_i | z_i)\} \prod_{i \in S_m} \{(1 - q_k)h_k(t_i | z_i)\}^{I[u_i=1, K_i=k]} \right\} \\ \prod_{i=1}^n \exp \left\{ - \sum_{k=1}^K \int_0^{t_i} h_k(v | z) dv \right\}, \quad (5.10)$$

where the indicator function $I[u_i = 1, K_i = k]$ is defined for masked set S_m and it takes value 1 if for $i \in S_m$, the true cause of failure is k otherwise 0.

5.3.3 Joint Inference of Longitudinal and Competing Risks Data

In the previous subsections, we have constructed the models for longitudinal data and competing risks data. To analyse the longitudinal data and competing risks data by formulating a joint model incorporating the case of cause-dependent masking, we assume that the association between both outcomes is induced by the random effects and common covariates. Hence, the association between longitudinal data and competing risks data is done with the shared-parameter model approach (see Sheikh et al. 2021). In this method, the random effects ϑ^R which serve as a parameter in the sub-model for longitudinal data, are induced in the sub-model for competing risks data as covariates. One can also use correlated random effects to link the two outcomes.

The density of random effects ϑ_i^R is given by

$$f(\vartheta_i^R) = \frac{1}{(2\pi)^{\frac{p+1}{2}}} \exp \left(-\frac{1}{2} \vartheta_i^{R'} \vartheta_i^R \right); \quad -\infty < \vartheta_i^R < \infty, \quad i = 1, \dots, n. \quad (5.11)$$

Assuming the conditional independence between the two outcomes, i.e., conditional on random effects and covariates, longitudinal data and competing risks data are assumed to be independent. The joint likelihood of $\Psi = (\Psi_1, \Psi_2)$ for observed data $(y_i, x_i, t_i, s_i, \delta_i, z_i)$ is given by

$$\begin{aligned} L(\Psi | data) &= L_{ld}(\Psi_1 | Y, \vartheta^R, \mathbf{X}) L_c(\Psi_2 | \delta, \vartheta^R, \mathbf{z}) L_R(\vartheta^R) \\ &= \prod_{i=1}^n \left[\left(\frac{1}{2\pi\sigma^2} \right)^{m_i/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{m_i} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma\vartheta_i^R) - X'_i\gamma)^2 \right\} \right] \\ &\quad \prod_{k=1}^K \left\{ \prod_{i \in S_k} \left\{ q_k \lambda_k t_i^{\lambda_k-1} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) \right\} \prod_{i \in S_m} \left\{ (1-q_k) \lambda_k t_i^{\lambda_k-1} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) \right\}^{I[u_i=1, K_i=k]} \right\} \\ &\quad \exp \left\{ -\sum_{k=1}^K t_i^{\lambda_k} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) \right\} \frac{1}{(2\pi)^{\frac{p+1}{2}}} \exp \left(-\frac{1}{2} \vartheta_i^R' \vartheta_i^R \right), \end{aligned} \quad (5.12)$$

where $L_{ld}(\Psi_1 | Y, \vartheta^R, \mathbf{X})$ is product of densities of repeated measurements on longitudinal response $Y_i; i = 1, \dots, n$, $L_c(\Psi_2 | \delta, \vartheta^R, \mathbf{z})$ is likelihood function for competing risks sub-model defined in (5.10) and $L_R(\vartheta^R)$ is density of multivariate standard normal distribution.

The unknown parameters of the joint model can be estimated by maximising the logarithmic of the likelihood function (5.12). The maximization can be done mostly by using Newton's method or the expectation-maximization (*EM*) algorithm. However, the Bayesian method has some advantages over the classical methods in terms of computational cost and implementation. Furthermore, the integrals of high dimension involved in the model do not need to be approximated with respect to unobserved random effects ϑ^R . The present form of the likelihood function (5.12) is convenient for the posterior analysis. Hence, we utilize the likelihood function in the next section to establish the posterior densities and the Bayesian estimates of the unknown parameters (Ψ_1, Ψ_2) of the joint model will be obtained accordingly.

5.4 Bayesian Estimation

In statistical inference, the Bayesian method is extensively used for parameter estimation and it becomes more reasonable when the dimension of the parameter space is quite large. As in the case of joint modelling, the parameter space involves the parameters of the longitudinal

sub-model as well as of the survival sub-model, we specify the priors and consequently, the posterior densities are derived for the parameters of both sub-models.

5.4.1 Prior Elicitation

For Bayesian inference, we choose informative priors for the parameters of the joint model given by

- (i) For b_{ll} ; $l = 1, \dots, p + 1$, the prior is normal distribution with mean $\mu_{b_{ll}}$ and variance $\sigma_{b_{ll}}^2$, truncated at point zero; i.e., $b_{ll} \sim \mathcal{TN}(\mu_{b_{ll}}, \sigma_{b_{ll}}^2, 0, \infty)$

$$\pi(b_{ll}) \propto \exp \left\{ -\frac{1}{2} \left(\frac{b_{ll} - \mu_{b_{ll}}}{\sigma_{b_{ll}}^2} \right)^2 \right\}; \quad b_{ll} > 0.$$

- (ii) For b_{lr} ; $l \neq r = 1, \dots, p + 1$, the prior is normal distribution; i.e., $b_{lr} \sim \mathcal{N}(\mu_{b_{lr}}, \sigma_{b_{lr}}^2)$

$$\pi(b_{lr}) \propto \exp \left\{ -\frac{1}{2} \left(\frac{b_{lr} - \mu_{b_{lr}}}{\sigma_{b_{lr}}^2} \right)^2 \right\}; \quad -\infty < b_{lr} < \infty.$$

- (iii) We choose independent normal priors for parameters $\gamma, \vartheta, \alpha_1, \alpha_2, \beta_1, \beta_2$. Specifically, we have

$$\eta \sim \mathcal{N}(\mu_\eta, \Sigma_\eta);$$

where $\eta \in \{\gamma, \vartheta, \alpha_1, \alpha_2, \beta_1, \beta_2\}$.

- (iv) For shape parameter λ_k ; $1, \dots, K$, we choose gamma prior, i.e., $\lambda_k \sim \mathcal{G}(\omega_k, \nu_k)$ as

$$\pi(\lambda_k) \propto \lambda_k^{\omega_k - 1} \exp(-\nu_k \lambda_k); \quad \omega_k, \nu_k > 0.$$

- (v) For variance of measurement error, we choose inverse gamma prior, i.e., $\sigma^2 \sim \mathcal{IG}(a_0, b_0)$

$$\pi(\sigma^2) \propto (\sigma^2)^{-(a_0+1)} \exp \left(-\frac{b_0}{\sigma^2} \right); \quad \sigma^2 > 0.$$

(vi) We choose beta prior for the identification probabilities, i.e., $q_k \sim \mathcal{B}(c_k, d_k)$,

$$\pi(q_k) \propto q_k^{c_k-1} (1-q_k)^{d_k-1}; \quad 0 < q_k < 1; \quad k = 1, \dots, K.$$

5.4.2 Posterior Inference

The joint posterior density can be written as

$$p(\Psi | data) \propto L(\Psi | data) \pi(\Psi). \quad (5.13)$$

Assuming a priori independence between Ψ_1 and Ψ_2 , $\pi(\Psi)$ can be written as

$$\pi(\Psi) = \pi(\Psi_1)\pi(\Psi_2);$$

and further, it can be written as the product of the priors of each parameter component.

The posterior samples on the model parameters (Ψ_1, Ψ_2) can be generated as follows

(i) The conditional posterior density of b_{ll} ; $l = 1, \dots, p + 1$, can be written as

$$\begin{aligned} p(b_{ll} | data) &\propto L(\Psi_1, \Psi_2 | data) \pi(b_{ll}) \\ &\propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma\vartheta_i^R) - \mathbf{X}'_i\gamma)^2 \right\} \exp \left\{ -\frac{1}{2} \left(\frac{b_{ll} - \mu_{b_{ll}}}{\sigma_{b_{ll}}} \right)^2 \right\} \\ &\propto \exp \left\{ b_{ll} \left(\frac{\mu_{b_{ll}}}{\sigma_{b_{ll}}} + \frac{1}{\sigma^2} \sum_{i=1}^n \vartheta_{il}^R \sum_{j=1}^{m_i} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma\vartheta_i^R) + b_{ll}\vartheta_{il}^R\tau_{ij}^{l-1} + \mathbf{X}'_i\gamma) \right) \right. \\ &\quad \left. - \frac{1}{2} b_{ll}^2 \left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{il}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + \frac{1}{\sigma_{b_{ll}}^2} \right) \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left(b_{ll} - \frac{\frac{\mu_{b_{ll}}}{\sigma_{b_{ll}}} + \sum_{i=1}^n \vartheta_{il}^R \sum_{j=1}^{m_i} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma\vartheta_i^R) + b_{ll}\vartheta_{il}^R\tau_{ij}^{l-1} + \mathbf{X}'_i\gamma)/\sigma^2}{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{il}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + \frac{1}{\sigma_{b_{ll}}^2} \right)} \right)^2 \right. \\ &\quad \left. \left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{il}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + 1 \right) \right\}. \end{aligned} \quad (5.14)$$

Hence, $b_{ll} \sim \mathcal{T}\mathcal{N}(\mu_{b_{ll|}}, \sigma_{b_{ll|}}^2, 0, \infty)$; where

$$\mu_{b_{ll|}} = \frac{\frac{\mu_{b_{ll}}}{\sigma_{b_{ll}}} + \sum_{i=1}^n \vartheta_{il}^R \sum_{j=1}^{m_i} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma \vartheta_i^R) + b_{ll} \vartheta_{il}^R \tau_{ij}^{l-1} + \mathbf{X}'_i \gamma) / \sigma^2}{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{il}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + \frac{1}{\sigma_{b_{ll}}^2} \right)},$$

and

$$\sigma_{b_{ll|}}^2 = \frac{1}{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{il}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + \frac{1}{\sigma_{b_{ll}}^2} \right)}.$$

(ii) The conditional posterior density of b_{lr} ; $l \neq r = 1, \dots, p+1$, is given by

$$p(b_{lr} | data) \propto \exp \left\{ -\frac{1}{2} \left(b_{lr} - \frac{\frac{\mu_{b_{lr}}}{\sigma_{b_{lr}}} + \sum_{i=1}^n \vartheta_{ir}^R \sum_{j=1}^{m_i} \tau_{ij} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma \vartheta_i^R) + b_{lr} \vartheta_{ir}^R \tau_{ij}^{l-1} + \mathbf{X}'_i \gamma) / \sigma^2}{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{ir}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + 1 \right)} \right)^2 \right. \\ \left. \left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{ir}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + 1 \right) \right\}. \quad (5.15)$$

Hence, $b_{lr} \sim \mathcal{N}(\mu_{b_{lr}}, \sigma_{b_{lr}}^2)$; where

$$\mu_{b_{lr}} = \frac{\frac{\mu_{b_{lr}}}{\sigma_{b_{lr}}} + \sum_{i=1}^n \vartheta_{ir}^R \sum_{j=1}^{m_i} \tau_{ij} (y_{ij} - \mathcal{A}_i(\tau)'(\vartheta + \Gamma \vartheta_i^R) + b_{lr} \vartheta_{ir}^R \tau_{ij}^{l-1} + \mathbf{X}'_i \gamma) / \sigma^2}{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{ir}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + \frac{1}{\sigma_{b_{lr}}^2} \right)},$$

and

$$\sigma_{b_{lr}}^2 = \frac{1}{\left(\frac{\sum_{i=1}^n \sum_{j=1}^{m_i} \vartheta_{ir}^{R2} \tau_{ij}^{2(l-1)}}{\sigma^2} + \frac{1}{\sigma_{b_{lr}}^2} \right)}.$$

(iii) The conditional posterior density of γ is defined by

$$\gamma \sim \mathcal{N} \left(V_\gamma \left\{ \frac{1}{\sigma^2} \sum_{i=1}^n \mathbf{X}_i (\mathbf{y}_i - \mathcal{A}_i(\tau)'(\vartheta + \Gamma \vartheta_i^R)) \right\}, V_\gamma \right), \quad (5.16)$$

where,

$$V_\gamma = \left(\frac{1}{\sigma^2} \sum_{i=1}^n \mathbf{X}_i \mathbf{X}'_i + \Sigma_\gamma^{-1} \right)^{-1}.$$

(iv) The conditional posterior density of ϑ is defined by

$$\vartheta \sim \mathcal{N} \left(V_\vartheta \left\{ \frac{1}{\sigma^2} \sum_{i=1}^n \mathcal{A}_i(\tau) (\mathbf{y}_i - \mathcal{A}_i(\tau)' \Gamma \vartheta_i^R - \mathbf{X}'_i \gamma) \right\}, V_\vartheta \right), \quad (5.17)$$

where,

$$V_\vartheta = \left(\frac{1}{\sigma^2} \sum_{i=1}^n \mathcal{A}_i(\tau) \mathcal{A}_i(\tau)' + \Sigma_\vartheta^{-1} \right)^{-1}.$$

(v) The conditional posterior density of σ^2 is given by

$$\begin{aligned} p(\sigma^2 | data) &\propto L(\Psi_1, \Psi_2 | data) \pi(\sigma^2) \\ &\propto \exp \left\{ -\frac{1}{\sigma^2} \left(b_0 + \frac{1}{2} \sum_{i=1}^n (y_i - \mathcal{A}_i(\tau)' (\vartheta + \Gamma \vartheta_i^R) - \mathbf{X}'_i \gamma)' \right. \right. \\ &\quad \left. \left. (y_i - \mathcal{A}_i(\tau)' (\vartheta + \Gamma \vartheta_i^R) - \mathbf{X}'_i \gamma) \right) \right\} \left(\frac{1}{\sigma^2} \right)^{(a_0 + \frac{1}{2} \sum_{i=1}^n m_i + 1)}. \end{aligned} \quad (5.18)$$

Thus,

$$\sigma^2 | data \sim \mathcal{IG} \left(a_0 + \frac{1}{2} \sum_{i=1}^n m_i, b_0 + b_0^* \right),$$

where, $b_0^* = \frac{1}{2} \sum_{i=1}^n (y_i - \mathcal{A}_i(\tau)' (\vartheta + \Gamma \vartheta_i^R) - \mathbf{X}'_i \gamma)' (y_i - \mathcal{A}_i(\tau)' (\vartheta + \Gamma \vartheta_i^R) - \mathbf{X}'_i \gamma)$.

(vi) The conditional posterior density for λ_k ; $k = 1, \dots, K$ is

$$\begin{aligned} p(\lambda_k | \alpha_k, \beta_k, data) &\propto \lambda_k^{n_k + n_k^* + \omega_k - 1} \prod_{i \in S_k} t_i^{\lambda_k - 1} \prod_{i \in S_m} \left\{ t_i^{\lambda_k - 1} \right\}^{I[u_i=1, K_i=k]} \\ &\quad \exp \left\{ - \left(\nu_k \lambda_k + \sum_{i=1}^n t_i^{\lambda_k} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) \right) \right\} \end{aligned} \quad (5.19)$$

where $n_k^* = \sum_{i=1}^{n_m} I[u_i = 1, K_i = k]$ is the number of individuals who are masked due to cause k .

(vii) The conditional posterior density for α_k ; $k = 1, \dots, K$ is

$$p(\alpha_k | h_{0k}, \beta_k, \vartheta^R, data) \propto \exp \left\{ \sum_{i \in S_k} (\alpha'_k \vartheta_i^R + z'_i \beta_k) + \sum_{i \in S_m} I[u_i = 1, K_i = k] (\alpha'_k \vartheta_i^R + z'_i \beta_k) \right\}$$

$$\left. - \sum_{i=1}^n t_i^{\lambda_k} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) - \frac{1}{2} (\alpha_k - \mu_{\alpha_k})' \Sigma_{\alpha_k}^{-1} (\alpha_k - \mu_{\alpha_k}) \right\}. \quad (5.20)$$

(viii) The conditional posterior density for $\beta_k; k = 1, \dots, K$ is

$$\begin{aligned} p(\beta_k | h_{0k}, \alpha_k, \vartheta^R, data) \propto & \exp \left\{ \sum_{i \in S_k} (\alpha'_k \vartheta_i^R + z'_i \beta_k) + \sum_{i \in S_m} I[u_i = 1, K_i = k] (\alpha'_k \vartheta_i^R + z'_i \beta_k) \right. \\ & \left. - \sum_{i=1}^n t_i^{\lambda_k} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) - \frac{1}{2} (\beta_k - \mu_{\beta_k})' \Sigma_{\beta_k}^{-1} (\beta_k - \mu_{\beta_k}) \right\}. \end{aligned} \quad (5.21)$$

(ix) The conditional posterior density for identification probabilities, $q_k; k = 1, \dots, K$ is given by

$$\begin{aligned} p(q_k | data) &\propto q_k^{n_k} (1 - q_k)^{n_k^*} q_k^{c_k - 1} (1 - q_k)^{d_k - 1} \\ &\propto q_k^{n_k + c_k - 1} (1 - q_k)^{n_k^* + d_k - 1}, \end{aligned} \quad (5.22)$$

Thus one can directly generate samples on q_k as

$$q_k | data \sim \mathcal{B}(n_k + c_k, n_k^* + d_k); \quad k = 1, \dots, K.$$

(x) The posterior density for $\vartheta_i^R; i = 1, \dots, n$ is given by

$$\begin{aligned} p(\vartheta_i^R | \Psi_1, \Psi_2, data) \propto & \exp \left\{ \delta_i \sum_{k=1}^K I[u_i = 0, K_i = k] (\alpha'_k \vartheta_i^R + z'_i \beta_k) - \frac{1}{2} \vartheta_i^{R'} \vartheta_i^R \right. \\ & + \delta_i \sum_{k=1}^K I[u_i = 1, K_i = k] (\alpha'_k \vartheta_i^R + z'_i \beta_k) - \sum_{k=1}^K t_i^{\lambda_k} \exp(\alpha'_k \vartheta_i^R + z'_i \beta_k) \\ & \left. - \frac{1}{2\sigma^2} (y_i - \mathcal{A}_i(\tau)'(\vartheta + \Gamma \vartheta_i^R) - X_i' \gamma)' (y_i - \mathcal{A}_i(\tau)'(\vartheta + \Gamma \vartheta_i^R) - X_i' \gamma) \right\}. \end{aligned} \quad (5.23)$$

The Bayesian method of estimation of parameters involves drawing posterior samples from the full conditionals and we use the M-H algorithm and Gibbs sampler (Chen et al. 2000) for the computational purpose. Since the full conditional distributions of the parameters

$\Gamma, \gamma, \vartheta, \sigma^2, q_1$ and q_2 are standard distributions, so we can draw posterior samples from their full conditional distributions directly. For the rest of the parameters and the random effects θ_i^R , we use the M-H algorithm with the normal approximation to their full conditional distributions as the candidate distribution. The initial values of the parameters can be obtained by modelling the longitudinal and competing risks data separately by a *LMM* and Cox proportional cause-specific hazard model, respectively.

5.5 Simulation Study

In this section, we perform an extensive simulation study in order to demonstrate the empirical properties of our proposed joint model constructed in Subsection (5.3.3) under the Bayesian paradigm. For this, we define the values of the design parameters of the two sub-models in order to generate the longitudinal data and competing risks data with cause-dependent masking.

The trajectory function of longitudinal response Y , is taken to be linear, i.e., $\mathcal{A}_i(\tau) = (\mathbf{1}^\top \tau)$. The elements of the variance-covariance matrix Σ are constrained by the fact that the matrix should be positive definite. This leads to slow convergence of *MCMC* sample in Bayesian estimation of Σ as it depends on the unobserved random effects. Further, the problem becomes more complex if the dimension of the random effects becomes large. Hence, we use Cholesky decomposition of Σ , and write $\Sigma = \Gamma\Gamma'$, so that re-parametrized quantity is given by $\vartheta_i^* = \Gamma\vartheta_i^R$; where $\vartheta_i^R \sim \mathcal{N}(0, I_p)$ and Γ is lower triangular matrix with positive diagonal elements. The values of longitudinal and survival parameters are chosen arbitrarily. The design values of parameters of the longitudinal sub-model are given as $\boldsymbol{\gamma} = (\gamma_1, \gamma_2) = (0.15, 0.3)$, $\boldsymbol{\theta} = (\theta_0, \theta_1) = (0.5, 1)$, $\boldsymbol{\Gamma} = \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.8 \end{bmatrix}$ and $\sigma^2 = 0.05$. Similarly, the design values of parameters of the competing risks sub-model are given as $\boldsymbol{\lambda}_1 = 4.8, \boldsymbol{\lambda}_2 = 5.8, \boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12}) = (0.5, 0.6), \boldsymbol{\beta}_2 = (\beta_{21}, \beta_{22}) = (0.2, -0.5), \boldsymbol{\alpha}_1 = (\alpha_{11}, \alpha_{12}) = (0.2, 0.8), \boldsymbol{\alpha}_2 = (\alpha_{21}, \alpha_{22}) = (0.5, 1)$ and we take $(q_1, q_2) = \{(0.85, 0.78), (0.65, 0.55), (0.35, 0.25)\}$ to illustrate the performance of joint model for different proportion of masked observations.

- (i) The longitudinal measurements are simulated from the sub-model

$$Y_i(\tau) = (\mathbf{1} \ \tau)(\vartheta + \Gamma \vartheta_i^R) + \mathbf{X}'_i \gamma + \epsilon_i(\tau), \quad (5.24)$$

where the measurements are generated for i^{th} individual at observation points at $\tau_{ij} = 28(j - 1)/365$; $i = 1, \dots, n$ and $j = 1, \dots, m_i$.

- (ii) The dimension of the random effects is taken to be 2, i.e., $p = 1$ and we generate $\vartheta_i^R = (\vartheta_{i0}^R, \vartheta_{i1}^R) \sim \mathcal{N}(0, \mathcal{J}_2)$. Further, the number of covariates is fixed at 2, i.e., $q = 2$ and generate the time-invariant covariates $\mathbf{X}_i = (X_{i1}, X_{i2})$ with $x_{i1} \sim \mathcal{N}(0, 1)$ and $x_{i2} | x_{i1} \sim \mathcal{N}(0.2x_{i1}, 1)$.

For the competing risks data sub-model, the methodology is developed for K competing causes, but we fix $K = 2$ to carry out the numerical study. Generate the time-to-event data corresponding to k^{th} cause,

$$T_{ik}^* \sim \mathcal{W} \left(\lambda_k, \exp \left(-\frac{\alpha'_k \vartheta_i^R + z'_k \beta_k}{\lambda_k} \right) \right); k = 1, 2.$$

For the competing risks data, we consider the same covariates as in longitudinal data, i.e., $Z_i = X_i$. The corresponding censoring time C_i is generated from the exponential distribution with rate parameter 0.5 so that the censoring proportion in the simulated data is maintained at around 30% as per the sample of size n . Thus, the observed time to failure is given by $t_i = \min(t_{i1}^*, t_{i2}^*, c_i)$; where

$$c_i^* = \min(c_i, \max_{1 \leq j \leq m_i} (\tau_{ij}))$$

and the corresponding cause of failure is also noted. From the set of individuals failed from cause 1 (cause 2), we randomly mask around 25% (40%) individuals accordingly as the identification probabilities are prefixed as $(q_1, q_2) = (0.65, 0.45)$. We simulated data on $n = 100$ individuals and for the longitudinal data, we have considered the balanced study design and generate $m_i = 10$ ($i = 1, \dots, 100$) repeated observations on longitudinal response Y_i for each individual. Based on the *MCMC* samples drawn from the conditional posteriors as defined in (3.2), the inferences on posterior distribution of the interested parameters

are made. The convergence of the posterior samples are governed by the cumsumplot, *ACF*, and Gelman and Rubin's test statistics (Gelman and Rubin 1992). We calculated the estimates of parameters under the Bayesian approach, along with their *MSE* and bias as well as the *AL* of 95% *HPD* intervals with their *CP* are reported from Table 5.1 to Table 5.3. The simulation study is based on 1000 samples to average out the sampling fluctuation and consequently to achieve consistency in the estimates. Figure 5.1 shows the trajectory

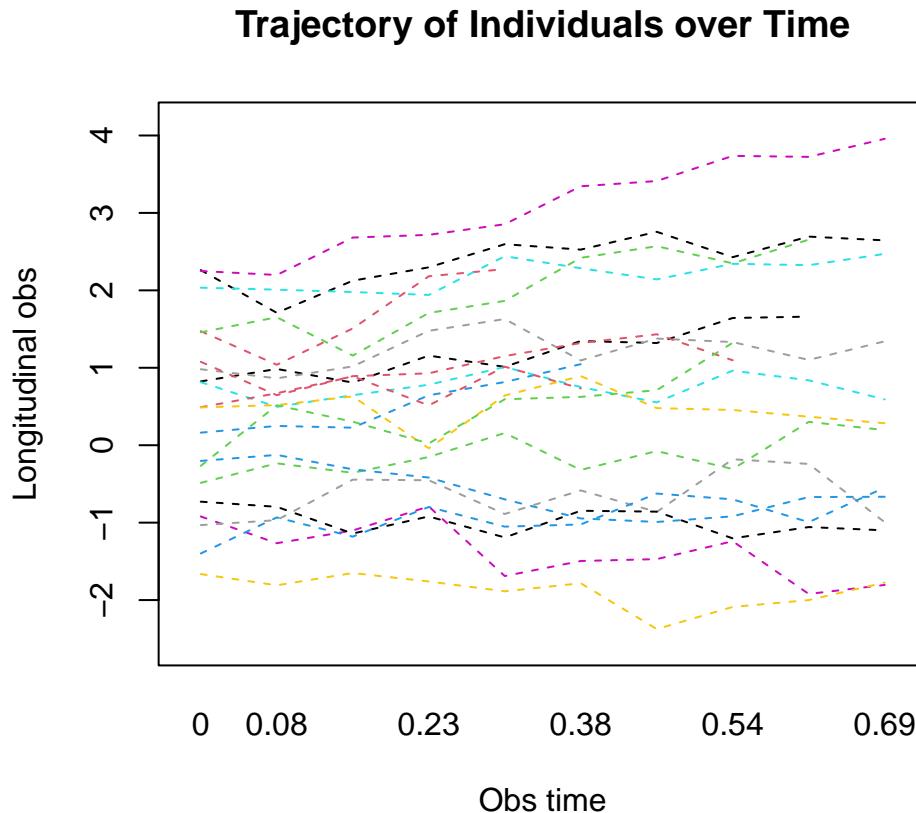


Figure 5.1: Trajectory of the longitudinal response for 20 randomly selected individuals from simulated data

of longitudinal response over time for 20 randomly selected individuals from the simulated data. It depicts the deviation in the pattern of changes in response variable over time among the individuals.

From Table 5.1 to Table 5.3, it is clear that the all the estimators of fixed effect covariates for both the processes performing well. Also, the *CP* of 95% *HPD* for longitudinal sub-model parameters converges to nominal level. For the parameters of competing risks sub-model,

Table 5.1: Bayes estimates and their bias, *MSE*, *AL* and *CP* of the parameters of the joint model for $(q_1, q_2) = (0.85, 0.78)$.

Longitudinal Sub-model						
Parameter	True	Est.	Bias	MSE	AL	CP
b_{11}	1.00	1.00	-0.0001	0.0000	0.0330	0.95
b_{21}	0.50	0.50	-0.0009	0.0002	0.0867	0.99
b_{22}	0.80	0.80	0.0014	0.0005	0.0896	0.96
γ_1	0.15	0.15	0.0006	0.0001	0.1508	0.95
γ_2	0.30	0.30	0.0000	0.0001	0.1503	0.96
θ_0	0.50	0.50	-0.0007	0.0002	0.4502	0.95
θ_1	1.00	0.99	0.0016	0.0015	0.5452	0.96
σ^2	0.05	0.05	0.0002	0.0000	0.0103	0.95
Competing Risks Sub-Model						
Parameter	True	Est.	Bias	MSE	AL	CP
λ_1	4.80	4.65	-0.1493	0.0653	1.3736	0.78
λ_2	5.80	5.56	-0.2417	0.0961	1.5079	0.81
α_{11}	0.20	0.20	0.0025	0.0109	0.3495	0.88
α_{12}	0.80	0.84	0.0382	0.0140	0.3515	0.86
α_{21}	0.50	0.50	0.0025	0.0113	0.3562	0.89
α_{22}	1.00	1.02	0.0201	0.0112	0.3563	0.91
β_{11}	0.50	0.51	0.0139	0.0213	0.3262	0.78
β_{12}	0.60	0.63	0.0267	0.0184	0.3289	0.81
β_{21}	0.20	0.19	-0.0055	0.0225	0.3091	0.74
β_{22}	-0.50	-0.50	-0.0007	0.0217	0.3178	0.77
q_1	0.85	0.81	-0.0414	0.0066	0.3109	0.98
q_2	0.78	0.73	-0.0506	0.0094	0.3930	0.97

Table 5.2: Bayes estimates and their bias, MSE , AL and CP of the parameters of the joint model for $(q_1, q_2) = (0.65, 0.55)$.

Longitudinal Sub-model						
Parameter	True	Est.	Bias	MSE	AL	CP
b_{11}	1.00	1.00	0.0002	0.0001	0.0330	0.94
b_{21}	0.50	0.50	0.0005	0.0002	0.0869	0.99
b_{22}	0.80	0.80	0.0008	0.0005	0.0894	0.95
γ_1	0.15	0.15	0.0002	0.0001	0.1505	0.96
γ_2	0.30	0.30	0.0000	0.0001	0.1499	0.93
θ_0	0.50	0.50	0.0003	0.0002	0.4523	0.96
θ_1	1.00	1.00	0.0005	0.0015	0.5473	0.95
σ^2	0.05	0.05	0.0003	0.0000	0.0103	0.93
Competing Risks Sub-Model						
Parameter	True	Est.	Bias	MSE	AL	CP
λ_1	4.80	4.66	-0.1354	0.0600	1.3904	0.80
λ_2	5.80	5.56	-0.2398	0.0922	1.5350	0.80
α_{11}	0.20	0.21	0.0076	0.0101	0.3530	0.90
α_{12}	0.80	0.83	0.0336	0.0133	0.3505	0.88
α_{21}	0.50	0.51	0.0070	0.0105	0.3536	0.88
α_{22}	1.00	1.01	0.0131	0.0133	0.3533	0.90
β_{11}	0.50	0.52	0.0205	0.0189	0.3263	0.80
β_{12}	0.60	0.63	0.0273	0.0205	0.3290	0.80
β_{21}	0.20	0.20	-0.0026	0.0233	0.3088	0.73
β_{22}	-0.50	-0.50	0.0026	0.0217	0.3164	0.75
q_1	0.65	0.66	0.0111	0.0087	0.3894	0.95
q_2	0.55	0.58	0.0314	0.0110	0.4447	0.95

Table 5.3: Bayes estimates and their bias, MSE , AL and CP of the parameters of the joint model for $(q_1, q_2) = (0.35, 0.25)$.

Longitudinal Sub-model						
Parameter	True	Est.	Bias	MSE	AL	CP
b_{11}	1.00	1.00	0.0000	0.0001	0.0330	0.94
b_{21}	0.50	0.50	0.0002	0.0001	0.0868	0.99
b_{22}	0.80	0.80	0.0018	0.0005	0.0893	0.95
γ_1	0.15	0.15	-0.0003	0.0001	0.1496	0.95
γ_2	0.30	0.30	0.0003	0.0001	0.1493	0.95
θ_0	0.50	0.50	-0.0002	0.0002	0.4522	0.95
θ_1	1.00	1.00	0.0002	0.0015	0.5472	0.96
σ^2	0.05	0.04	0.0002	0.0000	0.0103	0.95
Competing Risks Sub-model						
Parameter	True	Est.	Bias	MSE	AL	CP
λ_1	4.80	4.67	-0.1311	0.0575	1.4227	0.79
λ_2	5.80	5.55	-0.2466	0.0919	1.5761	0.78
α_{11}	0.20	0.21	0.0148	0.0116	0.3480	0.89
α_{12}	0.80	0.84	0.0428	0.0129	0.3511	0.88
α_{21}	0.50	0.51	0.0100	0.0078	0.3604	0.92
α_{22}	1.00	1.03	0.0257	0.0107	0.3630	0.93
β_{11}	0.50	0.51	0.0146	0.0200	0.3244	0.79
β_{12}	0.60	0.62	0.0225	0.0197	0.3230	0.78
β_{21}	0.20	0.20	-0.0048	0.0216	0.3180	0.76
β_{22}	-0.50	-0.51	-0.0096	0.0177	0.3278	0.79
q_1	0.35	0.43	0.0766	0.0157	0.4124	0.91
q_2	0.25	0.38	0.1339	0.0263	0.4406	0.87

CP does not converge to the nominal level. This may be due to the fact that random effects θ^R are the function of unobserved data and it is considered as a covariate in the sub-model for competing risks. As we increase the proportion of masked observations in the simulated data, the performance of estimators of the parameters of sub-model for competing risks data deteriorates.

5.6 Conclusion

In this chapter, we have extended the usual model for competing risks data to incorporate the effect of covariates on the time to occurrence of the event. We have established a flexible joint model to analyse the competing risks data and longitudinal data simultaneously. The model is flexible in the sense that it is capable of handling the missing causes of the failure in competing risks data along with the cause-dependent masking. The dependence structure between competing risks and longitudinal data is defined with the help of a shared random effect parameter approach. In the joint modelling framework, longitudinal data are analysed using the linear mixed-effect model and the competing risks data is analysed with the Cox-proportional cause-specific hazard model to incorporate the effects of covariates into the risk process. The estimation of parameters of the joint model is performed under the Bayesian framework. An extensive simulation study has been carried out to illustrate the performance of the established model under the Bayesian paradigm. The Bayesian computation in joint modelling concerning posterior sample generation for θ^R is quite challenging. However, Cholesky decomposition of Σ and parallel computing in **R** help reduce the overall computation time efficiently. The problem will become more complex in the case of non-linear trajectory and high dimensional random effects and covariates. There is a need to develop advanced Bayesian algorithms for the efficient computation of the model parameters.

In this chapter, we have assumed that the longitudinal trajectory is linear with respect to time and the measurement error has Gaussian distribution. These assumptions are quite straightforward with respect to the longitudinal response measurements. For future studies,

one can relax these assumptions and explore alternative approaches, such as incorporating the non-linear trajectories and asymmetric assumptions for measurement errors.

References

- Aarset, M. V. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability* **36(1)**: 106–108 (1987).
- Ajit Chaturvedi, G. S., Khumukcham. Bayesian estimation procedures for a family of lifetime distributions under squared-error and entropy losses. *Metron* **64(2)**: 179–198 (2006).
- Alsefri, M., Sudell, M., García-Fiñana, M. and Kolamunnage-Dona, R. Bayesian joint modelling of longitudinal and time to event data: A methodological review. *BMC Medical Research Methodology* **20(1)**: 1–17 (2020).
- Andrinopoulou, E.-R., Rizopoulos, D., Takkenberg, J. J. and Lesaffre, E. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Statistical Methods in Medical Research* **26(4)**: 1787–1801 (2017).
- Balakrishnan, N. On the maximum likelihood estimation of the location and scale parameters of exponential distribution based on multiply type ii censored samples. *Journal of Applied Statistics* **17(1)**: 55–61 (1990).
- Balakrishnan, N. and Aggarwala, R. *Progressive censoring: Theory, methods, and applications*. Springer Science & Business Media (2000).
- Barlow, R. E. and Proschan, F. Statistical theory of reliability and life testing: Probability models. Technical report, Florida State Univ Tallahassee (1975).
- Barnwal, V. and Panwar, M. Competing risks analysis for dependent causes using Marshall-Olkin bivariate generalized lifetime family. *Communications in Statistics-Theory and Methods* pp. 1–29 (2022).

- Basu, A. and Ghosh, J. Identifiability of the multinormal and other distributions under competing risks model. *Journal of Multivariate Analysis* **8(3)**: 413–429 (1978).
- Berger, J. The case for objective Bayesian analysis. *Bayesian Analysis* **1(3)**: 385–402 (2006).
- Berger, J. O. and Bernardo, J. M. On the development of the reference prior method. *Bayesian Statistics* **4(4)**: 35–60 (1992).
- Berger, J. O., De Oliveira, V. and Sansó, B. Objective Bayesian analysis of spatially correlated data. *Journal of the American Statistical Association* **96(456)**: 1361–1374 (2001).
- Bernardo, J. M. Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society: Series B (Methodological)* **41(2)**: 113–128 (1979).
- Byar, D. P. and Green, S. B. The choice of treatment for cancer patients based on covariate information. *Bulletin du cancer* **67(4)**: 477 (1980).
- Cai, J., Shi, Y. and Liu, B. Analysis of incomplete data in the presence of dependent competing risks from Marshall–Olkin bivariate Weibull distribution under progressive hybrid censoring. *Communications in Statistics-Theory and Methods* **46(13)**: 6497–6511 (2017).
- Chen, D. G. D., Sun, J. and Peace, K. E. *Interval-censored time-to-event data: Methods and applications*. CRC Press (2012).
- Chen, L. M., Ibrahim, J. G. and Chu, H. Sample size and power determination in joint modeling of longitudinal and survival data. *Statistics in Medicine* **30(18)**: 2295–2309 (2011).
- Chen, M. H., Müller, P., Sun, D., Ye, K. and Dey, D. K. *Frontiers of statistical decision making and Bayesian analysis: In Honor of James O. Berger*. Springer Science & Business Media (2010).
- 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).

- Chen, M. H., Shao, Q. M. and Ibrahim, J. G. *Monte Carlo Methods in Bayesian Computation*. Springer-Verlag, New York, Inc. (2000).
- Colosimo, E. A., Fausto, M. A., Freitas, M. A. and Pinto, J. A. Practical modeling strategies for unbalanced longitudinal data analysis. *Journal of Applied Statistics* **39**(9): 2005–2013 (2012).
- 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).
- Cox, D. R. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* **34**(2): 187–202 (1972).
- Crowder, M. On the identifiability crisis in competing risks analysis. *Scandinavian Journal of Statistics* pp. 223–233 (1991).
- Crowder, M. J. *Classical competing risks*. CRC Press (2001).
- 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* pp. 607–620 (1986).
- Elashoff, R., Li, N. et al. *Joint modeling of longitudinal and time-to-event data*. CRC press (2016).
- Elashoff, R. M., Li, G. and Li, N. An approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine* **26**(14): 2813–2835 (2007).
- Elashoff, R. M., Li, G. and Li, N. A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics* **64**(3): 762–771 (2008).
- Feizjavadian, S. and Hashemi, R. Analysis of dependent competing risks in the presence of progressive hybrid censoring using Marshall–Olkin bivariate Weibull distribution. *Computational Statistics & Data Analysis* **82**: 19–34 (2015).

- Fu, J., Xu, A. and Tang, Y. Objective Bayesian analysis of Pareto distribution under progressive type-II censoring. *Statistics & Probability Letters* **82**(10): 1829–1836 (2012).
- Gelfand, A. E. and Ghosh, S. K. Model choice: A minimum posterior predictive loss approach. *Biometrika* **85**(1): 1–11 (1998).
- Gelman, A., Meng, X.-L. and Stern, H. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica sinica* pp. 733–760 (1996).
- 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).
- Geman, S. and Geman, D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence* **(6)**: 721–741 (1984).
- Ghosh, J. and Mukerjee, R. Noninformative priors (with discussion). In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.). *Bayesian Statistics* **4**: 195–210 (1992).
- Ghosh, M. Objective priors: An introduction for frequentists. *Statistical Science* **26**(2): 187–202 (2011).
- 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).
- Goldman, A. I., Carlin, B. P., Crane, L. R., Launer, C., Korvick, J. A., Deyton, L. and Abrams, D. I. Response of CD4 lymphocytes and clinical consequences of treatment using ddi or ddc in patients with advanced hiv infection. *JAIDS Journal of Acquired Immune Deficiency Syndromes* **11**(2): 161–169 (1996).
- Guan, Q., Tang, Y. and Xu, A. Objective Bayesian analysis for bivariate Marshall–Olkin exponential distribution. *Computational Statistics & Data Analysis* **64**: 299–313 (2013).
- Guess, F. M., Usher, J. S. and Hodgson, T. J. Estimating system and component reliabilities under partial information on cause of failure. *Journal of Statistical Planning and Inference* **29**(1-2): 75–85 (1991).

- Guttman, I. The use of the concept of a future observation in goodness-of-fit problems. *Journal of the Royal Statistical Society: Series B (Methodological)* **29**(1): 83–100 (1967).
- Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1): 97–109 (1970).
- Henderson, R., Diggle, P. and Dobson, A. Joint modelling of longitudinal measurements and event time data. *Biostatistics* **1**(4): 465–480 (2000).
- Hope, A. C. A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society: Series B (Methodological)* **30**(3): 582–598 (1968).
- Hu, W., Li, G. and Li, N. A Bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine* **28**(11): 1601–1619 (2009).
- Huang, X., Li, G., Elashoff, R. M. and Pan, J. A general joint model for longitudinal measurements and competing risks survival data with heterogeneous random effects. *Lifetime Data Analysis* **17**(1): 80–100 (2011a).
- Huang, Y., Dagne, G. and Wu, L. Bayesian inference on joint models of HIV dynamics for time-to-event and longitudinal data with skewness and covariate measurement errors. *Statistics in Medicine* **30**(24): 2930–2946 (2011b).
- Ibrahim, J. G., Chen, M. H. and Sinha, D. Joint models for longitudinal and survival data. In *Bayesian Survival Analysis*, pp. 262–289. Springer (2001).
- IIPS and ICF. India National Family Health Survey NFHS-4 2015–16. *Mumbai: IIPS and ICF* pp. 1255–9 (2017).
- Jaynes, E. T. Prior probabilities. *IEEE Transactions on systems science and cybernetics* **4**(3): 227–241 (1968).
- Jeffreys, H. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **186**(1007): 453–461 (1946).

- Jewell, N. P. and van der Laan, M. Current status data: Review, recent developments and open problems. *Handbook of Statistics* **23**: 625–642 (2003).
- Kass, R. E. and Wasserman, L. The selection of prior distributions by formal rules. *Journal of the American Statistical Association* **91(435)**: 1343–1370 (1996).
- 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).
- Kodell, R. L. and Chen, J. J. Handling cause of death in equivocal cases using the EM algorithm. *Communications in Statistics-Theory and Methods* **16(9)**: 2565–2585 (1987).
- Kundu, D. and Dey, A. K. Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm. *Computational Statistics & Data Analysis* **53(4)**: 956–965 (2009).
- Kundu, D. and Gupta, A. K. Bayes estimation for the Marshall-Olkin bivariate Weibull distribution. *Computational Statistics & Data Analysis* **57(1)**: 271–281 (2013).
- Kundu, D., Kannan, N. and Balakrishnan, N. Analysis of progressively censored competing risks data. *Handbook of Statistics* **23**: 331–348 (2003).
- Kuo, L. and Yang, T. Y. Bayesian reliability modeling for masked system lifetime data. *Statistics & Probability Letters* **47(3)**: 229–241 (2000).
- Laird, N. M. and Ware, J. H. Random-effects models for longitudinal data. *Biometrics* **38(4)**: 963–974 (1982).
- Lawless, J. F. *Statistical Models and Methods for Lifetime Data*. John Wiley and Sons Inc, New York, 2nd edition (2003).
- Li, X., Tang, Y. and Xu, A. Objective Bayesian analysis of Weibull mixture cure model. *Quality Engineering* **32(3)**: 449–464 (2020).
- Lindqvist, B. H. and Skogsrud, G. Modeling of dependent competing risks by first passage times of Wiener processes. *IIE Transactions* **41(1)**: 72–80 (2008).

- Lu, T. Bayesian inference on longitudinal-survival data with multiple features. *Computational Statistics* **32(3)**: 845–866 (2017).
- Marshall, A. W. and Olkin, I. A multivariate exponential distribution. *Journal of the American Statistical Association* **62(317)**: 30–44 (1967).
- 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).
- Miyakawa, M. Analysis of incomplete data in competing risks model. *IEEE Transactions on Reliability* **33(4)**: 293–296 (1984).
- Moeschberger, M. Life tests under dependent competing causes of failure. *Technometrics* **16(1)**: 39–47 (1974).
- Moore, A. H. and Bilikam, J. E. Bayesian estimation of parameters of life distributions and reliability from type II censored samples. *IEEE Transactions on Reliability* **27(1)**: 64–67 (1978).
- Pinheiro, J. C. and Bates, D. M. Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing* **6(3)**: 289–296 (1996).
- Proschan, F. and Sullo, P. Estimating the parameters of a certain multivariate exponential distribution. Technical report, Dept of Statistics, Florida State University Tallahassee (1973).
- 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).
- Rossky, P. J., Doll, J. D. and Friedman, H. L. Brownian dynamics as smart Monte Carlo simulation. *The Journal of Chemical Physics* **69(10)**: 4628–4633 (1978).

- Rue, M., Andrinopoulou, E. R., Alvares, D., Armero, C., Forte, A. and Blanch, L. Bayesian joint modeling of bivariate longitudinal and competing risks data: An application to study patient-ventilator asynchronies in critical care patients. *Biometrical Journal* **59(6)**: 1184–1203 (2017).
- 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).
- Scheffe, H. *The Analysis of Variance*, volume 72. John Wiley & Sons (1999).
- Sen, A., Banerjee, M., Li, Y. and Noone, A. M. A Bayesian approach to competing risks analysis with masked cause of death. *Statistics in medicine* **29(16)**: 1681–1695 (2010).
- Seo, J. I. Objective Bayesian analysis for the Weibull distribution with partial information under the generalized type-II progressive hybrid censoring scheme. *Communications in Statistics-Simulation and Computation* pp. 1–17 (2020).
- Seo, J. I. and Kim, Y. Objective Bayesian analysis based on upper record values from two-parameter Rayleigh distribution with partial information. *Journal of Applied Statistics* **44(12)**: 2222–2237 (2017).
- Sheikh, M. T., Ibrahim, J. G., Gelfond, J. A., Sun, W. and Chen, M.-H. Joint modelling of longitudinal and survival data in the presence of competing risks with applications to prostate cancer data. *Statistical Modelling* **21(1-2)**: 72–94 (2021).
- Shen, Y. and Xu, A. On the dependent competing risks using Marshall–Olkin bivariate Weibull model: Parameter estimation with different methods. *Communications in Statistics-Theory and Methods* **47(22)**: 5558–5572 (2018).
- Sinha, S. K. *Reliability and Life Testing*. Wiley Eastern Ltd, New Delhi (1986).
- Sullivan, L. M. Repeated measures. *Circulation* **117(9)**: 1238–1243 (2008).
- Sun, D. and Berger, J. O. Reference priors with partial information. *Biometrika* **85(1)**: 55–71 (1998).

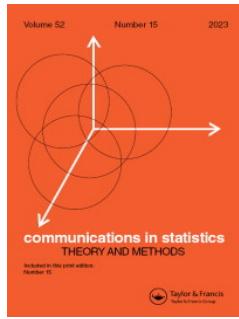
- Sun, D. and Ye, K. Reference prior Bayesian analysis for normal mean products. *Journal of the American Statistical Association* **90**(430): 589–597 (1995).
- Sun, J. *The statistical analysis of interval-censored failure time data*, volume 3. Springer (2006).
- Sun, J. and Kalbfleisch, J. D. The 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).
- Tiao, G. C. and Box, G. E. Some comments on “Bayes” estimators. *The American Statistician* **27**(1): 12–14 (1973).
- 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).
- Tsiatis, A. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences* **72**(1): 20–22 (1975).
- Upadhyay, S. and Smith, A. A Bayesian approach to model comparison in reliability via predictive simulation. Technical report, Dept. of Mathematics, Imperial College, London (1993).
- Upadhyay, S., Vasishta, N. and Smith, A. Bayes inference in life testing and reliability via Markov chain Monte Carlo simulation. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* **63**(1): 15–40 (2001).
- Upadhyay, S. K. and Smith, A. F. Modelling complexities in reliability, and the role of simulation in Bayesian computation. *International Journal of Continuing Engineering Education and Life Long Learning* **4**(1-2): 93–104 (1994).

- 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).
- Wada, C. Y., Sen, P. K. and Shimakura, S. E. A bivariate exponential model with covariates in competing risk data. *Calcutta Statistical Association Bulletin* **46**(3-4): 197–210 (1996).
- Wang, C. P. and Ghosh, M. Bayesian analysis of bivariate competing risks models with covariates. *Journal of Statistical Planning and Inference* **115**(2): 441–459 (2003).
- Wang, L. Estimation for exponential distribution based on competing risk middle censored data. *Communications in Statistics-Theory and Methods* **45**(8): 2378–2391 (2016).
- Xu, A. and Tang, Y. Bayesian analysis of Birnbaum-Saunders distribution with partial information. *Computational Statistics & Data Analysis* **55**(7): 2324–2333 (2011).
- Xu, A. and Zhou, S. Bayesian analysis of series system with dependent causes of failure. *Statistical Theory and Related Fields* **1**(1): 128–140 (2017).
- Yadav, C. P., Tomer, S. K. and Panwar, M. S. A competing risk study of menarcheal age distribution based on non-recall current status data. *Journal of Applied Statistics* **0**(0): 1–23 (2022).
- Yates, F. Complex experiments. *Supplement to the Journal of the Royal Statistical Society* **2**(2): 181–247 (1935).
- Zeger, S. L., Liang, K.-Y. and Albert, P. S. Models for longitudinal data: A generalized estimating equation approach. *Biometrics* pp. 1049–1060 (1988).
- Zhang, F., Chen, M.-H., Cong, X. J. and Chen, Q. Assessing importance of biomarkers: A Bayesian joint modelling approach of longitudinal and survival data with semi-competing risks. *Statistical Modelling* **21**(1-2): 30–55 (2021).
- Zhang, Z., Charalambous, C. and Foster, P. Joint modelling of longitudinal measurements and survival times via a multivariate copula approach. *Journal of Applied Statistics* pp. 1–21 (2022).

List of Papers

Paper Published

1. M. S. Panwar and Vikas Barnwal, **Objective Bayesian analysis of Marshall-Olkin bivariate Weibull distribution with partial information**, *Communications in Statistics - Theory and Methods*, 2023.
2. Vikas Barnwal and M. S. Panwar, **Competing risks analysis for dependent causes using Marshall-Olkin bivariate generalized lifetime family**, *Communications in Statistics - Theory and Methods*, 2022.



Objective Bayesian analysis of Marshall-Olkin bivariate Weibull distribution with partial information

M. S. Panwar & Vikas Barnwal

To cite this article: M. S. Panwar & Vikas Barnwal (2023): Objective Bayesian analysis of Marshall-Olkin bivariate Weibull distribution with partial information, *Communications in Statistics - Theory and Methods*, DOI: [10.1080/03610926.2023.2219418](https://doi.org/10.1080/03610926.2023.2219418)

To link to this article: <https://doi.org/10.1080/03610926.2023.2219418>



Published online: 06 Jun 2023.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Objective Bayesian analysis of Marshall-Olkin bivariate Weibull distribution with partial information

M. S. Panwar and Vikas Barnwal

Department of Statistics, Institute of Science, Banaras Hindu University, Varanasi, Uttar Pradesh, India

ABSTRACT

In competing risks problem, a subset of risks is needed more attention for inferential purposes. In the objective Bayesian paradigm, reference priors enable to achieve such inferential objectives. In this article, the Marshall-Olkin bivariate Weibull distribution is considered to model the competing risks data. In the availability of partial information for some of the parameters, the reference priors are derived as per the importance of the parameters. The Dirichlet prior is taken as a conditional subjective prior and the marginal reference prior has been derived. Also, the propriety of the resulting posterior density has been proved. The Bayesian estimates of the parameters are obtained under squared error and linear-exponential loss functions. Further, the derived reference prior is used for the computation of Bayes factors or posterior odds in testing the hypothesis that the competing risks are identical. The performance of established Bayesian estimators is illustrated using the Diabetic Retinopathy Study (DRS) and Prostate Cancer data sets. Finally, the model compatibility is done for the considered data sets under Bayesian Paradigm.

ARTICLE HISTORY

Received 1 September 2022
Accepted 19 May 2023

KEYWORDS

Objective Bayesian; partial information; reference prior; Bayes factor; posterior predictive p-value

1. Introduction

In life-testing experiments, the main concern lies in the unknown parameters of distribution assumed for the variable under study; for example, in testing the light bulbs, researchers are interested in the mean life length, rate of failure of the bulbs, or shape of the distribution, and so on. Hence, the unknown parameters of the model are always keen of interest. Thus, it becomes inevitable to perform statistical inferences to elicit more information about these parameters. The inferential procedure can be carried out using two major phenomena in the statistical literature- classical inference and Bayesian inference. The latter phenomenon considers the unknown parameter as a random variable, unlike in classical inference, which considers it as an unknown but fixed quantity. In the Bayesian approach, prior beliefs about the unknown parameter can be formulated with a specific probability distribution whenever some prior information is available. After observing the data, the prior distribution is nested with the likelihood to give the posterior distribution of the parameter. Hence, the prior belief about the parameter is updated in light of observed data. Moreover, this posterior information can be utilized as prior information for future data. Based on the prior information available

with the unknown parameters, the Bayesian analysis is done either with informative priors or non informative priors. Whenever the prior information about the parameter of interest is sufficiently available, the Bayesian analysis is preferred under informative priors such as conjugate priors. However, when the information available is vague or inadequate, non informative priors are used. For a detailed discussion on the selection of priors under the objective Bayesian paradigm, one can see Ghosh and Mukerjee (1992), Kass and Wasserman (1996), and Ghosh (2011).

In addition, reference priors are one of the non informative priors, which maximizes the divergence between prior and posterior density asymptotically. Reference prior works in such a way that the impact of the prior diffuses, and posterior inference becomes data-driven. Bernardo (1979) derived the reference priors and later Berger and Bernardo (1992) extended it giving an algorithm for the derivation of reference priors for the parameters of a number of distributions. Apart from all the above, sometimes a situation arises where adequate information is available for some parameters and the information is missing or vague for other parameters; for example, for a $N(\mu, \sigma^2)$ population, it is possible that one has prior information about the unknown parameter σ^2 , but there is no information about the mean parameter μ . In such a case of partial prior information available with some of the parameters, the marginal or conditional reference prior becomes indispensable to be derived for Bayesian analysis as suggested by Sun and Berger (1998).

In time-to-event analysis, the event of interest may be the death of an individual, relapse of a disease after remission, failure of a system, etc. It is often the case that the event of interest may occur with multiple endpoints. This phenomenon is known as competing risks in the statistical literature. While many authors considered competing risks to be independent, resulting in mathematical treatment much easier. However, in most situations, the independence assumption is quite susceptible. One of the available methods to model the dependency among competing risks is a multivariate probability distribution. Marshall and Olkin (1967) introduced one of the multivariate probability distribution models when there is a non zero probability of simultaneous occurrence of two causes of failure. Shen and Xu (2018) analyzed the competing risks data for dependent causes considering Marshall-Olkin bivariate Weibull distribution under three different methods. While Zhang et al. (2022) obtained the system reliability under the Bayesian paradigm for a multicomponent stress-strength model assuming Marshall-Olkin bivariate Weibull distribution. Recently, Barnwal and Panwar (2022) considered the Marshall-Olkin bivariate generalized lifetime family of distribution to model the dependency between competing risks and performed the classical and Bayesian inference with reference priors derived for different ordered grouping of the parameters of interest. Also, Kundu and Gupta (2013) performed the Bayes estimation of the parameters of Marshall-Olkin bivariate Weibull distribution considering the Gamma prior for scale parameters. Objective Bayesian analysis with partial information has been done by several authors for parameters of univariate distributions in different scenarios. Xu and Tang (2011) performed Bayesian analysis for Birnbaum-Saunders's distribution with partial information for a type-II censored sample and proved the propriety of the obtained posterior. Further, Seo and Kim (2017) analyzed the upper record value data under the Bayesian paradigm using partial information for unknown parameters of Rayleigh Distribution. Moreover, Seo (2020) performed the same for Weibull distribution with partial information for generalized type-II progressive hybrid censored data.

The novelty of the work is that objective Bayesian inference is performed using reference prior with partial information for competing risks data using Marshall-Olkin bivariate Weibull distribution (MOBWD). The purpose of this article is twofold: first, we are concerned with the changes in the behavior of Bayesian estimators when information is available on some of the parameters, that is, in case of partial prior information; and second to perform the testing of the hypothesis under the Bayesian paradigm using Bayes factor for testing whether the two risks are identical or not. In [Section 2](#), we have introduced the well-known Marshall-Olkin bivariate Weibull distribution along with its joint survival and joint density function. The likelihood formulation in the classical framework for the unknown parameters has been given in [Section 3](#). In [Section 4](#), the reference priors for unknown parameters of MOBWD have been derived using partial prior information available with parameters and propriety of the posterior has been proved. The simulation study has been performed to illustrate the performance of estimators derived in [Section 5](#). Finally, two real data sets have been analyzed in [Section 6](#).

2. Preliminaries

Let a lifetime random variable X follow the Weibull distribution with scale parameter α and shape parameter β . Consider the probability density function (*pdf*) and survival function of $\mathcal{W}(\beta, \alpha)$ as follows

$$f(x; \beta, \alpha) = \alpha\beta x^{\beta-1} \exp(-\alpha x^\beta) \quad (2.1)$$

$$S(x; \beta, \alpha) = \exp(-\alpha x^\beta); \quad x, \beta, \alpha > 0. \quad (2.2)$$

Assume three independent random variables U , V , and W each following Weibull distribution with the same shape parameter β and scale parameters α_i ; $i = 0, 1, 2$. After defining random variables $X_1 = \min(U, W)$ and $X_2 = \min(V, W)$, the bivariate random vector (X_1, X_2) is said to follow the well-known Marshall-Olkin bivariate Weibull distribution.

If $(X_1, X_2) \sim MOBWD(\beta, \alpha_0, \alpha_1, \alpha_2)$, then the following results are easy to establish

- (i) $X_1 \sim \mathcal{W}(\beta, \alpha_0 + \alpha_1)$ and $X_2 \sim \mathcal{W}(\beta, \alpha_0 + \alpha_2)$;
- (ii) $P(X_1 = X_2) = \frac{\alpha_0}{\alpha}$, $P(X_1 > X_2) = \frac{\alpha_2}{\alpha}$, $P(X_1 < X_2) = \frac{\alpha_1}{\alpha}$ and
- (iii) $\min\{X_1, X_2\} \sim \mathcal{W}(\beta, \alpha)$, where $\alpha = \alpha_0 + \alpha_1 + \alpha_2$.

The joint survival function of (X_1, X_2) for MOBWD can also be defined such as (Kundu and Dey [2009](#))

$$S(x_1, x_2) = \begin{cases} S_1(x_1, x_2) = S(x_1; \beta, \alpha_0 + \alpha_1)S(x_2; \beta, \alpha_2) & \text{if } x_1 > x_2, \\ S_2(x_1, x_2) = S(x_2; \beta, \alpha_0 + \alpha_2)S(x_1; \beta, \alpha_1) & \text{if } x_2 > x_1, \\ S_0(x_1, x_2) = S(x; \beta, \alpha) & \text{if } x_1 = x_2 = x. \end{cases} \quad (2.3)$$

Similarly, the joint pdf for (X_1, X_2) for MOBWD is given by (Kundu and Dey [2009](#))

$$f(x_1, x_2) = \begin{cases} f_1(x_1, x_2) = f(x_1; \beta, \alpha_0 + \alpha_1)f(x_2; \beta, \alpha_2) & \text{if } x_1 > x_2, \\ f_2(x_1, x_2) = f(x_2; \beta, \alpha_0 + \alpha_2)f(x_1; \beta, \alpha_1) & \text{if } x_2 > x_1, \\ f_0(x_1, x_2) = \frac{\alpha_0}{\alpha}f(x; \beta, \alpha) & \text{if } x_1 = x_2 = x. \end{cases} \quad (2.4)$$

3. Likelihood construction for dependent competing risks model

In a life-testing experiment, assume that n independent series systems with two components are being tested. The notion of component independence is untenable in this case. Let $T_j =$

$\min\{X_{1j}, X_{2j}\}$ be failure time for the j^{th} series system; $j = 1, 2, \dots, n$; where X_{ij} ; ($i = 1, 2$) denotes the failure of j^{th} system due to i^{th} component and components are assumed to be dependent. The dependence arises from the simultaneous failure of both components. The observed data is of the form $\{T_j, \delta_{1j}, \delta_{2j}\}$; where δ_{ij} is an indicator function which takes the value 1 if j^{th} system's failure is caused by i^{th} component and value 0 if system's failure is caused by both the components. Hence, failure of the system may be caused by simultaneous failure of both components with a non zero probability. Now further assume that (X_{1j}, X_{2j}) follows Marshall-Olkin bivariate Weibull distribution, that is, $(X_{1j}, X_{2j}) \sim MOBWD(\beta, \alpha_0, \alpha_1, \alpha_2)$.

To carry out inference for the unknown parameters, it is always an interest to infer about the characteristics which are more important than the parameters themselves. In competing risks analysis, the researcher is more interested in the cause of failure, which is more severe than other causes. So, a particular set after re-parameterization has been considered such as, say,

$$\lambda_1 = \alpha = \alpha_0 + \alpha_1 + \alpha_2, \quad \lambda_2 = \frac{\alpha_1}{\alpha}, \quad \lambda_3 = \frac{\alpha_2}{\alpha}.$$

The re-parameterized quantities $\lambda_2 = P(X_1 < X_2)$ and $\lambda_3 = P(X_2 < X_1)$ represent the relative risk due to cause 1 and 2, respectively. The transformation $(\beta, \Lambda) = (\beta, \lambda_1, \lambda_2, \lambda_3)$ from $(\beta, \alpha_0, \alpha_1, \alpha_2)$ is one-to-one with the inverse transformation where

$$\beta = \beta, \quad \alpha_1 = \lambda_1 \lambda_2, \quad \alpha_2 = \lambda_1 \lambda_3, \quad \alpha_0 = \lambda_1(1 - \lambda_2 - \lambda_3).$$

Now the likelihood function under re-parametrization is given by

$$L(t; \beta, \Lambda) = \beta^n \lambda_1^{n_1} \lambda_2^{n_2} \lambda_3^{n_3} (1 - \lambda_2 - \lambda_3)^{n_0} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}, \quad (3.1)$$

where n_0, n_1 and n_2 are the numbers of systems which have failed due to the failure of both components, the first component and the second component, respectively. We can denote them such as:

$$\begin{aligned} \sum_{j=1}^n I(X_{1j} = X_{2j}) &= \sum_{j=1}^n \delta_{1j} \delta_{2j} = n_0, \\ \sum_{j=1}^n I(X_{1j} < X_{2j}) &= \sum_{j=1}^n \delta_{1j} (1 - \delta_{2j}) = n_1, \\ \sum_{j=1}^n I(X_{1j} > X_{2j}) &= \sum_{j=1}^n \delta_{2j} (1 - \delta_{1j}) = n_2. \end{aligned}$$

Here $n = n_0 + n_1 + n_2$, as it is assumed that no information is censored or missing in the experiment. The maximum likelihood estimators (MLEs) of the unknown parameters (β, Λ) can be obtained by taking the partial derivative of the natural logarithm of the likelihood function (3.1) with respect to (β, Λ) and equating them to zero. In a similar way, the Fisher information matrix, say Σ can be obtained as

$$\Sigma = -E \left[\frac{\partial^2 \log L(t; \beta, \Lambda)}{\partial (\beta, \Lambda)^2} \middle| (\beta, \Lambda) \right]$$

Hence, the Fisher information matrix has the following form:

$$\Sigma = n \begin{bmatrix} \frac{k(\lambda_1)}{\beta^2} & \frac{r_1+1-\log\lambda_1}{\beta\lambda_1} & 0 & 0 \\ \frac{r_1+1-\log\lambda_1}{\beta\lambda_1} & \frac{1}{\lambda_1^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}, \quad (3.2)$$

where $k(\lambda_1) = r_2 + 2r_1(1 - \log\lambda_1) - \log\lambda_1(2 - \log\lambda_1) + 1$ and $r_u = \int_0^\infty (\log z)^u \exp\{-z\} dz$; $u = 1, 2$. The present form of the likelihood function (3.1) is convenient for the posterior analysis. Hence, we utilize the likelihood function in the next section to establish the posterior density and the Bayesian estimates of the unknown parameters (β, Λ) of the dependent competing risks model will be obtained accordingly.

4. Bayesian estimation

The Bayesian inference for unknown parameters has become inevitable in life-testing experiments. To update the information about parameters in light of the observed data, the prior information is nested with the likelihood function. If one has only partial prior information on some of the parameters and the information available to other parameters is vague, then the Bayesian estimation is carried out with partial prior information available as suggested by Sun and Berger (1998). In case when the available information about the parameter is vague, one of the most popular priors is the reference prior proposed by Berger and Bernardo (1992). The idea behind reference prior is that it maximizes the Kullback-Leibler divergence between the prior and posterior density asymptotically. Consequently, the posterior inference becomes data-dominated and the effect of the prior is diffused. We derive reference prior for the unknown parameter of interest using the partial prior information available with the other parameter as proposed by Sun and Berger (1998).

Suppose one is interested in the ordered group (θ_1, θ_2) of parameters, where parameters θ_1 and θ_2 are of dimensions d_1 and d_2 , respectively.

- (i) Given the marginal subjective prior $\pi^s(\theta_2)$, the conditional reference prior for θ_1 given θ_2 , say $\pi^r(\theta_1|\theta_2)$ can be derived.
- (ii) Given the subjective conditional prior $\pi^s(\theta_2|\theta_1)$, the marginal reference prior for θ_1 is given by Theorem 1 of Sun and Berger (1998) as follows

$$\pi^r(\theta_1) \propto \left(\frac{\det \Sigma}{\det \Sigma_{22}} \right)^{1/2}, \quad (4.1)$$

where Σ is the Fisher Information matrix for (θ_1, θ_2) and Σ_{22} is Fisher information matrix for θ_2 for fixed θ_1 .

4.1. Reference prior with partial information (RPPI)

In this subsection, we derive reference prior in the case when partial information is available to some of the unknown parameters. Also, we establish that the derived reference prior leads to the proper posterior density.

- (i) When partial prior information is available for parameter λ_1 only

Assume that parameter λ_1 has a Gamma prior, say $\mathcal{G}(a, b)$ given by

$$\pi_1^s(\lambda_1 | a, b) \propto \lambda_1^{a-1} e^{-b\lambda_1}; \quad \lambda_1 > 0; \quad a, b > 0. \quad (4.2)$$

Here a and b are the hyper-parameters specifying the prior distribution of λ_1 . The choice of hyper-parameters is very important to carry out the Bayesian inference.

Now to obtain the joint prior for (β, Λ) , rearranging the rows and columns of matrix Σ given in (3.2), we get

$$\Sigma_1 = n \begin{bmatrix} \frac{1}{\lambda_1^2} & \frac{r_1+1-\log\lambda_1}{\beta\lambda_1} & 0 & 0 \\ \frac{r_1+1-\log\lambda_1}{\beta\lambda_1} & \frac{k(\lambda_1)}{\beta^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}. \quad (4.3)$$

The determinant of matrix Σ_1 in (4.3) is

$$\det \Sigma_1 = \frac{n}{\beta^2} \frac{k(\lambda_1)}{\lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)},$$

which can be written as

$$\det \Sigma_1 \propto g_1(\theta_1).g_2(\theta_2), \quad (4.4)$$

where

$$g_1(\theta_1) = k(\lambda_1) \text{ and } g_2(\theta_2) = \frac{1}{\beta^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)}.$$

Then according to Theorem 2 given in Sun and Berger (1998), the conditional reference prior for $\theta_2 = (\beta, \lambda_2, \lambda_3)$ given $\theta_1 = \lambda_1$ is

$$\begin{aligned} \pi_1^r(\theta_2 | \theta_1) &\propto \sqrt{g_2(\theta_2)} \\ &\propto \beta^{-1} (\lambda_2 \lambda_3)^{-1/2} (1 - \lambda_2 - \lambda_3)^{-1/2}. \end{aligned} \quad (4.5)$$

The joint prior for (β, Λ) is given by

$$\begin{aligned} \xi_1(\beta, \Lambda) &\propto \pi_1^r(\beta, \lambda_2, \lambda_1 | \lambda_3). \pi_1^s(\lambda_1 | a, b) \\ &\propto \beta^{-1} \lambda_1^{a-1} e^{-b\lambda_1} (\lambda_2 \lambda_3)^{-1/2} (1 - \lambda_2 - \lambda_3)^{-1/2}. \end{aligned} \quad (4.6)$$

In respect to the prior (4.6), the posterior density will be as follows:

$$\begin{aligned} p_1(\beta, \Lambda | t) &\propto \xi_1(\beta, \Lambda). L(t; \beta, \Lambda) \\ &\propto \beta^{n-1} \lambda_1^{n+a-1} \lambda_2^{(n_1-\frac{1}{2})} \lambda_3^{(n_2-\frac{1}{2})} (1 - \lambda_2 - \lambda_3)^{(n_0-\frac{1}{2})} \\ &\cdot \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\}. \end{aligned} \quad (4.7)$$

(ii) *When partial prior information is available for parameters (λ_2, λ_3) only*

Given (β, λ_1) , the conditional subjective prior for (λ_2, λ_3) has a Dirichlet prior, say $Dir(a_0, a_1, a_2)$ given by

$$\pi_2^s(\lambda_2, \lambda_3 | \lambda_1, a_0, a_1, a_2) \propto \lambda_2^{a_1-1} \lambda_3^{a_2-1} (1 - \lambda_2 - \lambda_3)^{a_0-1}; \quad a_i > 0; \forall i = 0, 1, 2. \quad (4.8)$$

The partition of matrix Σ is given by

$$\Sigma_{11} = n \begin{bmatrix} \frac{k(\lambda_1)}{\beta^2} & \frac{r_1+1-\log\lambda_1}{\beta\lambda_1} \\ \frac{r_1+1-\log\lambda_1}{\beta\lambda_1} & \frac{1}{\lambda_1^2} \end{bmatrix}, \quad (4.9)$$

and

$$\Sigma_{22} = n \begin{bmatrix} \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix} \quad (4.10)$$

The determinant of Σ_{11} is

$$\det \Sigma_{11} = \frac{n(r_2 - r_1^2)}{\beta^2 \lambda_1^2}.$$

The determinant of Σ_{22} is

$$\det \Sigma_{22} = \frac{n}{\lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)}.$$

The determinant of Σ is

$$\det \Sigma = \det \Sigma_{11} \cdot \det \Sigma_{22}$$

Since the quantity $\frac{\det \Sigma}{\det \Sigma_{22}}$ does not depend on (λ_2, λ_3) , the marginal reference prior for (β, λ_1) is given by using (4.1)

$$\pi_2^r(\beta, \lambda_1) \propto \frac{1}{\beta \lambda_1}. \quad (4.11)$$

The joint prior for (β, Λ) is given by

$$\begin{aligned} \xi_2(\beta, \Lambda) &\propto \pi_2^r(\beta, \lambda_1) \cdot \pi_2^s(\lambda_2, \lambda_3 | \lambda_1, a_0, a_1, a_2) \\ &\propto (\beta \lambda_1)^{-1} \lambda_2^{a_1-1} \lambda_3^{a_2-1} (1 - \lambda_2 - \lambda_3)^{a_0-1}. \end{aligned} \quad (4.12)$$

In respect to the prior (4.12), the posterior density will be as follows:

$$\begin{aligned} p_2(\beta, \Lambda | \underline{t}) &\propto \xi_2(\beta, \Lambda) \cdot L(\underline{t}; \beta, \Lambda) \\ &\propto \beta^{n-1} \lambda_1^{n-1} \lambda_2^{(n_1+a_1-1)} \lambda_3^{(n_2+a_2-1)} (1 - \lambda_2 - \lambda_3)^{(n_0+a_0-1)} \\ &\cdot \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}. \end{aligned} \quad (4.13)$$

4.2. Posterior propriety

In this subsection, the propriety of posterior densities $p_1(\beta, \Lambda | \underline{t})$ and $p_2(\beta, \Lambda | \underline{t})$ obtained under the joint priors $\xi_1(\beta, \Lambda)$ and $\xi_2(\beta, \Lambda)$, respectively, has been established.

Theorem 4.1. *The posterior density $p_1(\beta, \Lambda | \underline{t})$ under the joint prior $\xi_1(\beta, \Lambda)$ is proper for $n_0 > 0, n_1 > 0, n_2 > 0$ with at least two failures.*

Proof. The posterior density $p_1(\beta, \Lambda | \underline{t})$ is given by (4.7). To prove the posterior density proper, we have to show that it gives a finite value after integration, that is, $\int p_1(\beta, \Lambda | \underline{t}) d\beta d\Lambda < \infty$.

We have to calculate the following integral

$$I^1 = \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1 - \frac{1}{2}} \lambda_3^{n_2 - \frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0 - \frac{1}{2}} d\lambda_2 d\lambda_3 \right\} \right. \\ \left. \cdot \lambda_1^{n+a-1} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta.$$

If $\mathcal{B}(.,.)$ denotes the beta function then we have

$$\mathcal{B}(a+1, b+c+2) \cdot \mathcal{B}(b+1, c+1) = \int_{0 < x+y < 1} x^a y^b (1-x-y)^c dx dy.$$

Now using this result, the above integral can be written such as

$$I^1 = A_1 \int_0^\infty \left[\int_0^\infty \lambda_1^{n+a-1} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ = C_1 \int_0^\infty \beta^{n-1} \frac{\prod_{j=1}^n t_j^\beta}{\left(b + \sum_{j=1}^n t_j^\beta \right)^{n+a}} d\beta \leq C_1 \int_0^\infty \beta^{n-1} \prod_{i=1}^n \left(\frac{t_i}{b + t_{(n)}} \right)^\beta d\beta < \infty,$$

where $t_{(n)} = \max_{1 \leq i \leq n} t_i$, $A_1 = \mathcal{B}(n_1 + \frac{1}{2}, n_0 + n_2 + 1) \cdot \mathcal{B}(n_2 + \frac{1}{2}, n_0 + \frac{1}{2})$ and $C_1 = A_1 \Gamma(n+a) \left(\prod_{j=1}^n t_j \right)^{-1}$. The last inequality holds because $\prod_{i=1}^n \frac{t_i}{b+t_{(n)}} < 1$. Hence the posterior density $p_1(\beta, \Lambda | \underline{t})$ is proper. \square

Theorem 4.2. *The posterior density $p_2(\beta, \Lambda | \underline{t})$ under the joint prior $\xi_2(\beta, \Lambda)$ is proper for $n_0 > 0, n_1 > 0, n_2 > 0$ with at least two failures.*

Proof. The posterior density $p_2(\beta, \Lambda | \underline{t})$ is given by (4.13). To prove the posterior density proper, we have to show that it gives a finite value after integration, that is, $\int p_2(\beta, \Lambda | \underline{t}) d\beta d\Lambda < \infty$.

We have to calculate the following integral

$$I^* = \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1+a_1-1} \lambda_3^{n_2+a_2-1} (1 - \lambda_2 - \lambda_3)^{n_0+a_0-1} d\lambda_2 d\lambda_3 \right\} \right. \\ \left. \cdot \lambda_1^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta.$$

If $\mathcal{B}(.,.)$ denotes the beta function then we have

$$\mathcal{B}(a+1, b+c+2) \cdot \mathcal{B}(b+1, c+1) = \int_{0 < x+y < 1} x^a y^b (1-x-y)^c dx dy.$$

Now using this result, the above integral can be written such as

$$\begin{aligned} I &= A_2 \int_0^\infty \left[\int_0^\infty \lambda_1^{n-1} e^{-\lambda_1 \sum_{i=1}^n t_j^\beta} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ &= C_2 \int_0^\infty \beta^{n-1} \frac{\prod_{j=1}^n t_j^\beta}{\left(\sum_{i=1}^n t_j^\beta \right)^n} d\beta \leq C_2 \int_0^\infty \beta^{n-1} \prod_{i=1}^n \left(\frac{t_j}{t_{(n)}} \right)^\beta d\beta < \infty, \end{aligned}$$

where $t_{(n)} = \max_{1 \leq i \leq n} t_i$, $A_2 = \mathcal{B}(n_1 + a_1, n_0 + n_2 + a_0 + a_2) \cdot \mathcal{B}(n_2 + a_2, n_0 + a_0)$ and $C_2 = A_2 \Gamma(n) \left(\prod_{j=1}^n t_j \right)^{-1}$. The last inequality holds because $\prod_{i=1}^n \frac{t_j}{t_{(n)}} < 1$. Hence the posterior density $p_2(\beta, \Lambda | t)$ is proper. \square

4.3. Posterior estimation

The posterior analysis of the unknown parameters has been done using conditional distributions. Since the obtained distributions are not in standard form, we use the Gibbs sampling method, a Markov chain Monte Carlo (MCMC) technique (Chen, Shao, and Ibrahim 2000).

The full conditionals of the parameters can be obtained using (4.7) as follows:

$$p_3(\beta | \lambda_1, \text{data}) \propto \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\}, \quad (4.14)$$

$$p_4(\lambda_1 | \beta, \text{data}) \propto \lambda_1^{n+a-1} \exp \left\{ -\lambda_1 \left(b + \sum_{j=1}^n t_j^\beta \right) \right\}, \quad (4.15)$$

$$p_5(\lambda_2, \lambda_3 | \text{data}) \propto \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}}. \quad (4.16)$$

Also, the full conditionals of the parameters using (4.13) are:

$$p_6(\beta | \lambda_1, \text{data}) \propto \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}, \quad (4.17)$$

$$p_7(\lambda_1 | \beta, \text{data}) \propto \lambda_1^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\}, \quad (4.18)$$

$$p_8(\lambda_2, \lambda_3 | \text{data}) \propto \lambda_2^{n_1+a_1-1} \lambda_3^{n_2+a_2-1} (1 - \lambda_2 - \lambda_3)^{n_0+a_0-1}. \quad (4.19)$$

To obtain the posterior samples on the parameters from the posteriors given in (4.7) and (4.13), we follow the procedure given below:

- The posterior sample for β from (4.14) and (4.16) can be obtained using Metropolis-Hastings (M-H) algorithm (Metropolis and Ulam 1949; Hastings 1970) by considering normal distribution as a proposal density such as the candidate point $\beta^c \sim \mathcal{N}(\hat{\beta}, \text{Var}(\hat{\beta}))$.
- From (4.17) and (4.19), it is clear that a sample can be generated directly for λ_1 from $\mathcal{G}\left(n, \sum_{j=1}^n t_j^\beta\right)$ and $\mathcal{G}\left(n+a, (b + \sum_{j=1}^n t_j^\beta)\right)$. The values of hyperparameters a and b are chosen arbitrarily. Here it is to be noted that (β, λ_1) is a vector of parameters for which

MCMC is induced by the Gibbs sampler. First, we choose an arbitrary starting point (β^0, λ_1^0) and then proceed accordingly. So, as we collect the sample from both it is easy to obtain the Bayes estimates.

- The joint posterior density follows Dirichlet distribution, that is, $(\lambda_2, \lambda_3) \sim Dir(n_1 + a_1, n_2 + a_2, n_0 + a_0)$. Now if a $rv Y_i \sim \mathcal{G}(a_i, 1); i = 1, 2, \dots, K$, where $\mathcal{G}(\cdot, \cdot)$ stands for gamma density and assume that $X_i = Y_i / \sum_{i=1}^K Y_i$ then $(X_1, X_2, \dots, X_{K-1}) \sim Dir(a_1, a_2, \dots, a_K)$. Then it is obvious to write that $\lambda_2 = \frac{y_1}{y_1 + y_2 + y_3}$; $\lambda_3 = \frac{y_2}{y_1 + y_2 + y_3}$ where $y_1 \sim \mathcal{G}(n_1 + a_1, 1)$, $y_2 \sim \mathcal{G}(n_2 + a_2, 1)$ and $y_3 \sim \mathcal{G}(n_0 + a_0, 1)$ for suitable choice of hyperparameters (a_0, a_1, a_2) . It is quite easy to generate random numbers from y_1, y_2 , and y_3 and so for λ_2 and λ_3 .

After removing the burn-in from the generated chains, the stationarity and convergence can be checked using the cumsum plot, ACF, and Gelman and Rubin's test statistics (Gelman and Rubin 1992). Along with the selection of priors, the choice of the loss function is also important to draw the required inference related to unknown parameters of interest. The Bayesian estimator of unknown parameters under the commonly used squared error loss function (SELF) is the mean of the posterior distribution. Suppose one wants to obtain the Bayesian estimator of some function of (β, Λ) , say, $g(\beta, \Lambda)$ under SELF, then SELF is defined as

$$L_{SE}(\hat{g}(\beta, \Lambda), g(\beta, \Lambda)) = (\hat{g}(\beta, \Lambda) - g(\beta, \Lambda))^2,$$

where $\hat{g}(\beta, \Lambda)$ is an estimator of $g(\beta, \Lambda)$.

The Bayesian estimator of $g(\beta, \Lambda)$ under SELF is given by

$$\hat{g}(\beta, \Lambda)^{SE} = \frac{\int g(\beta, \Lambda)p(\beta, \Lambda | t) d\beta d\Lambda}{\int p(\beta, \Lambda | t) d\beta d\Lambda},$$

where $p(\beta, \Lambda | t)$ is the posterior density of (β, Λ) .

Since SELF is a symmetric loss function, it gives equal weight to underestimation as well as overestimation. But, in real-life scenarios, underestimation of a parameter may lead to more or less severe consequences than overestimation and vice versa. Hence, using SELF in this case is not appropriate. So, we use a useful asymmetric loss function which is the linear-exponential (LINEX) loss function to obtain Bayesian estimators. The LINEX loss function is defined as

$$L_L(\hat{g}(\beta, \Lambda), g(\beta, \Lambda)) = e^{\gamma(\hat{g}(\beta, \Lambda) - g(\beta, \Lambda))} - \gamma(\hat{g}(\beta, \Lambda) - g(\beta, \Lambda)) - 1, \quad \gamma \neq 0.$$

Under LINEX, the Bayesian estimator of $g(\beta, \Lambda)$ is given by

$$\hat{g}(\beta, \Lambda)^L = -\frac{1}{\gamma} \log \left[E(e^{-\gamma g(\beta, \Lambda)} | t) \right] = -\frac{1}{\gamma} \log \left(\int e^{-\gamma g(\beta, \Lambda)} p(\beta, \Lambda | t) d\beta d\Lambda \right).$$

Provided that $E(e^{-\gamma g(\beta, \Lambda)} | t)$ exists and is finite. The value of γ regulates the shape of the LINEX loss function. The positive (negative) value of γ gives more weight to underestimation (overestimation). If the value of $|\gamma|$ is very small, then the LINEX loss function is approx symmetric and hence close to SELF.

4.4. Testing under Bayesian paradigm

In the Bayesian framework, testing of hypotheses can be carried out using posterior odds. Consider the dependent model established in Section 3. Assume that one wants to test the

hypothesis $H_0 : \lambda_2 = \lambda_3$ against $H_1 : \lambda_2 \neq \lambda_3$. For doing this, the posterior odds are given by

$$\text{posterior odds} = \frac{p(\beta, \Delta | \underline{t}, H_1)}{p(\beta^*, \alpha^* | \underline{t}, H_1)}$$

which further can be written as

$$\text{posterior odds} = \frac{p(\underline{t} | H_1)}{p(\underline{t} | H_0)} \frac{\pi(H_1)}{\pi(H_0)} \quad (4.20)$$

The first term on the right-hand side of (4.20) is known as the Bayes Factor (BF). We use the BF to test the hypothesis stated earlier. Since under H_0 , the two competing risks are identical, hence suppose that the observed data (t_1, t_2, \dots, t_n) is coming from a Weibull distribution with shape and scale parameters β^* and α^* , respectively. It should be noted that the shape parameter for the Weibull distribution is the same under null and alternative hypotheses, that is, $\beta = \beta^*$. The likelihood function for the unknown parameters β^* and α^* given the data is

$$L(\beta^*, \alpha^* | \text{data}) = \beta^{*n} \alpha^{*n} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \exp \left(-\alpha^* \sum_{j=1}^n t_j^{\beta^*} \right). \quad (4.21)$$

The log-likelihood function is

$$\log L(\beta^*, \alpha^* | \text{data}) = n \log(\beta^*) + n \log(\alpha^*) + (\beta^* - 1) \sum_{j=1}^n t_j - \alpha^* \sum_{j=1}^n t_j^{\beta^*}. \quad (4.22)$$

In this case, a reference prior with partial prior information for the unknown parameters β^* and α^* is derived. From the log-likelihood Equation (4.22), the negative of second partial derivative for α^* is

$$-\frac{\partial^2}{\partial \alpha^*^2} \log L(\beta^*, \alpha^* | \text{data}) \propto \frac{1}{\alpha^{*2}}.$$

According to Sun and Berger (1998), the conditional reference prior for α^* is

$$\pi(\alpha^* | \beta^*) \propto \frac{1}{\alpha^*}.$$

A non informative prior for β^* can be taken as

$$\pi(\beta^*) \propto \frac{1}{\beta^*}; \quad \beta^* > 0.$$

Hence the joint prior for the unknown parameters β^* and α^* is given by

$$\pi(\beta^*, \alpha^*) \propto \frac{1}{\alpha^* \beta^{*\nu}}; \quad -\infty < \nu < \infty. \quad (4.23)$$

In respect to prior given in (4.23), the posterior density will be as follows:

$$\begin{aligned} p(\beta^*, \alpha^* | \underline{t}) &\propto L(\beta^*, \alpha^* | \text{data}).\pi(\beta^*, \alpha^*) \\ &\propto \beta^{*n-\nu} \alpha^{*n-1} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \exp \left(-\alpha^* \sum_{j=1}^n t_j^{\beta^*} \right). \end{aligned} \quad (4.24)$$

The marginal distribution of t under H_0 is given by

$$\begin{aligned} p(\underline{t} \mid H_0) &= \int_0^\infty \int_0^\infty p(\beta^*, \alpha^* \mid \underline{t}) d\alpha^* d\beta^* \\ &= \int_0^\infty \int_0^\infty \beta^{*n-\nu} \alpha^{*n-1} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \exp \left(-\alpha^* \sum_{j=1}^n t_j^{\beta^*} \right) d\alpha^* d\beta^* \\ &= \Gamma(n) I_1, \end{aligned} \quad (4.25)$$

where

$$I_1 = \int_0^\infty \beta^{*n-\nu} \left(\prod_{j=1}^n t_j^{\beta^*-1} \right) \left(\sum_{j=1}^n t_j^{\beta^*} \right)^{-n} d\beta^*.$$

Now, we calculate the BF using posterior densities $p_1(\beta, \Lambda \mid \underline{t})$ and $p_2(\beta, \Lambda \mid \underline{t})$ given by (4.7) and (4.13). So, using $p_1(\beta, \Lambda \mid \underline{t})$ the marginal distribution of t under H_1 is given by

$$\begin{aligned} p_1(\underline{t} \mid H_1) &= \int p_1(\beta, \Lambda \mid \underline{t}) d\beta d\Lambda \\ &= \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1 - \frac{1}{2}} \lambda_3^{n_2 - \frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0 - \frac{1}{2}} d\lambda_2 d\lambda_3 \right\} \right. \\ &\quad \left. \cdot \lambda_1^{n+a-1} \exp \left\{ -\lambda_1(b + \sum_{j=1}^n t_j^\beta) \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ &= \mathcal{B}(n_1 + \frac{1}{2}, n_0 + n_2 + 1) \mathcal{B}(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}) \Gamma(n+a) I_2, \end{aligned} \quad (4.26)$$

where

$$I_2 = \int_0^\infty \beta^{n-1} \left(\prod_{j=1}^n t_j^{\beta-1} \right) \left(b + \sum_{j=1}^n t_j^\beta \right)^{-(n+a)} d\beta.$$

Hence, using (4.25) and (4.26), the Bayes Factor is given by

$$BF_1 = \frac{p_1(\underline{t} \mid H_1)}{p_1(\underline{t} \mid H_0)} = \mathcal{B}(n_1 + \frac{1}{2}, n_0 + n_2 + 1) \mathcal{B}(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}) \frac{\Gamma(n+a) I_2}{\Gamma(n) I_1}. \quad (4.27)$$

Similarly, using $p_2(\beta, \Lambda \mid \underline{t})$, the marginal distribution of t under H_1 is given by

$$\begin{aligned} p_2(\underline{t} \mid H_1) &= \int p_2(\beta, \Lambda \mid \underline{t}) d\beta d\Lambda \\ &= \int_0^\infty \left[\int_0^\infty \left\{ \iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1+a_1-1} \lambda_3^{n_2+a_2-1} (1 - \lambda_2 - \lambda_3)^{n_0+a_0-1} d\lambda_2 d\lambda_3 \right\} \right. \\ &\quad \left. \cdot \lambda_1^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n t_j^\beta \right\} d\lambda_1 \right] \beta^{n-1} \left\{ \prod_{j=1}^n t_j^{\beta-1} \right\} d\beta \\ &= \mathcal{B}(n_1 + a_1, n_0 + n_2 + a_0 + a_2) \mathcal{B}(n_2 + a_2, n_0 + a_0) \Gamma(n) I_3, \end{aligned} \quad (4.28)$$

where

$$I_3 = \int_0^\infty \beta^{n-1} \left(\prod_{j=1}^n t_j^{\beta-1} \right) \left(\sum_{j=1}^n t_j^\beta \right)^{-n} d\beta.$$

Similarly, using (4.25) and (4.28), the Bayes Factor is given by

$$BF_2 = \frac{p_2(\underline{t} | H_1)}{p(\underline{t} | H_0)} = \mathcal{B}(n_1 + a_1, n_0 + n_2 + a_0 + a_2) \cdot \mathcal{B}(n_2 + a_2, n_0 + a_0) \frac{I_3}{I_1}. \quad (4.29)$$

The decision rule for Bayesian testing is that one would reject H_0 in favor of H_1 if the value of the Bayes Factor is very high. Hence in the context of Bayesian testing, if the null hypothesis would be rejected, then the two dependent competing risks would not be considered identical.

4.5. Posterior predictive compatibility of model

It is indispensable to perform the compatibility of the model (i.e., likelihood and prior) established in the previous section before applying it to real data. Model compatibility is a necessary step to show that the established model is rightly used for the data. The initial and simplest approach considers the predictive capability of the model with regard to some of its important characteristics such as hazard rate, survival rate, and mean time to failure (MTTF) etc., using graphical tools. In this section, the goodness of fit test of the data in the Bayesian framework is performed through posterior probability p -value (PPV) given by Guttman (1967). To obtain the value of PPV , define the chi-square statistic which is a measure of discrepancy and is given by,

$$\chi^2 = \sum_{i=1}^n \frac{(t_i - E(T_i | \beta, \Lambda))^2}{V(T_i | \beta, \Lambda)}. \quad (4.30)$$

Hence, the PPV based on the discrepancy measure defined in (4.30), can be obtained as

$$PPV = \int P[\chi_2^2 > \chi_1^2 | \beta, \Lambda] p(\beta, \Lambda | \underline{t}) d\beta d\Lambda,$$

where χ_1^2 and χ_2^2 are the calculated values of χ^2 for the observed data and predictive data sets, respectively. The PPV is defined as the expected value of the classical p -value, where the expectation is taken with respect to posterior density. Since we have considered the case of a complete sample, so to calculate the value of PPV , the procedure described by Gelman, Meng, and Stern (2000) is followed. First of all, we draw posterior samples on (β, Λ) from $p(\beta, \Lambda | \underline{t})$ and then obtain the value of χ_1^2 for the given data set. After that, the predictive data sets are generated each of the same size as that of the observed data. The value of $P[\chi_2^2 > \chi_1^2 | \beta, \Lambda]$ is calculated as the number of times χ_2^2 exceeds χ_1^2 . The procedure defined above is replicated a number of times with extracted values of parameters and thus PPV is obtained as the posterior expectation of $P[\chi_2^2 > \chi_1^2 | \beta, \Lambda]$.

5. Simulation study

In this section, we have carried out a simulation study to illustrate the performance of derived estimators in Section 4. The simulation study has been performed to obtain the Bayesian

Table 1. ARB and MSE of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$ under RPPI-I.

n		SELF		LINEX($\gamma=1.5$)		LINEX($\gamma=1.5$)	
		ARB	MSE	ARB	MSE	ARB	MSE
20	β	0.1219	0.0577	0.1360	0.0726	0.1134	0.0489
	λ_1	0.1816	0.1244	0.2155	0.1890	0.1626	0.0922
	λ_2	0.2954	0.0079	0.3027	0.0083	0.2880	0.0076
40	λ_3	0.2560	0.0096	0.2565	0.0099	0.2556	0.0094
	β	0.0950	0.0360	0.1006	0.0410	0.0915	0.0326
	λ_1	0.1276	0.0585	0.1388	0.0715	0.1209	0.0504
60	λ_2	0.2138	0.0041	0.2160	0.0042	0.2116	0.0041
	λ_3	0.1805	0.0049	0.1813	0.0049	0.1797	0.0048
	β	0.0784	0.0240	0.0815	0.0263	0.0763	0.0224
100	λ_1	0.1034	0.0368	0.1092	0.0422	0.0992	0.0333
	λ_2	0.1786	0.0030	0.1795	0.0030	0.1776	0.0029
	λ_3	0.1510	0.0034	0.1516	0.0034	0.1504	0.0034
	β	0.0611	0.0144	0.0627	0.0153	0.0600	0.0137
	λ_1	0.0798	0.0219	0.0828	0.0238	0.0779	0.0207
	λ_2	0.1388	0.0018	0.1397	0.0018	0.1386	0.0018
	λ_3	0.1181	0.0021	0.1187	0.0021	0.1180	0.0021

Table 2. ARB and MSE of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$ under RPPI-II.

n		SELF		LINEX($\gamma=1.5$)		LINEX($\gamma=1.5$)	
		ARB	MSE	ARB	MSE	ARB	MSE
20	β	0.1233	0.0592	0.1380	0.0748	0.1142	0.0498
	λ_1	0.2019	0.1626	0.2432	0.2763	0.1805	0.1164
	λ_2	0.2664	0.0065	0.2724	0.0068	0.2604	0.0061
40	λ_3	0.2272	0.0077	0.2342	0.0081	0.2231	0.0075
	β	0.0953	0.0363	0.1010	0.0414	0.0917	0.0462
	λ_1	0.1337	0.0645	0.1450	0.0789	0.1270	0.0611
60	λ_2	0.2012	0.0037	0.2031	0.0038	0.1993	0.0036
	λ_3	0.1690	0.0043	0.1705	0.0044	0.1683	0.0044
	β	0.0785	0.0241	0.0817	0.0264	0.0764	0.0225
100	λ_1	0.1061	0.0390	0.1121	0.0446	0.1025	0.0354
	λ_2	0.1709	0.0027	0.1723	0.0027	0.1700	0.0026
	λ_3	0.1445	0.0031	0.1451	0.0032	0.1440	0.0031
	β	0.0611	0.0143	0.0627	0.0153	0.0600	0.0137
	λ_1	0.0812	0.0227	0.0840	0.0245	0.0795	0.0214
	λ_2	0.1358	0.0017	0.1369	0.0017	0.1348	0.0017
	λ_3	0.1155	0.0020	0.1163	0.0020	0.1148	0.0020

estimates of the unknown parameters under reference priors with partial prior information, namely, RPPI-I and RPPI-II derived in [Subsection 4.1](#). The performance of estimators is assessed by absolute relative bias (ARB), mean square error (MSE), and coverage probability (CP). We have taken sample sizes $n = 20, 40, 60$, and 100 to evaluate and compare the efficiency of estimators as the sample size increases. The simulation study is based on 5,000 samples to remove undesired sampling fluctuations and to achieve consistency in the outputs. The output of the simulation performed is shown in [Tables 1–6](#) for two different values of shape parameter β as from $MOBWD(1.55, 0.65, 0.35, 0.45)$ and $MOBWD(0.75, 0.65, 0.35, 0.45)$, respectively. For Bayesian inference, hyperparameter values are chosen arbitrarily. For posterior sample generation and computational purposes, R software is used. After eliciting the priors, the Bayes estimates of the parameters are obtained using the M-H algorithm. The Markov chain of length 10^5 is generated and the first 2×10^4 observations are removed as a burn-in period.

Table 3. Average length of HPD interval and coverage probability of the Bayes estimates for $MOBWD(1.55, 0.65, 0.35, 0.45)$.

n	RPPI-I		RPPI-II		
	HPD	CP	HPD	CP	
20	β	0.9978	0.9726	1.0025	0.9738
	λ_1	1.2708	0.9584	1.3431	0.9458
	λ_2	0.3308	0.9404	0.3226	0.9452
	λ_3	0.3604	0.9338	0.3490	0.9456
40	β	0.7294	0.9540	0.7310	0.9522
	λ_1	0.8980	0.9498	0.9201	0.9402
	λ_2	0.2481	0.9274	0.2441	0.9574
	λ_3	0.2697	0.9376	0.2645	0.9512
60	β	0.5998	0.9528	0.6005	0.9532
	λ_1	0.7334	0.9538	0.7450	0.9508
	λ_2	0.2070	0.9474	0.2046	0.9538
	λ_3	0.2239	0.9364	0.2209	0.9474
100	β	0.4688	0.9548	0.4690	0.9552
	λ_1	0.5678	0.9500	0.5731	0.9448
	λ_2	0.1627	0.9442	0.1617	0.9458
	λ_3	0.1762	0.9446	0.1749	0.9490

Table 4. ARB and MSE of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$ under RPPI-I.

n	SELF		LINEX($\gamma=-1.5$)		LINEX($\gamma=1.5$)		
	ARB	MSE	ARB	MSE	ARB	MSE	
20	β	0.1480	0.0210	0.1567	0.0239	0.1409	0.0187
	λ_1	0.1796	0.1182	0.2126	0.1782	0.1613	0.0884
	λ_2	0.2948	0.0078	0.3024	0.0082	0.2873	0.0075
	λ_3	0.2455	0.0090	0.2458	0.0093	0.2454	0.0088
40	β	0.1011	0.0095	0.1040	0.0102	0.0987	0.0090
	λ_1	0.1308	0.0610	0.1424	0.0748	0.1235	0.0523
	λ_2	0.2198	0.0044	0.2221	0.0045	0.2176	0.0043
	λ_3	0.1795	0.0048	0.1803	0.0049	0.1787	0.0048
60	β	0.0823	0.0062	0.0839	0.0065	0.0811	0.0059
	λ_1	0.1048	0.0379	0.1105	0.0433	0.1013	0.0344
	λ_2	0.1794	0.0029	0.1804	0.0029	0.1784	0.0029
	λ_3	0.1487	0.0033	0.1494	0.0034	0.1480	0.0033
100	β	0.0644	0.0037	0.0650	0.0038	0.0639	0.0037
	λ_1	0.0803	0.0220	0.0830	0.0238	0.0786	0.0207
	λ_2	0.1393	0.0018	0.1402	0.0018	0.1390	0.0018
	λ_3	0.1163	0.0021	0.1168	0.0021	0.1162	0.0021

After that remaining observations were chosen at a gap of 15 to remove any serial correlation among them.

The average length of highest posterior density (HPD) and coverage probabilities of interval estimates with varying sample sizes are reported in Tables 3 and 6. The Bayes estimates of the unknown parameters are obtained under SELF as well as the LINEX loss function for $\gamma = (-1.5, 1.5)$ to capture the behavior of estimators in case of underestimation and overestimation. The important findings from the simulated results can be summarized. From simulated results, we can see that all the Bayes estimators perform efficiently with the increase in sample size. MSE and absolute relative bias of the estimators decrease as sample size increases. This is true for both SELF and LINEX loss functions under each reference prior with partial information. The coverage probability for parameters (β, Λ) tends near to the nominal level of 0.95 as the sample size increases. The Bayes estimates obtained under the LINEX loss function for $\gamma = 1.5$ outperform the Bayes estimates under remaining loss functions in terms

Table 5. ARB and MSE of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$ under RPPI-II.

n		SELF		LINEX($\gamma=1.5$)		LINEX($\gamma=1.5$)	
		ARB	MSE	ARB	MSE	ARB	MSE
20	β	0.1500	0.0216	0.1591	0.0247	0.1426	0.0192
	λ_1	0.1995	0.1521	0.2388	0.2461	0.1789	0.1106
	λ_2	0.2666	0.0064	0.2728	0.0068	0.2604	0.0061
40	λ_3	0.2181	0.0072	0.2252	0.0076	0.2140	0.0070
	β	0.1014	0.0096	0.1044	0.0103	0.0990	0.0091
	λ_1	0.1370	0.0674	0.1489	0.0829	0.1298	0.0579
60	λ_2	0.2070	0.0039	0.2090	0.0041	0.2050	0.0038
	λ_3	0.1678	0.0043	0.1693	0.0044	0.1672	0.0042
	β	0.0825	0.0062	0.0840	0.0065	0.0812	0.0059
100	λ_1	0.1080	0.0403	0.1136	0.0458	0.1048	0.0367
	λ_2	0.1718	0.0027	0.1733	0.0027	0.1709	0.0026
	λ_3	0.1424	0.0031	0.1429	0.0031	0.1417	0.0030
	β	0.0644	0.0037	0.0651	0.0038	0.0639	0.0037
	λ_1	0.0817	0.0227	0.0843	0.0245	0.0802	0.0215
	λ_2	0.1367	0.0017	0.1378	0.0017	0.1356	0.0017
	λ_3	0.1137	0.0020	0.1145	0.0020	0.1130	0.0019

Table 6. Average length of HPD interval and coverage probability of the Bayes estimates for $MOBWD(0.75, 0.65, 0.35, 0.45)$.

n		RPPI-I		RPPI-II	
		HPD	CP	HPD	CP
20	β	0.5317	0.9540	0.5350	0.9542
	λ_1	1.2688	0.9656	1.3397	0.9498
	λ_2	0.3321	0.9460	0.3236	0.9534
40	λ_3	0.3616	0.9428	0.3497	0.9532
	β	0.3685	0.9548	0.3691	0.9522
	λ_1	0.9014	0.9522	0.9239	0.9452
60	λ_2	0.2484	0.9250	0.2445	0.9558
	λ_3	0.2697	0.9410	0.2646	0.9532
	β	0.2991	0.9482	0.2997	0.9490
100	λ_1	0.7330	0.9494	0.7443	0.9456
	λ_2	0.2068	0.9428	0.2045	0.9500
	λ_3	0.2238	0.9390	0.2210	0.9504
	β	0.2301	0.9470	0.2304	0.9454
	λ_1	0.5674	0.9496	0.5725	0.9476
	λ_2	0.1627	0.9464	0.1618	0.9468
	λ_3	0.1762	0.9440	0.1749	0.9464

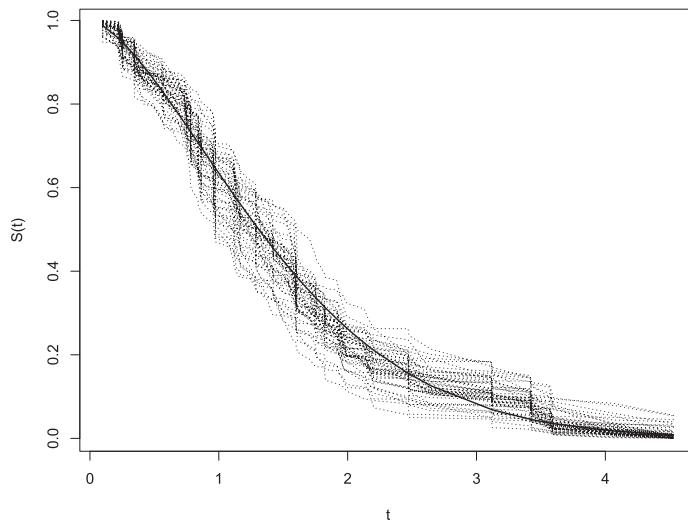
of MSE. The HPD length for the parameter λ_1 under RPPI-I is shorter than the HPD length for the parameter λ_1 under RPPI-II and the pattern remains the same with an increase in sample size. The HPD length for the parameters λ_2 and λ_3 under RPPI-II is shorter than the HPD length for the parameters λ_2 and λ_3 under RPPI-I and the pattern remains the same with an increase in sample size.

6. Real data application

In this section, to elucidate the performance of estimators established in the previous section, two real data sets are utilized. The first data set is from a diabetic retinopathy study, while the second data set is from prostate cancer. These two data sets are analyzed using MOBWD under competing risks set up.

Table 7. Time to blindness in days and causes for 71 patients.

Cause	Time
Eye under Laser Treatment	266, 583, 79, 93, 805, 344, 306, 415, 178, 1484, 315, 1252, 642, 407, 356, 699, 667, 126, 350, 84, 392, 901, 276, 520, 503, 584, 355, 1302
Eye under No Treatment	91, 154, 547, 707, 469, 1313, 790, 125, 777, 307, 637, 577, 517, 287, 717, 141, 427, 36, 588, 350, 567, 1140, 448, 904, 485, 248, 423, 285, 315, 727, 210, 409, 227
Both Eyes	285, 622, 272, 1137, 1653, 471, 663, 966, 203, 1247

**Figure 1.** Estimated survival plots for observed DRS study (solid line) and predictive data sets.

6.1. Diabetic retinopathy study data

A diabetic retinopathy study (DRS) was conducted by National Eye Institute to evaluate the effect of laser treatment on delaying the onset of blindness in patients diagnosed with diabetic retinopathy. One can see Xu and Zhou (2017) for more details on this data set given in **Table 7**. A total of 71 patients with diabetic retinopathy were monitored in this study. Laser treatment was administered to one eye at random, while the other was left untreated. The time to blindness for the patients was recorded for each eye. It was observed that, for 10 individuals (approx 14%), the time to blindness was determined to be the same in both eyes. Considering this data set as competing risks data, the event of interest is blindness of the eye system and the two risks of blindness are treatment and absence of treatment. As a result, the two risks are not independent and are being analyzed using MOBWD. Let us denote X_1 as the time to the blindness of the eye under laser treatment and X_2 being the time to the blindness of the eye in the absence of treatment.

To perform the model predictive capability using some graphical tools, we first generate 50 posterior samples of the same size as the size of DRS study data corresponding to the dependent competing risks model using the distribution of time to the blindness of eyes of diabetic patients. For each generated predictive sample, the survival rate is calculated.

Table 8. Bayes estimates of parameters of DRS data under SELF and LINEX loss function under RPPI-I.

Parameters	SELF	LINEX		95% HPD		
		$\gamma = -1.5$	$\gamma = 1.5$	Lower	Upper	Length
β	1.5548	1.5629	1.5467	1.3479	1.7557	0.4078
λ_1	0.4713	0.4747	0.4680	0.3414	0.5991	0.2578
λ_2	0.3949	0.3973	0.3925	0.2857	0.5057	0.2200
λ_3	0.4646	0.4672	0.4621	0.3495	0.5777	0.2283

Table 9. Bayes estimates of parameters of DRS data under SELF and LINEX loss function under RPPI-II.

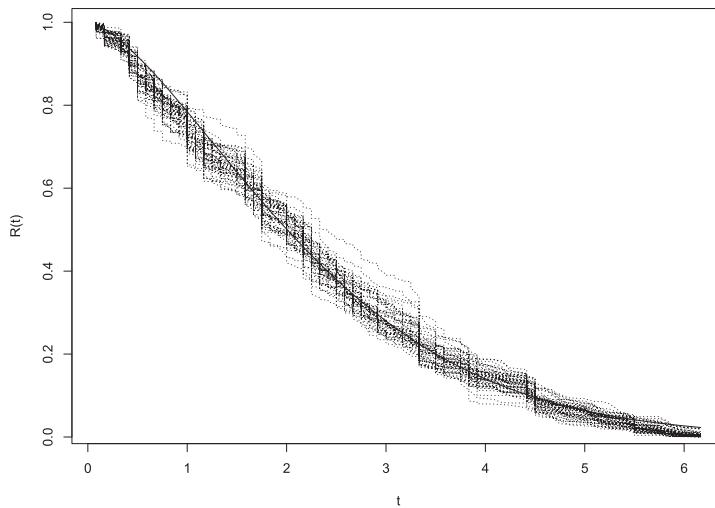
Parameters	SELF	LINEX		95% HPD		
		$\gamma = -1.5$	$\gamma = 1.5$	Lower	Upper	Length
β	1.5559	1.5643	1.5476	1.3431	1.7546	0.4115
λ_1	0.4710	0.4746	0.4676	0.3371	0.6031	0.2660
λ_2	0.3943	0.3967	0.3918	0.2854	0.5077	0.2224
λ_3	0.4642	0.4668	0.4616	0.3531	0.5828	0.2297

Figure 1 shows that the survival rate of the blindness of eyes for diabetic patients is greatly superimposed by the survival rates of the 50 predictive samples. Now we perform the model compatibility in the form of PPV discussed in Subsection 4.5. We generate 200 posterior samples of the same size as the size of DRS study data corresponding to the dependent competing risks model. Further, we obtain the value of χ^2_1 for DRS study data for each generated posterior sample on parameters (β, Λ) . In the next step, we generated 1,000 predictive samples from the model, each of the same sizes as of DRS study data, for each value of the set of parameters (β, Λ) . Based on these 1000 predictive samples, we calculated χ^2_2 . On comparing the value of χ^2_2 with χ^2_1 , we obtain $P[\chi^2_2 > \chi^2_1]$. Finally, the above procedure is repeated a sufficient number of times to calculate PPV. We found the estimated value of PPV equals 0.383. Thus, with a greater value of PPV, we can conclude that the considered model (i.e., likelihood and prior) is compatible with the considered data set.

The Bayesian analysis of the considered data set is carried out under reference priors RPPI-I and RPPI-II derived in Subsection 4.1. The Bayesian estimates are obtained considering SELF and LINEX loss functions under both priors. The values of hyperparameters are chosen arbitrary. Since we do not have any prior information for the real data, we have taken all the hyperparameters to be 0.001. Using the full conditionals (4.14) to (4.19), sample values are generated for parameters (β, Λ) following the MCMC procedure defined in Subsection 4.3. The MCMC algorithm is applied to generate 10^5 posterior samples on the parameters and the first 2×10^4 samples are discarded to eliminate the effect of the initial value used for iteration. After that for removing the serial correlation of sampled draws, one value is chosen in every 10 iterations. Thus, the Bayesian estimates are based on 8,000 posterior samples. The Bayesian estimates of the parameters (β, Λ) along with their 95% HPD interval and HPD length under RPPI-I and RPPI-II are presented in Tables 8 and 9, respectively. We further perform the Bayesian testing of hypotheses to test whether the two competing risks are identical or not. To do so, we calculate the Bayes factor in (4.27) and (4.29) obtained under RPPI-I and RPPI-II, respectively. The calculated Bayes factor comes out to be very low, so the null hypothesis of the two risks being identical can not be rejected in the Bayesian context; hence, due to the lack of proper evidence, we can conclude that the laser treatment is not effective in delaying the blindness of eyes in diabetic patients.

Table 10. Patients summary statistics of prostate cancer data.

Cause of death	<i>n</i>	Minimum	First quartile	Median	Mean	Third quartile	Maximum	SD
Prostatic cancer	127	2.00	12.00	24.00	26.26	37.00	71.00	17.35
Cerebrovascular	31	3.00	21.00	33.00	31.32	42.00	60.00	15.14
Other causes	24	1.00	8.25	20.00	25.75	35.50	74.00	22.79

**Figure 2.** Estimated survival plots for observed prostate cancer data (solid line) and predictive data sets.

6.2. Prostate cancer data

Here, we have used the prostate cancer data set, which was first reported in Byar and Green (1980). The data set originally consists of 502 patients. We consider only 182 patients for our purpose corresponding to the major causes of death related to patients. The major causes of death are distributed for these 182 patients as 127 patients died of prostatic cancer; 31 patients died of cerebrovascular accident, and 24 patients died of causes other than these two. The main purpose of analyzing this data set is that we can compare the risk of death under all causes except cerebrovascular accident with all causes except prostatic cancer. Before analyzing the data set, we first form the groups of patients as per their cause of death to make the data suitable for our dependent competing risks model. The first group includes those patients who died of prostatic cancer or other causes, and the second group includes those who died of a cerebrovascular accident or other causes. The summary statistics of the patients according to their cause of death are given in Table 10.

For j^{th} patient, let X_{1j} and X_{2j} be the latent time of death (in months) due to group 1 and group 2 of causes, respectively. So that $T_j = \min(X_{1j}, X_{2j})$ be the observed death time of the j^{th} patient. Hence, if the patient dies of prostatic cancer, then only X_{1j} is observed; if the patient dies of cerebrovascular accident, then only X_{2j} is observed; and if the patient dies of other causes, then X_{1j} and X_{2j} both are observed simultaneously.

We first go for the model predictive capability using some graphical tool based on survival rate. For this, we first generate 50 posterior samples of the same size as the size of prostate cancer data corresponding to the dependent competing risks model using the distribution of

Table 11. Bayes estimates of parameters of prostate cancer data under SELF and LINEX loss function under RPPI-I.

Parameters	LINEX			95% HPD		
	SELF	$\gamma = -1.5$	$\gamma = 1.5$	Lower	Upper	Length
β	1.5062	1.5090	1.5035	1.3810	1.6201	0.2391
λ_1	0.2531	0.2536	0.2525	0.2025	0.3047	0.1021
λ_2	0.6980	0.6989	0.6971	0.6281	0.7632	0.1352
λ_3	0.1701	0.1707	0.1695	0.1172	0.2259	0.1086

Table 12. Bayes estimates of parameters of prostate cancer data under SELF and LINEX loss function under RPPI-II.

Parameters	LINEX			95% HPD		
	SELF	$\gamma = -1.5$	$\gamma = 1.5$	Lower	Upper	Length
β	1.5068	1.5096	1.5039	1.3814	1.6235	0.2421
λ_1	0.2527	0.2532	0.2522	0.2025	0.3037	0.1012
λ_2	0.6971	0.6979	0.6962	0.6303	0.7618	0.1315
λ_3	0.1707	0.1712	0.1701	0.1201	0.2274	0.1073

time to death for patients. For each generated predictive sample, the survival rate is calculated. Figure 2 shows that the survival rate of patients is greatly superimposed by the survival rates of the 50 predictive samples. Before performing an estimation procedure for unknown parameters for the prostate cancer data, we check the model compatibility in the form of PPV discussed in Subsection 4.5. To calculate the PPV, we adopted the same procedure as in the case of the DRS study data. We found the estimated value of PPV equals 0.619. Thus, with a greater value of PPV, we can conclude that the considered model (i.e., likelihood and prior) in the Bayesian paradigm is compatible with the considered data set.

Now, we are in a position to perform the Bayesian analysis of prostate cancer data. The Bayesian estimation procedure is carried out under reference priors RPPI-I and RPPI-II derived in Subsection 4.1. Since, for the real data, we do not have any prior information, the value of hyperparameters is taken to be 0.001. The MCMC algorithm is applied to generate 10^5 posterior samples on the parameters and the first 2×10^4 samples are discarded to eliminate the effect of the initial value used for iteration. After that for removing the serial correlation of sampled draws, one value is chosen in every 10 iterations. Thus, the Bayesian estimates are based on 8,000 posterior samples. The Bayesian estimates are obtained considering SELF and LINEX loss functions under both priors. The values of hyperparameters are chosen arbitrary. Using the full conditionals (4.14) to (4.19), sample values are generated for parameters (β, Λ) following the MCMC procedure defined in Subsection 4.3. The Bayesian estimates of the parameters (β, Λ) along with their 95% HPD interval and HPD length under RPPI-I and RPPI-II are presented in Tables 11 and 12, respectively.

7. Conclusion

In this article, we have provided the Bayesian method for analyzing the dependent competing risks data arising from MOBWD distribution. We have derived the reference prior for the unknown parameters of MOBWD when partial prior information is available to some of the parameters only. The reference prior is derived for two cases when the prior information is

available with parameters (β, λ_1) only and when with (λ_2, λ_3) only. We also established that the derived prior, yet improper, leads to a suitable posterior distribution for the parameters. The Bayesian estimates of the unknown parameters are obtained under the two reference priors RPPI-I and RPPI-II for SELF and LINEX loss functions. As a result, if some of the parameters have no prior information, the reference prior with partial information obtained is a good option to carry out Bayesian estimation. Although this article focused on the estimation of unknown parameters of the Marshall-Olkin bivariate Weibull distribution, the proposed method can be applied to other bivariate distributions with partial prior information. The proposed method can also be extended for competing risks data under different censoring schemes, such as Type-I censoring, Type-II censoring, progressively hybrid censoring, etc., and also for incomplete data problems.

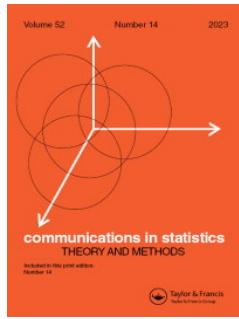
Funding

This work is partially sponsored by the Institute of Eminence Scheme no. 6031 of Banaras Hindu University, India.

References

- Barnwal, V., and M. S. Panwar. 2022. Competing risks analysis for dependent causes using Marshall-Olkin bivariate generalized lifetime family. *Communications in Statistics- Theory and Methods*. doi:[10.1080/03610926.2022.2094412](https://doi.org/10.1080/03610926.2022.2094412).
- Berger, J. O., and J. M. Bernardo. 1992. On the development of the reference prior method. *Bayesian Statistics* 4 (4):35–60.
- Bernardo, J. M. 1979. Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society: Series B (Methodological)* 41 (2):113–28. doi:[10.1111/j.2517-6161.1979.tb01066.x](https://doi.org/10.1111/j.2517-6161.1979.tb01066.x).
- Byar, D. P., and S. B. Green. 1980. The choice of treatment for cancer patients based on covariate information. *Bulletin Du Cancer* 67 (4):477.
- Chen, M. H., Q. M. Shao, and J. G. Ibrahim. 2000. *Monte Carlo methods in Bayesian computation*. New York: Springer-Verlag. doi:[10.1007/978-1-4612-1276-8](https://doi.org/10.1007/978-1-4612-1276-8)
- Gelman, A., X. L. Meng, and H. Stern. 1996. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* 6 (4):733–60.
- Gelman, A., and D. B. Rubin. 1992. A single series from the Gibbs sampler provides a false sense of security. *Bayesian Statistics* 4:625–31.
- Ghosh, J., and R. Mukerjee. 1992. Noninformative priors (with discussion). *Bayesian Statistics* 4: 195–210.
- Ghosh, M. 2011. Objective priors: An introduction for frequentists. *Statistical Science* 26 (2):187–202.
- Guttman, I. 1967. The use of the concept of a future observation in goodness-of-fit problems. *Journal of the Royal Statistical Society: Series B (Methodological)* 29 (1):83–100. doi:[10.1111/j.2517-6161.1967.tb00676.x](https://doi.org/10.1111/j.2517-6161.1967.tb00676.x).
- Hastings, W. K. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57 (1):97–109. doi:[10.1093/biomet/57.1.97](https://doi.org/10.1093/biomet/57.1.97).
- Kass, R. E., and L. Wasserman. 1996. The selection of prior distributions by formal rules. *Journal of the American Statistical Association* 91 (435):1343–70. doi:[10.1080/01621459.1996.10477003](https://doi.org/10.1080/01621459.1996.10477003).
- Kundu, D., and A. K. Dey. 2009. Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm. *Computational Statistics & Data Analysis* 53 (4):956–65. doi:[10.1016/j.csda.2008.11.009](https://doi.org/10.1016/j.csda.2008.11.009).
- Kundu, D., and A. K. Gupta. 2013. Bayes estimation for the Marshall–Olkin bivariate Weibull distribution. *Computational Statistics & Data Analysis* 57 (1):271–81. doi:[10.1016/j.csda.2012.06.002](https://doi.org/10.1016/j.csda.2012.06.002).
- Marshall, A. W., and I. Olkin. 1967. A multivariate exponential distribution. *Journal of the American Statistical Association* 62 (317):30–44. doi:[10.1080/01621459.1967.10482885](https://doi.org/10.1080/01621459.1967.10482885).

- Metropolis, N., and S. Ulam. 1949. The Monte Carlo method. *Journal of the American Statistical Association* 44 (247):335–41. doi:[10.1080/01621459.1949.10483310](https://doi.org/10.1080/01621459.1949.10483310). 18139350
- Seo, J. I. 2020. Objective Bayesian analysis for the Weibull distribution with partial information under the generalized type-II progressive hybrid censoring scheme. *Communications in Statistics-Simulation and Computation*, 51 (9):5157–73. doi:[10.1080/03610918.2020.1758138](https://doi.org/10.1080/03610918.2020.1758138).
- Seo, J. I., and Y. Kim. 2017. Objective Bayesian analysis based on upper record values from two-parameter Rayleigh distribution with partial information. *Journal of Applied Statistics* 44 (12):2222–37. doi:[10.1080/02664763.2016.1251886](https://doi.org/10.1080/02664763.2016.1251886).
- Shen, Y., and A. Xu. 2018. On the dependent competing risks using Marshall–Olkin bivariate Weibull model: Parameter estimation with different methods. *Communications in Statistics- Theory and Methods* 47 (22):5558–72. doi:[10.1080/03610926.2017.1397170](https://doi.org/10.1080/03610926.2017.1397170).
- Sun, D., and J. O. Berger. 1998. Reference priors with partial information. *Biometrika* 85 (1):55–71. doi:[10.1093/biomet/85.1.55](https://doi.org/10.1093/biomet/85.1.55).
- Xu, A., and Y. Tang. 2011. Bayesian analysis of Birnbaum–Saunders distribution with partial information. *Computational Statistics & Data Analysis* 55 (7):2324–33. doi:[10.1016/j.csda.2011.01.021](https://doi.org/10.1016/j.csda.2011.01.021).
- Xu, A., and S. Zhou. 2017. Bayesian analysis of series system with dependent causes of failure. *Statistical Theory and Related Fields* 1 (1):128–40. doi:[10.1080/24754269.2017.1348708](https://doi.org/10.1080/24754269.2017.1348708).
- Zhang, L., A. Xu, L. An, and M. Li. 2022. Bayesian Inference of System Reliability for Multicomponent Stress-Strength Model under Marshall-Olkin Weibull Distribution. *Systems* 10 (6):196. doi:[10.3390/systems10060196](https://doi.org/10.3390/systems10060196).



Competing risks analysis for dependent causes using Marshall-Olkin bivariate generalized lifetime family

Vikas Barnwal & M. S Panwar

To cite this article: Vikas Barnwal & M. S Panwar (2022): Competing risks analysis for dependent causes using Marshall-Olkin bivariate generalized lifetime family, *Communications in Statistics - Theory and Methods*, DOI: [10.1080/03610926.2022.2094412](https://doi.org/10.1080/03610926.2022.2094412)

To link to this article: <https://doi.org/10.1080/03610926.2022.2094412>



Published online: 04 Jul 2022.



Submit your article to this journal [↗](#)



Article views: 1041



View related articles [↗](#)



View Crossmark data [↗](#)



Competing risks analysis for dependent causes using Marshall-Olkin bivariate generalized lifetime family

Vikas Barnwal and M. S Panwar 

Department of Statistics, Banaras Hindu University, Varanasi, India

ABSTRACT

The assumption of independence of causes for modeling a competing risks scenario is not presumable always. In this article, cause dependent competing risks model has been analyzed under Marshall-Olkin set-up. A Marshall-Olkin generalized lifetime distribution has been established to address a competing risks model. The essential statistical properties have been derived for this distribution to utilize in a competing risks model. The estimators of unknown simple/transformed parameters are obtained by using two approaches: maximum likelihood method and Bayesian estimation through reference prior. To examine the performance of the five models under generalized lifetime family, a simulation study has been performed. For illustration, two real data sets namely, prostate cancer and diabetic retinopathy study are considered. For both data sets, analysis is performed using the most suitable model from Marshall-Olkin generalized lifetime family.

ARTICLE HISTORY

Received 17 September 2021
Accepted 14 June 2022

KEYWORDS

Aging class; competing risks; generalized family of lifetime distribution; Marshall-Olkin distribution; reference prior

1. Introduction

In multifarious circumstances, one is interested in the measurement of “risks”- a risk of breakdown, a risk of becoming ill, a risk of getting infected, and so forth. A specific problem arises when several risk factors are presented altogether. In such situations, the existence of one risk always entangles the inference of another. For an individual, when the several risks compete for occurrence of the event of interest, referred to as “competing risks” in literature. Competing risks set-up can be treated as a specific case of the multivariate model where causes or risks are considered as the multiple variables of an individual.

The information available in competing risks set-up is the time to event and the corresponding cause of failure for the individuals under observation. The collected information may be complete or sometimes may be incomplete with respect to time and/or cause. Sinha (1986) and Lawless (2003), described the conventional competing risks model to analyze complete data. In an experiment where the researcher is unable to obtain the complete information, then adopting a censoring procedure becomes an inescapable choice to analyze competing risks data. Assuming the lifetimes of the different causes as independent exponential distributions, Kundu, Kannan and Balakrishnan

CONTACT M. S Panwar  vir.panwar@gmail.com  Department of Statistics, Banaras Hindu University, Varanasi 221005, India.

© 2022 Taylor & Francis Group, LLC

(2003) provided the estimates under progressive Type-II censored data. In many situations, the time to failure is known but the cause of failure is masked/hidden and for such observations Miyakawa (1984), Kuo and Yang (2000) and Tomer, Singh and Panwar (2014) extended the competing risks model accordingly.

In real-life, it is almost impossible for an individual to have two or more independent risks of failure. However, in the literature, it can be noticed that a limited work has been done on the cause dependent competing risks modeling. Moeschberger (1974) defines that “the theoretical lifetime of an individual failing from one cause may be correlated with the theoretical lifetime of the same individual failing from a different cause”. A bivariate exponential lifetime model has been used by Wada, Sen and Shimakura (1996) for the analysis of competing risks data involving two dependent risk components. Wang and Ghosh (2003) studied the dependent competing risks model in the presence of covariates, whereas Lindqvist and Skogsrud (2008) analyzed if preventive maintenance has been taken for the components in between the inspection.

In competing risks problems, it is challenging to derive the distribution of marginal risks. Several times, if the marginals are completely known, then the joint distribution of individual failure times is not identifiable. For a competing risks problem, Tsiatis (1975) proved that the joint distribution of failure times is not identifiable by their minimum in the multivariate case. Specifically, the identifiability for the Marshall-Olkin model is discussed by Basu and Ghosh (1978). In literature, Crowder (1991) referred to this issue as the “identifiability crisis in competing risks analysis”.

Marshall and Olkin (1967) derived the bivariate exponential distribution and demonstrated that it has both an absolutely continuous and a singular part. Further, Proschan and Sullo (1973) obtained the likelihood function for the multivariate Marshall-Olkin bivariate exponential distribution (MOBED) parameters under the competing risks set-up. Kundu and Gupta (2013) developed the inferential procedure for Bayesian estimation of the parameters of Marshall-Olkin bivariate Weibull distribution (MOBWD) under informative priors. It is inevitable to perform Bayesian estimation under non informative priors when there is no prior information about the parameters. For the multiparameter case, one of the popular non informative priors is “reference prior” suggested by Berger and Bernardo (1992).

The key idea behind the formulation of reference prior is that it maximizes the Kullback-Leibler divergence between the prior and posterior density asymptotically, i.e.,

$$\pi_r(\theta) = \arg \max_{\pi(\theta)} K(\theta, X_n),$$

where

$$K(\theta, X_n) = E_X \left[\int \pi(\theta|X_n) \log \left\{ \frac{\pi(\theta|X_n)}{\pi(\theta)} \right\} d\theta \right],$$

which is the expected Kullback-Leibler divergence.

Cai, Shi and Liu (2017) analyzed MOBWD under Type-I progressive hybrid censored condition for dependent competing risks incomplete data. The Bayesian analysis performed with non informative prior is somewhat known as objective Bayesian analysis. Guan, Tang and Xu (2013) considered the objective Bayesian analysis for MOBED for dependent causes of failure. Recently, Feizjavadian and Hashemi (2015), Xu and Zhou

(2017), and Shen and Xu (2018) performed competing risks analysis using Marshall-Olkin set-up under different types of observations.

In this article, a bivariate probability distribution set-up is considered, where it is assumed that there is a dependency between the competing risks. A generalized lifetime distribution (GLD) family, considered by Moore and Bilikam (1978) and Chaturvedi and Khumukcham (2006), has been taken to capture the wide range of real data characteristics. We derive the Marshall-Olkin bivariate distribution for the generalized lifetime family, which accommodates various shapes of the hazard function. So we provide a general structure to obtain the estimators for different bivariate distributions' parameters into one framework. This makes it easier for anyone to analyze the distributions available on one hand. We use the term MOBGLD (Marshall-Olkin bivariate generalized lifetime distribution) for bivariate family of distributions obtained using the same process given by Marshall and Olkin (1967).

The rest of the article is organized as follows. In Section 2, we first establish the MOBGLD and then derive its important properties to utilize in the cause dependent competing risks analysis. The procedure to obtain the maximum likelihood (ML) estimators and variance-covariance matrix has been given for direct parameters and re-parameterize cases as well in Section 3. In Section 4, we approach through the ordered grouping of parameters in reference priors under the Bayesian method. Section 5 is dedicated to simulation study for the derived theory. In Section 6, two real data sets, the prostate cancer and diabetic retinopathy study have been taken to illustration purpose. Finally, we conclude the article with a summary in Section 7.

2. Model building for dependent competing risks

In this section, our objective is to propose a model which is suitable for the cause dependent competing risks set-up. All its important properties, required for the cause dependent modeling, have been derived.

2.1. Generalized lifetime distribution

In literature, a GLD family is used to achieve a wide range of applicability in a single effort. Due to this fact, a GLD family, introduced by Moore and Bilikam (1978) has been considered. A random variable (rv), X , is said to follow the GLD if its probability density function(pdf) $f_{GLD}(x; \beta, \theta)$ and cumulative distribution function(cdf) $F_{GLD}(x; \beta, \theta)$, $x \in \mathbb{R}^+$, respectively defined as

$$f_{GLD}(x; \beta, \theta) = \beta \theta g'(x) [g(x)]^{\beta-1} \exp \left\{ -\theta [g(x)]^\beta \right\}, \quad (2.1)$$

and

$$F_{GLD}(x; \beta, \theta) = 1 - \exp \left\{ -\theta [g(x)]^\beta \right\}, \quad (2.2)$$

where β and $\theta > 0$ are the parameters. The function $g(x)$ in (2.1) is a real-valued, differentiable and strictly increasing function of x . Here, it is assumed that the inverse of $g(x)$ exists where $g(0+) = 0$ and $g(\infty) = \infty$. For the specific choices of $g(x)$, the particular cases of GLD can be obtained as given in Table 1.

Table 1. Distributions for specific forms of $g(x)$ and parameters for GLD.

$g(x)$	(β, θ)	$f_{GLD}(x; \beta, \theta)$	Distribution
x	$(1, \theta)$	$f_E(x; \theta)$	Exponential Distribution
x	$(2, \theta)$	$f_R(x; \theta)$	Rayleigh Distribution
$\ln(x/a), a > 0$	$(1, \theta)$	$f_P(x; \theta)$	Pareto Distribution
$\ln(1 + x^b), b > 0$	$(1, \theta)$	$f_B(x; \theta)$	Burr Distribution
x	(β, θ)	$f_W(x; \beta, \theta)$	Weibull Distribution

In this article, for the inferential purpose, two cases, when θ and β both are unknown and when θ is unknown but β is known, are considered. The survival function $S_{GLD}(x; \beta, \theta)$ and hazard rate $h_{GLD}(x; \beta, \theta)$ from (2.1), respectively, given by

$$S_{GLD}(x; \beta, \theta) = \exp \left\{ -\theta [g(x)]^\beta \right\} \quad (2.3)$$

and

$$h_{GLD}(x; \beta, \theta) = \beta \theta g'(x) [g(x)]^{\beta-1}. \quad (2.4)$$

It is to be noticeable from Table 1 that GLD covers a variety of hazard functions, such as constant, increasing as well as decreasing hazard rate.

2.2. Marshall-Olkin bivariate generalized lifetime distribution

In this section, first we will establish Marshall-Olkin bivariate distribution (MOBD), proposed by Marshall and Olkin (1967), for the GLD family (2.1) and then its important properties will be derived. Let U_0 , U_1 and U_2 are three independent rv following the same distribution from the GLD family (2.1) with parameters β and $\theta_i; i = 0, 1, 2.$, i.e., $U_0 \sim GLD(\beta, \theta_0)$, $U_1 \sim GLD(\beta, \theta_1)$ and $U_2 \sim GLD(\beta, \theta_2)$. After defining $X_1 = \min(U_0, U_1)$ and $X_2 = \min(U_0, U_2)$, the bivariate set-up of rv (X_1, X_2) follows MOBGLD with unknown parameters $(\beta, \Theta) = (\beta, \theta_0, \theta_1, \theta_2)$. This set-up can be utilized in competing risks modeling successfully. If we assume that X_1 and X_2 denotes the observations of two components attached to a series system, then $T = \min(X_1, X_2)$ will represent the system failure time. In this approach, the correlation parameter θ_0 plays an important role which facilitates MOBGLD to model onto a cause dependent competing risks problems. As well as the above discussed competing system will reduce MOBGLD with independent marginals if the correlation parameter θ_0 becomes zero. First, for the future purpose, let derive some important properties for MOBGLD. These properties will be helpful for further derivations.

Theorem 2.1. *If $X_i; i = 1, 2$ is defined as $X_i = \min(U_0, U_i)$ where U_0, U'_i s are independent rv following $GLD(\beta, \theta_j) \forall j = 0, 1, 2$. Then for the bivariate rv $(X_1, X_2) \sim MOBGLD(\beta, \Theta)$ and $\Delta = \theta_0 + \theta_1 + \theta_2$, we have*

- (i) $X_i \sim GLD(\beta, \theta_0 + \theta_i); i = 1, 2,$
- (ii) $T = \min\{X_1, X_2\} \sim GLD(\beta, \Delta).$

Proof. (i) The survival function of rv $X_i; i = 1, 2$ is defined as

$$S(x_i) = Pr[X_i > x_i] = Pr[\min(U_0, U_i) > x_i]$$

Since U_0 and U'_i s are independent, so

$$\begin{aligned} S(x_i) &= \Pr[U_0 > x_i] \cdot \Pr[U_i > x_i] \\ &= \exp \left\{ -(\theta_0 + \theta_i) [g(x_i)]^\beta \right\}, \end{aligned}$$

which is the survival function of $GLD(\beta, \theta_0 + \theta_i)$ from (2.3) and hence $X_i \sim GLD(\beta, \theta_0 + \theta_i)$; $i = 1, 2$.

(ii) Let $T = \min(X_1, X_2)$ then the survival function of T is given by

$$\begin{aligned} S(t) &= \Pr[\min(X_1, X_2) > t] \\ &= \Pr[(\min(U_0, U_1) > t), (\min(U_0, U_2) > t)] \end{aligned}$$

Since U_0, U_1 and U_2 are independent, so

$$\begin{aligned} S(t) &= \Pr[U_0 > t] \cdot \Pr[U_1 > t] \cdot \Pr[U_2 > t] \\ &= \exp \left\{ -\Delta [g(t)]^\beta \right\}. \end{aligned}$$

Which is the survival function of $GLD(\beta, \Delta)$ from (2.3) and hence $T \sim GLD(\beta, \Delta)$, where $\Delta = \theta_0 + \theta_1 + \theta_2$. \square

Theorem 2.2. If a bivariate rv (X_1, X_2) follows the MOBGLD and $X_0 = \max(X_1, X_2)$, then the joint survival function of MOBGLD is given by

$$S(x_1, x_2) = \prod_{i=0}^2 \exp \left\{ -\theta_i [g(x_i)]^\beta \right\}. \quad (2.5)$$

Proof. By the definition of rv X_1 and X_2 , the joint survival function of (X_1, X_2) is given by

$$\begin{aligned} S(x_1, x_2) &= P(X_1 > x_1, X_2 > x_2) \\ &= P(U_0 > x_0, U_1 > x_1, U_2 > x_2), \quad \text{where } x_0 = \max\{x_1, x_2\} \\ &= P(U_0 > x_0)P(U_1 > x_1)P(U_2 > x_2) \\ &= \prod_{i=0}^2 \exp \left\{ -\theta_i [g(x_i)]^\beta \right\} \end{aligned}$$

Corollary 1. The joint survival function of (X_1, X_2) for MOBGLD can also be defined as [see Marshall and Olkin 1967]

$$S(x_1, x_2) = \begin{cases} S_0(x_1, x_2) = S_{GLD}(x; \beta, \Delta) & \text{if } x_1 = x_2 = x, \\ S_1(x_1, x_2) = S_{GLD}(x_1; \beta, \theta_0 + \theta_1)S_{GLD}(x_2; \beta, \theta_2) & \text{if } x_1 > x_2, \\ S_2(x_1, x_2) = S_{GLD}(x_2; \beta, \theta_0 + \theta_2)S_{GLD}(x_1; \beta, \theta_1) & \text{if } x_2 > x_1. \end{cases} \quad (2.6)$$

\square

Theorem 2.3. *The joint pdf for (X_1, X_2) for MOBGLD is given by*

$$f(x_1, x_2) = \begin{cases} f_0(x_1, x_2) = \frac{\theta_0}{\Delta} f_{GLD}(x; \beta, \Delta) & \text{if } x_1 = x_2 = x, \\ f_1(x_1, x_2) = f_{GLD}(x_1; \beta, \theta_0 + \theta_1) f_{GLD}(x_2; \beta, \theta_2) & \text{if } x_1 > x_2, \\ f_2(x_1, x_2) = f_{GLD}(x_2; \beta, \theta_0 + \theta_2) f_{GLD}(x_1; \beta, \theta_1) & \text{if } x_2 > x_1. \end{cases} \quad (2.7)$$

Proof. Let first consider the case when $x_1 > x_2$. The joint pdf $f_1(x_1, x_2)$ can be obtained by the second order partial differentiation of $S_1(x_1, x_2)$ from (2.6) as follows

$$\begin{aligned} f_1(x_1, x_2) &= \frac{\partial^2}{\partial x_1 \partial x_2} S_1(x_1, x_2) \\ &= \frac{\partial}{\partial x_1} S_{GLD}(x_1; \beta, \theta_0 + \theta_1) \frac{\partial}{\partial x_2} S_{GLD}(x_2; \beta, \theta_2) \\ &= f_{GLD}(x_1; \beta, \theta_0 + \theta_1) f_{GLD}(x_2; \beta, \theta_2) \end{aligned}$$

Similarly when $x_2 > x_1$, by the second order partial differentiation of $S_2(x_1, x_2)$ from (2.6), we have

$$f_2(x_1, x_2) = f_{GLD}(x_2; \beta, \theta_0 + \theta_2) f_{GLD}(x_1; \beta, \theta_1).$$

The joint pdf $f_0(x_1, x_2)$, when $x_1 = x_2 = x$ can be obtained by using the fact that

$$\int_0^\infty \int_{x_2}^\infty f_1(x_1, x_2) dx_2 dx_1 + \int_0^\infty \int_{x_1}^\infty f_2(x_1, x_2) dx_1 dx_2 + \int_0^\infty f_0(x) dx = 1 \quad (2.8)$$

From the first integral expression of left hand side of Equation (2.8), we have

$$\begin{aligned} \int_0^\infty \int_{x_2}^\infty f_1(x_1, x_2) dx_1 dx_2 &= \int_0^\infty \left[\int_{x_2}^\infty f_{GLD}(x_1; \beta, \theta_0 + \theta_1) f_{GLD}(x_2; \beta, \theta_2) dx_1 \right] dx_2 \\ &= \frac{\theta_2}{\Delta} \int_0^\infty f_{GLD}(x_2; \beta, \Delta) dx_2 \\ &= \frac{\theta_2}{\Delta}. \end{aligned}$$

Similarly from the second integral expression of Equation (2.8), we get

$$\int_0^\infty \int_{x_1}^\infty f_2(x_1, x_2) dx_2 dx_1 = \frac{\theta_1}{\Delta}.$$

By replacing the above obtained values in Equation (2.8) and as $rv X$ follows $GLD(\beta, \theta)$ given in (2.1), it is easy to obtain that

$$f_0(x_1, x_2) = \frac{\theta_0}{\Delta} f_{GLD}(x; \beta, \Delta).$$

Corollary 2. The joint hazard function, $h(x_1, x_2)$, for bivariate rv (X_1, X_2) of MOBGLD is

$$h(x_1, x_2) = \begin{cases} h_0(x_1, x_2) = \beta\theta_0g'(x)[g(x)]^{\beta-1} & \text{if } x_1 = x_2 = x, \\ h_1(x_1, x_2) = \beta^2(\theta_0 + \theta_1)\theta_2g'(x_1)g'(x_2)[g(x_1)g(x_2)]^{\beta-1} & \text{if } x_1 > x_2, \\ h_2(x_1, x_2) = \beta^2(\theta_0 + \theta_2)\theta_1g'(x_1)g'(x_2)[g(x_1)g(x_2)]^{\beta-1} & \text{if } x_2 > x_1. \end{cases} \quad (2.9)$$

Proof. The hazard function of a bivariate random vector (X_1, X_2) can be defined as $h(x_1, x_2) = f(x_1, x_2)/S(x_1, x_2)$. So let consider the case when $X_1 > X_2$ then the hazard function can be obtained as

$$\begin{aligned} h_1(x_1, x_2) &= \frac{f_{GLD}(x_1; \beta, \theta_0 + \theta_1)f_{GLD}(x_2; \beta, \theta_2)}{S_{GLD}(x_1; \beta, \theta_0 + \theta_1)S_{GLD}(x_2; \beta, \theta_2)} \\ &= h_{GLD}(x_1; \beta, \theta_0 + \theta_1)h_{GLD}(x_2; \beta, \theta_2) \\ &= \beta^2(\theta_0 + \theta_1)\theta_2g'(x_1)g'(x_2)[g(x_1)g(x_2)]^{\beta-1}. \end{aligned}$$

Similarly for $X_2 > X_1$ and $X_1 = X_2 = X$, respectively, we get

$$h_2(x_1, x_2) = \beta^2(\theta_0 + \theta_2)\theta_1g'(x_1)g'(x_2)[g(x_1)g(x_2)]^{\beta-1},$$

and

$$h_0(x_1, x_2) = \beta\theta_0g'(x)[g(x)]^{\beta-1}.$$

□

2.3. Aging properties for MOBGLD

If for an individual, multiple causes are present, then different causes may have non identical aging behavior with respect to their survival time. So it becomes necessary to inspect the aging behavior of such systems. If the bivariate random vector $(X_1, X_2) \sim MOBGLD(\beta, \Theta)$ in such a way that $S(x_1 + t, x_2 + t)/S(x_1, x_2)$ monotonically decreases/increases in $x_1, x_2, t \in \mathbb{R}^+$, we say that (X_1, X_2) have multivariate increasing/decreasing failure rate (MIFR/MDFR) property [see Barlow and Proschan 1975]. Obviously, constant hazard rate (CFR) is a special class of both. Based on this, a classification of MOBGLD has been presented in Table 2.¹

2.4. Dependent competing risks for MOBGLD

As the objective of this article is to inference the lifetime of cause dependent competing risks, so let first match with the competing risks set-up for MOBGLD. Suppose that there are n independent series systems with two causes of failure in a lifetime experiment. Here, we assume that the causes are not independent. Let $T_j = \min\{X_{1j}, X_{2j}\}$ be failure time for the j^{th} series system; $j = 1, 2, \dots, n$; where X_{ij} denotes the failure time of j^{th} system due to i^{th} cause, $i = 1, 2$. Further assume that (X_{1j}, X_{2j}) are the bivariate random vector such as $(X_{1j}, X_{2j}) \sim MOBGLD(\beta, \Theta)$. Let $(\delta_{1j}, \delta_{2j})$ is the indicator for the components causing failure for the system, i.e., when $T_j = X_{1j} = X_{2j}$, $T_j = X_{1j} < X_{2j}$ and $T_j = X_{2j} < X_{1j}$ then one get $(1, 1), (1, 0)$ and $(0, 1)$, respectively. Here it is to be noted

Table 2. Aging classification of MOBGLD for specific distributions.

Model	Parameter	Survival function	Class
MOBE	$\beta = 1$	$\prod_{i=0}^2 \exp\{-\theta_i x_i\}$	MCFR
MOBR	$\beta = 2$	$\prod_{i=0}^2 \exp\{-\theta_i x_i^2\}$	MIFR
MOBP	$\beta = 1$	$\prod_{i=0}^2 \exp\{-\theta_i \log(x_i/a)\}$	MDFR
MOBB	$\beta = 1$	$\prod_{i=0}^2 \exp\{-\theta_i \log(1+x_i^b)\}$	MDFR
MOBW	$\beta > 1 (\beta < 1)$	$\prod_{i=0}^2 \exp\{-\theta_i x_i^\beta\}$	MIFR(MDFR)

Note: $x_0 = \max(x_1, x_2)$

that T_j and $(\delta_{1j}, \delta_{2j})$ are independent. In this study, we are considering complete sample analysis, i.e., none of the observations are censored and no information is missing partially or completely.

3. Classical inference for competing risks model

In this section, two objectives have been fulfilled separately. In first, a simple parameter estimation has been done called as direct maximum likelihood estimation by considering the parameter β known as well as unknown. However, sometimes one may also be interested in the inference of other characteristics; those can be acquired by the combination or transformation of parameters. Hence, in the second attempt, the ML estimation has been done for re-parameterize set of parameters according to the statistic of keen interest.

3.1. Maximum likelihood estimation of MOBGLD

As discussed in Section 2.4, let n independent series systems with two causes of failure are considered in the experiment. The likelihood function can be formulated for mentioned bivariate competing risks observations $(T_j, \delta_{1j}, \delta_{2j}); j = 1, 2, \dots, n$ as follows [Lawless 2003]

$$L(\underline{t}; \beta, \Theta) = \prod_{j=1}^n \left[f(t_j, t_j) \right]^{\delta_{1j}\delta_{2j}} \left[-\frac{\partial S(x_{1j}, x_{2j})}{\partial x_1} \Big|_{(t_j, t_j)} \right]^{\delta_{1j}(1-\delta_{2j})} \left[-\frac{\partial S(x_{1j}, x_{2j})}{\partial x_2} \Big|_{(t_j, t_j)} \right]^{\delta_{2j}(1-\delta_{1j})}. \quad (3.1)$$

Then the likelihood Equation (3.1) for MOBGLD(β, Θ) given data $(T_j, \delta_{1j}, \delta_{2j}); j = 1, 2, \dots, n$ comes out to be[see Appendix A, from (A1)]

$$L(\underline{t}; \beta, \Theta) = \beta^n \theta_0^{n_0} \theta_1^{n_1} \theta_2^{n_2} \prod_{j=1}^n \left\{ g'(t_j) [g(t_j)]^{\beta-1} \right\} \exp \left\{ -\Delta \sum_{j=1}^n [g(t_j)]^\beta \right\}, \quad (3.2)$$

where n_0, n_1 and n_2 are such as

$$\begin{aligned} \sum_{j=1}^n I(X_{1j} = X_{2j}) &= \sum_{j=1}^n \delta_{1j}\delta_{2j} = n_0, \\ \sum_{j=1}^n I(X_{1j} < X_{2j}) &= \sum_{j=1}^n \delta_{1j}(1 - \delta_{2j}) = n_1, \\ \sum_{j=1}^n I(X_{1j} > X_{2j}) &= \sum_{j=1}^n \delta_{2j}(1 - \delta_{1j}) = n_2. \end{aligned}$$

Here, $n = n_0 + n_1 + n_2$, as it is assumed that no information is censored or missing in the experiment. Taking the logarithm of (3.2), the log-likelihood function comes out to be

$$l(\underline{t}; \beta, \Theta) \propto n \ln \beta + n_0 \ln \theta_0 + n_1 \ln \theta_1 + n_2 \ln \theta_2 + (\beta - 1) \sum_{j=1}^n \ln g(t_j) - \Delta \sum_{j=1}^n [g(t_j)]^\beta. \quad (3.3)$$

First order partial derivative of (3.3) with respect to β and $\theta_i; i = 0, 1, 2$ and equating to zero, we get

$$\hat{\beta} = \left[\frac{1}{n} \left(\Delta \sum_{j=1}^n [g(t_j)] \hat{\beta} \ln g(t_j) - \sum_{j=1}^n \ln g(t_j) \right) \right]^{-1}, \quad (3.4)$$

$$\hat{\theta}_i = \left[\frac{1}{n_i} \sum_{j=1}^n [g(t_j)] \hat{\beta} \right]^{-1}; i = 0, 1, 2. \quad (3.5)$$

Here, it can be seen that the ML estimators are not in closed form. So fixed point iteration (an iterative method) has been used to obtain the ML estimates.

Note: It can be seen from (3.5) that n_i should be greater than zero for existence of MLE's of $\theta_i; \forall (i = 0, 1, 2)$.

Let first calculate the expectation of second-order partial derivatives of the log-likelihood function (3.3), which are as follows (see *Appendix B*, by using (B1) and (B2)):

$$\begin{aligned} -E\left[\frac{\partial^2 l}{\partial \theta_i^2}\right] &= \frac{1}{\theta_i \Delta} n; i = 0, 1, 2. \\ -E\left[\frac{\partial^2 l}{\partial \beta \partial \theta_i}\right] &= \frac{n}{\Delta \beta} (r_1 + (1 - \ln \Delta)); i = 0, 1, 2. \\ -E\left[\frac{\partial^2 l}{\partial \beta^2}\right] &= \frac{n}{\beta^2} (r_2 + 2r_1(1 - \ln \Delta) - \ln \Delta(2 - \ln \Delta) + 1); \\ -E\left[\frac{\partial^2 l}{\partial \theta_i \partial \theta_k}\right] &= 0; i \neq k = 0, 1, 2, \text{ where, } r_u = \int_0^\infty (\ln z)^u \exp\{-z\} dz. \end{aligned}$$

It is noticeable that the GLD family can be bifurcated on the behalf of parameter β . If β is known, then we get exponential, Burr, Pareto and Rayleigh distribution, but for unknown β , Weibull distribution will be the particular case of GLD. So, accordingly, we can formulate the Fisher information matrix. Let consider two different cases:

Case (i): When β is known

The Fisher information matrix for $\Theta = (\theta_0, \theta_1, \theta_2)$ is given by

$$\Sigma_1 = \begin{bmatrix} \frac{n}{\Delta} \frac{1}{\theta_0} & 0 & 0 \\ 0 & \frac{n}{\Delta} \frac{1}{\theta_1} & 0 \\ 0 & 0 & \frac{n}{\Delta} \frac{1}{\theta_2} \end{bmatrix}. \quad (3.6)$$

Case (ii): When β is unknown

The Fisher information matrix for (β, Θ) is given by

$$\Sigma_2 = \begin{bmatrix} \frac{n}{\beta^2} K(\Delta) & \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] \\ \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & \frac{n}{\Delta} \frac{1}{\theta_0} & 0 & 0 \\ \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & 0 & \frac{n}{\Delta} \frac{1}{\theta_1} & 0 \\ \frac{n}{\Delta} \frac{1}{\beta} [r_1 + (1 - \ln \Delta)] & 0 & 0 & \frac{n}{\Delta} \frac{1}{\theta_2} \end{bmatrix}, \quad (3.7)$$

where $K(\Delta) = r_2 + 2r_1(1 - \ln \Delta) - \ln \Delta(2 - \ln \Delta) + 1$. Now a $100(1 - \alpha)\%$ asymptotic confidence interval for any function of parameter say, $h(\theta)$, can be obtained by

$$h(\hat{\theta}) \pm Z_{\alpha/2} \sqrt{Var(h(\hat{\theta}))},$$

where $h(\hat{\theta})$ and $Var(h(\hat{\theta}))$ is the ML estimate and variance of $h(\hat{\theta})$, respectively. The $Z_{\alpha/2}$ is the $(\alpha/2)^{th}$ quantile of the standard normal distribution with $0 < \alpha < 1$.

3.2. ML estimation under re-parametrization

In competing risks analysis, always a keen interest would be in the cause which is more severe (performing poorly) in respect to others and subsequently in the study of its characteristics. So, if parameters are re-parameterized then many more inferential objectives can be achieved. So, a particular set after re-parameterization has been considered, such as, say,

$$\lambda_1 = \Delta = \theta_0 + \theta_1 + \theta_2, \quad \lambda_2 = \frac{\theta_1}{\Delta}, \quad \lambda_3 = \frac{\theta_2}{\Delta}.$$

The re-parameterized quantities $\lambda_2 = P(X_1 < X_2)$ and $\lambda_3 = P(X_2 < X_1)$ represent the relative risk due to cause 1 and cause 2, respectively. The transformation $(\beta, \Lambda) = (\beta, \lambda_1, \lambda_2, \lambda_3)$ from $(\beta, \Theta) = (\beta, \theta_0, \theta_1, \theta_2)$ is one-to-one with the inverse transformation where

$$\beta = \beta, \quad \theta_1 = \lambda_1 \lambda_2, \quad \theta_2 = \lambda_1 \lambda_3, \quad \theta_0 = \lambda_1(1 - \lambda_2 - \lambda_3).$$

Now the likelihood function under re-parameterization is given by

$$L(\underline{t}; \beta, \Lambda) = \beta^n \lambda_1^n \lambda_2^{n_1} \lambda_3^{n_2} (1 - \lambda_2 - \lambda_3)^{n_0} \prod_{j=1}^n \left\{ g'(t_j) [g(t_j)]^{\beta-1} \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\}. \quad (3.8)$$

Taking the logarithmic of Equation (3.8) and differentiating with respect to unknown parameters and equating to zero, we have the ML estimators of (β, Λ) as follows:

$$\hat{\beta} = \left[\frac{1}{n} \left(\hat{\lambda}_1 \sum_{j=1}^n [g(t_j)] \hat{\beta} \ln g(t_j) - \sum_{j=1}^n \ln g(t_j) \right) \right]^{-1}, \quad (3.9)$$

$$\hat{\lambda}_1 = \left[\frac{1}{n} \sum_{j=1}^n [g(t_j)] \hat{\beta} \right]^{-1}, \quad (3.10)$$

$$\hat{\lambda}_2 = \frac{n_1}{n} \quad \text{and} \quad \hat{\lambda}_3 = \frac{n_2}{n}. \quad (3.11)$$

The ML estimators for parameters λ_2 and λ_3 are in closed form, but to obtain the ML estimate for β and λ_1 , an iterative method (i.e., fixed iteration) can be used. The inverse of Fisher information matrix can be obtained by utilizing matrix Σ and respective Jacobian transformations, such as

Case (i): When β is known

From [Appendix C: case \(i\)](#), the inverse of the Fisher information matrix for Λ is given by

$$\Sigma_{R_1}^{-1} = \begin{bmatrix} \frac{\lambda_1^2}{n} & 0 & 0 \\ 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}. \quad (3.12)$$

Case (ii): When β is unknown

From [Appendix C: case \(ii\)](#), the inverse of the Fisher information matrix for (β, Λ) is given by

$$\Sigma_{R_2}^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2-r_1^2)} & -\frac{\beta\lambda_1(r_1+1-\ln\lambda_1)}{n(r_2-r_1^2)} & 0 & 0 \\ -\frac{\beta\lambda_1(r_1+1-\ln\lambda_1)}{n(r_2-r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2-r_1^2)} & 0 & 0 \\ 0 & 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}. \quad (3.13)$$

The diagonal elements of the above matrices give the variances of the ML estimators of the parameter under consideration.

4. Bayesian inference for competing model

In Bayesian inference, the prior information regarding the parameter may be quite limited or not much credible sometimes. In such situations, it is expected that the observed data have enough freedom to express itself in the sense of information it contains. In such a case, considering a non informative prior becomes a better choice for Bayesian inference. One of the popular non informative priors is the reference prior, which is discussed with specific steps of the algorithm by Berger and Bernardo (1992). The reference prior is equivalent to Jeffrey's prior in the single parameter case and it overcomes the deficiencies of Jeffrey's prior in the multi parameter case. When some of the parameters have to be given more weight according to their inferential importance in the model, then it becomes inevitable to design the Bayesian inference accordingly. The reference prior is quite useful to fulfill this requirement.

As we discussed earlier, in competing risks analysis, we are interested in a cause wise inference. And many times we may have more concerns for a particular cause or a subset of causes presented in study. For example, in a clinical trial data is collected from cancer patients and the two possible causes of death of an individual may be cancer or any other factor. So in this study, we are focused on deaths due to cancer and related statistics. So parameters related to the cause distribution of cancer must be estimated

Table 3. Reference priors for parameters, Λ , of $MOBGLD(\Lambda)$ and for parameters, Θ , of $MOBGLD(\Theta)$.

Ordered grouping	Reference prior for Λ	Reference prior for Θ
$\{\lambda_1, (\lambda_2, \lambda_3)\}$	$\omega_{11}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{11}(\Theta)$
$\{(\lambda_2, \lambda_3), \lambda_1\}$	$\omega_{11}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{11}(\Theta)$
$\{\lambda_1, \lambda_2, \lambda_3\}$	$\omega_{12}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{12}(\Theta)$
$\{\lambda_1, \lambda_3, \lambda_2\}$	$\omega_{12}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{12}(\Theta)$
$\{\lambda_3, \lambda_2, \lambda_1\}$	$\omega_{13}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_3)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{13}(\Theta)$
$\{\lambda_2, \lambda_3, \lambda_1\}$	$\omega_{13}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_3)(1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{13}(\Theta)$

with highest precision. Even though if all causes are of equal importance, then in competing risks analysis, researchers are primarily interested toward those causes which are performing very poorly and lastly to those who are good enough. The reference prior based study provides such a facility to researchers. In reference priors set-up, we distribute the unknown parameters in different subsets. These sets are designed in an ordered manner by the expert according to the importance of sets(parameters) for the inferential purpose. So accordingly, parameters (i.e., causes) can be elected in different sets having more significance in respect to others [see Ghosh and Mukerjee 1992]. Earlier Guan, Tang and Xu (2013) discussed about the suitability of reference priors for the competing risks analysis.

In this article, some preferable choices of ordered grouping have been taken and respective reference priors are shown in Table 3 and Table 4 when the parameter β is considered known and unknown, respectively. These ordered groupings have been done on the basis of their inferential significance. The reference prior ω_{11} for known β and ω_{21} and ω_{24} for unknown β have been derived [see Appendix D]. Following the same procedure, one can also obtain other reference priors. A triplet indicator (l_1, l_2, l_3) has been used to represent all the considered reference priors and their respective posterior densities. On the basis of this triplet, a generalized form of reference prior as well as posterior density is given as follows:

Case (i): When β is known

The generalized form of reference prior for parameter Λ as derived in (D1) is given by

$$\xi(\Lambda) = \left[\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)^{l_1} (1 - \lambda_3)^{l_2} (1 - \lambda_2 - \lambda_3) k(\lambda_1)^{l_3} \right]^{-1}. \quad (4.1)$$

Hence for the expression of reference prior given in (4.1), we get the posterior density as follows:

$$\begin{aligned} p(\Lambda | t) &\propto \xi(\Lambda) L(t; \Lambda) \\ &\propto \lambda_1^{n-1} \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2)^{-\frac{l_1}{2}} (1 - \lambda_3)^{-\frac{l_2}{2}} k(\lambda_1)^{-\frac{l_3}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}} \\ &\quad \cdot \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^{\beta^*} \right\}, \end{aligned} \quad (4.2)$$

where β^* is a known value of β .

Case (ii): When β is unknown

The generalized form of reference prior for (β, Λ) as derived in (D2) and (D3) is given by

Table 4. Reference priors for parameters, (β, Λ) , of $MOBGLD(\beta, \Lambda)$ and for parameters, (β, Θ) , of $MOBGLD(\beta, \Theta)$.

Ordered grouping	Reference Prior for (β, Λ)	Reference Prior for (β, Θ)
$\{\beta, (\lambda_1, \lambda_2, \lambda_3)\}$	$\omega_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{21}(\beta, \Theta)$
$\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$	$\omega_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{21}(\beta, \Theta)$
$\{(\beta, \lambda_1), (\lambda_2, \lambda_3)\}$	$\omega_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{21}(\beta, \Theta)$
$\{(\lambda_2, \lambda_3), \lambda_1, \beta\}$	$\omega_{22}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3) k(\lambda_1)]^{-1/2}$	$\pi_{22}(\beta, \Theta)$
$\{\lambda_3, \lambda_2, \lambda_1, \beta\}$	$\omega_{23}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_3) (1 - \lambda_2 - \lambda_3) k(\lambda_1)]^{-1/2}$	$\pi_{23}(\beta, \Theta)$
$\{(\beta, \lambda_1, \lambda_2, \lambda_3)\}$	$\omega_{24}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2) (1 - \lambda_2 - \lambda_3)]^{-1/2}$	$\pi_{24}(\beta, \Theta)$

$$\xi(\beta, \Lambda) = \left[\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)^{l_1} (1 - \lambda_3)^{l_2} (1 - \lambda_2 - \lambda_3) k(\lambda_1)^{l_3} \right]^{-1}. \quad (4.3)$$

In respect to reference prior (4.3), the posterior density will be as follows:

$$\begin{aligned} p(\beta, \Lambda \mid \underline{t}) &\propto \xi(\beta, \Lambda) L(\underline{t}; \beta, \Lambda) \\ &\propto \beta^{n-1} \lambda_1^{n-1} \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2)^{-\frac{l_1}{2}} (1 - \lambda_3)^{-\frac{l_2}{2}} k(\lambda_1)^{-\frac{l_3}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}} \\ &\quad \cdot \prod_{j=1}^n \left\{ [g(t_j)]^{\beta-1} g'(t_j) \right\} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\}. \end{aligned} \quad (4.4)$$

The reference prior ' $\xi(\cdot)$ ' defined in (4.1) and (4.3) can take different form of reference priors according to triplet (l_1, l_2, l_3) where, $l_i = 0$ or 1 ; $i = 1, 2, 3$. All considered reference priors with corresponding triplet indicators are given in Table 5 for both the case when the parameter β is known and when it is unknown.

Theorem 4.1. *The posterior density $p(\Lambda \mid \underline{t})$ under the reference prior $\omega(\Lambda)$ is proper for $n_0 > 0, n_1 > 0, n_2 > 0$ with at least two failures.*

Proof. For a reference prior ω , we have posterior density, such as

$$p(\Lambda \mid \underline{t}) \propto \omega(\Lambda) L(\underline{t} \mid \Lambda)$$

To prove the posterior density proper, we have to show that it gives a finite value after integration, i.e., $\int p(\Lambda \mid \underline{t}) d\Lambda < \infty$. Let first prove for the posterior density $p_{11}(\Lambda \mid \underline{t})$ under reference prior $\omega_{11}(\Lambda)$. To find posterior density $p_{11}(\Lambda \mid \underline{t})$ consider the triplet $(0, 0, 0)$ in (4.2). Now we have to calculate

$$\begin{aligned} I_{11} &= \int_0^\infty \left[\iint_{0 < \lambda_2 + \lambda_3 < 1} \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}} d\lambda_2 d\lambda_3 \right] \\ &\quad \cdot \lambda_1^{n-1} \left(\prod_{i=1}^n g'(t_j) \right) e^{-\lambda_1 \sum_{i=1}^n g(t_j)} d\lambda_1. \end{aligned}$$

If $\mathcal{B}(., .)$ denotes the beta function, then we have

$$\mathcal{B}(a+1, b+2) \cdot \mathcal{B}(b+1, c+1) = \int_{0 < x+y < 1} x^a y^b (1-x-y)^c dx dy.$$

Now using this result, the above integral can be written, such as

Table 5. All reference priors for the parameter Λ when β is known and for (β, Λ) when β is unknown with corresponding triplets (l_1, l_2, l_3) .

Triplet (l_1, l_2, l_3)	β known		Triplet (l_1, l_2, l_3)	β unknown	
	Reference Prior (Λ)	Reference Prior (Θ)		Reference Prior (β, Λ)	Reference Prior (β, Θ)
(0, 0, 0)	$\omega_{11}(\Lambda)$	$\pi_{11}(\Theta)$	(0, 0, 0)	$\omega_{21}(\beta, \Lambda)$	$\pi_{21}(\beta, \Theta)$
(1, 0, 0)	$\omega_{12}(\Lambda)$	$\pi_{12}(\Theta)$	(0, 0, 1)	$\omega_{22}(\beta, \Lambda)$	$\pi_{22}(\beta, \Theta)$
(0, 1, 0)	$\omega_{13}(\Lambda)$	$\pi_{13}(\Theta)$	(0, 1, 1)	$\omega_{23}(\beta, \Lambda)$	$\pi_{23}(\beta, \Theta)$
			(1, 0, 0)	$\omega_{24}(\beta, \Lambda)$	$\pi_{24}(\beta, \Theta)$

$$\begin{aligned} I_{11} &= \mathcal{B}\left(n_1 + \frac{1}{2}, n_0 + n_2 + 1\right) \mathcal{B}\left(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}\right) \left(\prod_{i=1}^n g'(t_j) \right) \int_0^\infty \lambda_1^{n-1} e^{-\lambda_1} \sum_{i=1}^n g(t_j) d\lambda_1 \\ &= \mathcal{B}\left(n_1 + \frac{1}{2}, n_0 + n_2 + 1\right) \mathcal{B}\left(n_2 + \frac{1}{2}, n_0 + \frac{1}{2}\right) \left(\prod_{i=1}^n g'(t_j) \right) \frac{\Gamma(n)}{\left(\sum_{i=1}^n g(t_j)\right)^n} < \infty. \end{aligned}$$

Hence, the posterior density $p_{11}(\Lambda | t)$ is proper. Similar procedure can be used for other cases to prove them proper as well. \square

4.1. Posterior estimation

The posterior analysis of the parameters can be done by using the conditional posterior distributions for different reference priors based on values of triplet (l_1, l_2, l_3) . We use Gibbs sampling method, a Markov chain Monte Carlo (MCMC) technique Chen, Shao and Ibrahim (2000), to obtain the Bayesian estimator of the parameters. A step-by-step procedure is defined for obtaining the samples (estimates) for all parameters in both the cases and for all considered priors in Table 5.

Case (i): When β is known

The full conditional posterior in presence of known β are as follows:

$$p(\lambda_1 | \beta^*, \text{data}) \propto \lambda_1^{n-1} [k(\lambda_1)]^{-\frac{l_3}{2}} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^{\beta^*} \right\}, \quad (4.5)$$

$$p(\lambda_2, \lambda_3 | \text{data}) \propto \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2)^{-\frac{l_1}{2}} (1 - \lambda_2)^{-\frac{l_2}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}}. \quad (4.6)$$

- As in Table 5, we have three reference priors, ω_{11}, ω_{12} and ω_{13} with $l_3 = 0$. So, in all three priors, we get exact gamma distribution to generate a sample for λ_1 , i.e., $\lambda_1 \sim \mathcal{G}\left(n, \sum_{j=1}^n [g(t_j)]^{\beta^*}\right)$.
- It is easy to see that all three reference priors ω_{11}, ω_{12} and ω_{13} in Table 5 have different set of l_1 and l_2 . So, in respect to each reference prior a unique procedure needed to apply.
 - In case of reference prior ω_{11} , the joint posterior density follows Dirichlet distribution, i.e., $(\lambda_2, \lambda_3) \sim \text{Dir}(n_1 + 1/2, n_2 + 1/2, n_0 + 1/2)$. Now if a rv $Y_i \sim \mathcal{G}(\alpha_i, 1); i = 1, 2, \dots, K$, where $\mathcal{G}(\cdot, \cdot)$ stands for gamma density and assume that $X_i = Y_i / \sum_{i=1}^K Y_i$, then $(X_1, X_2, \dots, X_{K-1}) \sim \text{Dir}(\alpha_1, \alpha_2, \dots, \alpha_K)$. Then it is

obvious to write that $\lambda_2 = \frac{y_1}{y_1+y_2+y_3}$; $\lambda_3 = \frac{y_2}{y_1+y_2+y_3}$, where $y_1 \sim \mathcal{G}(n_1 + 1/2, 1)$, $y_2 \sim \mathcal{G}(n_2 + 1/2, 1)$ and $y_3 \sim \mathcal{G}(n_0 + 1/2, 1)$. It is quite easy to generate random number from y_1 , y_2 and y_3 and so for λ_2 and λ_3 .

- In case of the reference priors ω_{12} , another transformation procedure can be applied. Let assume that $z_1 = \lambda_2$ and $z_2 = \lambda_3/(1 - \lambda_2)$. Then it can be directly obtained from (4.2) that z_1 and z_2 are independent as well as both are following beta distribution, i.e., $Z_1 \sim \mathcal{B}(n_1 + 1/2, n_0 + n_2 + 0.5)$ and $Z_2 \sim \mathcal{B}(n_2 + 0.5, n_0 + n_2 + 0.5)$. So a sample can be easily generated from beta distribution, so as from λ_2 and λ_3 .
- In case of ω_{13} , assume that $z_1 = \lambda_3$ and $z_2 = \lambda_2/(1 - \lambda_3)$, then by the similar procedure as for ω_{12} , samples can be obtained for λ_2 and λ_3 .

Case (ii): When β is unknown

The full conditional of the parameters can be obtained from (4.4) as follows:

$$p(\beta|\lambda_1, data) \propto \beta^{n-1} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\} \prod_{j=1}^n [g(t_j)]^{\beta-1}, \quad (4.7)$$

$$p(\lambda_1|\beta, data) \propto \lambda_1^{n-1} [k(\lambda_1)]^{-\frac{l_3}{2}} \exp \left\{ -\lambda_1 \sum_{j=1}^n [g(t_j)]^\beta \right\}, \quad (4.8)$$

$$p(\lambda_2, \lambda_3|data) \propto \lambda_2^{n_1-\frac{1}{2}} \lambda_3^{n_2-\frac{1}{2}} (1 - \lambda_2)^{-\frac{l_1}{2}} (1 - \lambda_2)^{-\frac{l_2}{2}} (1 - \lambda_2 - \lambda_3)^{n_0-\frac{1}{2}}. \quad (4.9)$$

- To obtain a posterior sample for β from (4.7), we use the Metropolis-Hastings (M-H) algorithm by considering normal distribution as a proposal density, such as the candidate point $\beta^c \sim \mathcal{N}(\hat{\beta}, Var(\hat{\beta}))$.
- As in Table 5, we consider four different reference priors, $\omega_{21}, \omega_{22}, \omega_{23}$ and ω_{24} , for unknown β . As for reference priors ω_{21} and ω_{24} , we have $l_3 = 0$ and hence a sample can be generated directly for λ_1 from $\mathcal{G}(n, \sum_{j=1}^n [g(t_j)]^\beta)$. But for the reference prior ω_{22} and ω_{23} , where $l_3 = 1$, we use M-H algorithm using gamma distribution as a proposal density such as candidate point $\lambda_1^c \sim \mathcal{G}(n, \sum_{j=1}^n [g(t_j)]^\beta)$. Here, it is to be noted that (β, λ_1) be a vector of parameters for which MCMC induced by the Gibbs sampler. First, we choose an arbitrary starting point (β^0, λ_1^0) and then proceed accordingly. So, as we collect the sample from both, it is easy to obtain the Bayes estimates.
- As no expression of β is presented in the joint posterior distribution of (λ_2, λ_3) hence the same steps will be followed in *case (ii)* as discussed earlier in *case (i)*.

The procedure defined above provides a general form of full conditional posteriors with respect to MOBGLD. For any particular member of the family, the same can be drawn by setting the value of $g(t)$ and the parameters accordingly.

5. Simulation study

In this section, simulation studies have been performed for the maximum likelihood estimators and the Bayesian estimators under different reference priors of the unknown

Table 6. The AB, RMSE and CP of the estimators for MOBED (0.87, 1.30, 1.50).

n	MLE			ω_{11}			ω_{12}			ω_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.7105	0.9412	0.9280	0.7128	0.9432	0.9390	0.7101	0.9403	0.9410	0.7108	0.9420	0.9410
	λ_2	0.0835	0.1047	0.9844	0.0779	0.0974	0.9420	0.0794	0.1001	0.9450	0.0789	0.0974	0.9420
	λ_3	0.0841	0.1044	0.9828	0.0790	0.0973	0.9420	0.0800	0.0980	0.9510	0.0795	0.0994	0.9530
30	λ_1	0.5769	0.7420	0.9330	0.5773	0.7428	0.9440	0.5777	0.7431	0.9460	0.5776	0.7439	0.9450
	λ_2	0.0718	0.0800	0.9350	0.0684	0.0857	0.9550	0.0695	0.0873	0.9370	0.0683	0.0858	0.9370
	λ_3	0.0726	0.0911	0.9590	0.0695	0.0867	0.9360	0.0700	0.0871	0.9450	0.0700	0.0883	0.9390
40	λ_1	0.4648	0.6054	0.9410	0.4643	0.6056	0.9540	0.4651	0.6048	0.9540	0.4646	0.6051	0.9570
	λ_2	0.0624	0.0783	0.9420	0.0602	0.0755	0.9420	0.0608	0.0766	0.9380	0.0605	0.0755	0.9400
	λ_3	0.0662	0.0819	0.9320	0.0638	0.0789	0.9320	0.0639	0.0791	0.9320	0.0646	0.0800	0.9320
50	λ_1	0.4157	0.5360	0.9500	0.4161	0.5358	0.9470	0.4156	0.5354	0.9520	0.4157	0.5357	0.9520
	λ_2	0.0557	0.0695	0.9370	0.0540	0.0675	0.9370	0.0546	0.0682	0.9370	0.0539	0.0675	0.9370
	λ_3	0.0563	0.0701	0.9580	0.0547	0.0681	0.9570	0.0547	0.0682	0.9430	0.0552	0.0687	0.9520

Table 7. The AB, RMSE and CP of the estimators for MOBRD (0.33, 0.66, 0.76)..

n	MLE			ω_{11}			ω_{12}			ω_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.3254	0.4273	0.9290	0.3257	0.4273	0.9520	0.3256	0.4276	0.9540	0.3255	0.4279	0.9500
	λ_2	0.0856	0.1064	0.9420	0.0794	0.0989	0.9660	0.0817	0.1018	0.9470	0.0792	0.0992	0.9570
	λ_3	0.0887	0.1092	0.9640	0.0824	0.1020	0.9640	0.0829	0.1031	0.9580	0.0845	0.1041	0.9550
30	λ_1	0.2623	0.3404	0.9370	0.2623	0.3400	0.9470	0.2618	0.3402	0.9460	0.2624	0.3401	0.9460
	λ_2	0.0719	0.0901	0.9380	0.0687	0.0859	0.9380	0.0693	0.0873	0.9390	0.0693	0.0864	0.9460
	λ_3	0.0731	0.0924	0.9590	0.0704	0.0881	0.9480	0.0713	0.0887	0.9330	0.0710	0.0894	0.9440
40	λ_1	0.2198	0.2834	0.9460	0.2201	0.2840	0.9530	0.2196	0.2832	0.9550	0.2199	0.2838	0.9520
	λ_2	0.0587	0.0734	0.9610	0.0567	0.0707	0.9610	0.0574	0.0718	0.9600	0.0570	0.0707	0.9600
	λ_3	0.0613	0.0754	0.9640	0.0593	0.0729	0.9640	0.0595	0.0733	0.9640	0.0597	0.0737	0.9630
50	λ_1	0.1973	0.2552	0.9410	0.1970	0.2551	0.9530	0.1975	0.2557	0.9520	0.1979	0.2559	0.9560
	λ_2	0.0552	0.0686	0.9380	0.0535	0.0666	0.9380	0.0544	0.0674	0.9410	0.0536	0.0666	0.9440
	λ_3	0.0590	0.0731	0.9500	0.0573	0.0711	0.9280	0.0573	0.0713	0.9330	0.0581	0.0718	0.9410

parameters. The performance of an estimator has been shown using absolute bias (AB), root mean square error (RMSE) and coverage probability (CP).

Under simulation study, we take sample size $n = 20, 30, 40$ and 50 to evaluate and compare the efficiency of the estimators derived in Sections 3.2 and 4. All the simulation studies are repeated 1000 times to achieve consistency in the results. We have applied the given set-up for all five particular cases of MOBGLD.

The results are shown for MOBED(0.87, 1.30, 1.50), MOBRD(0.33, 0.66, 0.76) and MOBD(1.15, 1.65, 1.71) when $a = 0.50$, MOBBD(1.25, 1.64, 1.72) when $b = 1.50$ and MOBWD(0.75, 0.55, 1.25, 1.35) in Table 6–10, respectively.

The important establishments from the simulated results are as follows:

- All the estimators are performing well for all the members of MOBGLD. Also, as the sample size is increasing, the AB's and RMSE's of estimators are gradually decreasing.
- Here, it is to be noted that no single reference prior is best for all the parameters of interest. It is because the formulation of reference priors is solely based on the inferential interest of the researcher. For example, it can be seen from the simulation table of MOBED, ω_{13} performs better for λ_1 for larger n and ω_{11}, ω_{12} and ω_{13} all perform well for λ_2 and λ_3 .

Table 8. The AB, RMSE and CP of the estimators for MOBPD (1.15, 1.65, 1.71) when $a = 0.5$.

n	MLE			ω_{11}			ω_{12}			ω_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.8693	1.1704	0.9100	0.8690	1.1702	0.9460	0.8703	1.1702	0.9450	0.8686	1.1694	0.9430
	λ_2	0.0910	0.1110	0.9340	0.0848	0.1033	0.9340	0.0859	0.1057	0.9350	0.0853	0.1035	0.9470
	λ_3	0.0857	0.1058	0.9390	0.0796	0.0983	0.9650	0.0793	0.0985	0.9530	0.0817	0.1010	0.9480
30	λ_1	0.6975	0.9229	0.9370	0.6966	0.9216	0.9510	0.6985	0.9249	0.9500	0.6981	0.9240	0.9510
	λ_2	0.0663	0.0836	0.9730	0.0632	0.0796	0.9490	0.0647	0.0808	0.9570	0.0636	0.0796	0.9510
	λ_3	0.0685	0.0848	0.9530	0.0653	0.0808	0.9530	0.0653	0.0808	0.9620	0.0662	0.0823	0.9550
40	λ_1	0.6005	0.7854	0.9310	0.5999	0.7851	0.9510	0.6011	0.7854	0.9510	0.6008	0.7858	0.9460
	λ_2	0.0612	0.0752	0.9640	0.0590	0.0725	0.9540	0.0597	0.0734	0.9580	0.0589	0.0727	0.9470
	λ_3	0.0610	0.0733	0.9500	0.0590	0.0745	0.9500	0.0592	0.0744	0.9490	0.0594	0.0757	0.9490
50	λ_1	0.5225	0.6834	0.9320	0.5219	0.6836	0.9430	0.5245	0.6852	0.9490	0.5231	0.6844	0.9440
	λ_2	0.0536	0.0673	0.9400	0.0521	0.0653	0.9580	0.0523	0.0659	0.9450	0.0523	0.0655	0.9470
	λ_3	0.0535	0.0680	0.9370	0.0519	0.0660	0.9410	0.0520	0.0658	0.9410	0.0529	0.0669	0.9390

Table 9. The AB, RMSE and CP of the estimators for MOBBD (1.25, 1.64, 1.72) when $b = 1.5$.

n	MLE			ω_{11}			ω_{12}			ω_{13}			
	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	AB	RMSE	CP	
20	λ_1	0.8742	1.2158	0.9360	0.8751	1.2163	0.9470	0.8737	1.2165	0.9470	0.8741	1.2143	0.9470
	λ_2	0.0836	0.1057	0.9740	0.0781	0.0983	0.9450	0.0793	0.1008	0.9480	0.0790	0.0983	0.9460
	λ_3	0.0851	0.1047	0.9540	0.0793	0.0974	0.9640	0.0793	0.0978	0.9670	0.0810	0.1001	0.9580
30	λ_1	0.7095	0.9162	0.9370	0.7103	0.9181	0.9410	0.7101	0.9164	0.9350	0.7097	0.9156	0.9390
	λ_2	0.0692	0.0868	0.9510	0.0658	0.0827	0.9630	0.0671	0.0842	0.9510	0.0656	0.0827	0.9480
	λ_3	0.0711	0.0885	0.9420	0.0679	0.0843	0.9420	0.0682	0.0844	0.9460	0.0686	0.0858	0.9440
40	λ_1	0.5930	0.7596	0.9420	0.5927	0.7589	0.9450	0.5921	0.7579	0.9440	0.5929	0.7597	0.9440
	λ_2	0.0610	0.0747	0.9560	0.0589	0.0720	0.9560	0.0593	0.0729	0.9540	0.0590	0.0721	0.9510
	λ_3	0.0599	0.0756	0.9520	0.0575	0.0728	0.9490	0.0578	0.0728	0.9470	0.0588	0.0739	0.9440
50	λ_1	0.5314	0.7006	0.9350	0.5312	0.7015	0.9490	0.5323	0.7016	0.9460	0.5319	0.7003	0.9450
	λ_2	0.0526	0.0658	0.9510	0.0510	0.0628	0.9510	0.0516	0.0645	0.9500	0.0510	0.0640	0.9550
	λ_3	0.0563	0.0702	0.9550	0.0546	0.0681	0.9320	0.0544	0.0681	0.9350	0.0553	0.0689	0.9360

- For all members of MOBGLD except MOBED, all priors perform better than MLE of β, λ_2 and λ_3 with respect to their ABs and RMSEs. For MOBWD, all reference priors under-perform than MLE for λ_1 .
- For MOBPD and MOBRD, ω_{13} performs satisfactorily than MLE, and for MOBBD, ω_{11} performed more desirable than that of MLE based on their ABs and RMSEs. Also, the coverage probability of the Bayesian estimators is close to the nominal level.

The simulation study can be concluded that all estimators are performing good in both estimation approaches. It can also be remarked that the Bayesian estimators based on reference priors resulting more effective than the classical estimator with respect to their AB, RMSE and coverage probability for different sample sizes.

6. Real data analysis

In this section, two data sets are analyzed to illustrate the usefulness of the models we proposed in the earlier section. The first data is of prostate cancer, whereas the second data is of diabetic retinopathy study. Assuming the cause dependent competing risks set-up is suitable, the analysis of both data sets has been done using MOBGLD.

Table 10. The AB, RMSE and CP of the estimators for MOBWD when $(\beta, \theta_0, \theta_1, \theta_2) = (0.75, 0.55, 1.25, 1.35)$.

n		MLE						α_{21}						α_{22}						α_{23}						α_{24}							
		AB			RMSE			CP			AB			RMSE			CP			AB			RMSE			CP							
		β	λ_1	λ_2	λ_3	β	λ_1	λ_2	λ_3	β	λ_1	λ_2	λ_3	β	λ_1	λ_2	λ_3	β	λ_1	λ_2	λ_3	β	λ_1	λ_2	λ_3	β	λ_1	λ_2	λ_3				
20	β	0.1203	0.1695	0.8670	0.1174	0.1612	0.09080	0.1174	0.1613	0.9170	0.1175	0.1615	0.9170	0.1175	0.1613	0.1613	0.8427	λ_1	0.9569	1.8194	0.9660	0.9774	1.9253	0.9380	0.9898	2.0136	0.9310	0.9779	1.9728	0.9420	0.9726	1.8892	0.9150
	λ_1	0.0831	0.1066	0.9410	0.0828	0.1047	0.09910	0.0831	0.1051	0.9920	0.0794	0.1006	0.9460	0.0798	0.1015	0.1015	0.9580																
	λ_2	0.0867	0.1088	0.9440	0.0845	0.1069	0.09910	0.0845	0.1065	0.9930	0.0824	0.1035	0.9490	0.0812	0.1031	0.1031	0.9480																
	λ_3	0.0987	0.1401	0.9050	0.0935	0.1250	0.09210	0.0935	0.0156	0.9190	0.0936	0.1248	0.9160	0.0943	0.1260	0.1260	0.8581																
	β	0.6533	0.0603	0.9770	0.6478	1.0392	0.9440	0.6476	1.0817	0.9440	0.6495	1.0458	0.9410	0.6554	1.0553	0.9310	0.9310																
	λ_1	0.0678	0.0848	0.9690	0.0677	0.0842	0.9990	0.0682	0.0071	0.9980	0.0654	0.0813	0.9640	0.0660	0.0822	0.0822	0.9640																
40	λ_2	0.0678	0.0855	0.9360	0.0679	0.0850	0.9930	0.0674	0.0071	0.9930	0.0660	0.0829	0.9610	0.0653	0.0821	0.0821	0.9540																
	λ_3	0.0759	0.1015	0.9130	0.0746	0.0941	0.9270	0.0745	0.0940	0.9330	0.0745	0.0941	0.9270	0.0748	0.0946	0.0946	0.8840																
	β	0.4885	0.6587	0.9500	0.4922	0.6631	0.9510	0.4902	0.6619	0.9550	0.4901	0.6608	0.9560	0.4943	0.6669	0.9340	0.9340																
	λ_1	0.0577	0.0745	0.9440	0.0579	0.0742	0.9920	0.0578	0.0741	0.9920	0.0560	0.0723	0.9430	0.0567	0.0729	0.0729	0.9460																
	λ_2	0.0584	0.0744	0.9560	0.0586	0.0744	0.9950	0.0587	0.0747	0.9950	0.0571	0.0727	0.9430	0.0569	0.0722	0.0722	0.9490																
	λ_3	0.0703	0.0958	0.9310	0.0688	0.0880	0.9250	0.0689	0.0882	0.9180	0.0686	0.0879	0.9230	0.0689	0.0883	0.0883	0.8740																
50	β	0.4343	0.5815	0.9570	0.4334	0.5806	0.9510	0.4346	0.5816	0.9510	0.4350	0.5820	0.9470	0.4370	0.5845	0.9400	0.9400																
	λ_1	0.0544	0.0670	0.9550	0.0546	0.0671	0.9950	0.0543	0.0668	0.9950	0.0529	0.0654	0.9600	0.0535	0.0657	0.0657	0.9640																
	λ_2	0.0536	0.0692	0.9420	0.0553	0.0690	0.9950	0.0549	0.0687	0.9960	0.0545	0.0678	0.9510	0.0541	0.0675	0.0675	0.9480																

Table 11. Patients summary statistics of prostate cancer data.

Cause of Death	n	Minimum	First Quartile	Median	Mean	Third Quartile	Maximum	SD
Prostate cancer	127	2.00	12.00	24.00	26.26	37.00	71.00	17.35
Cerebrovascular	31	3.00	21.00	33.00	31.32	42.00	60.00	15.14
Other causes	24	1.00	8.25	20.00	25.75	35.50	74.00	22.79

6.1. Prostate cancer data

The prostate cancer data set originally consists of 502 patients, reported by Byar and Green (1980). We only use the major causes of death related to patients: prostate cancer (127 patients), cerebrovascular accident (31 patients) and other causes (24 patient), a sum of 182 patients [see Table 11]. To make the data suitable to our model derived in Section 3, we initially grouped the patients as per their cause of death. The first group consists of those patients whose death is caused by prostate cancer or other causes; and the second group consists of those patients whose death is caused by cerebrovascular or other causes. One of the advantages of grouping the patients in such a way is that one can compare the risk of death from all causes except cerebrovascular and with all causes except prostate cancer.

For j^{th} patient, let X_{1j} and X_{2j} be the latent time of death (in months) due to group 1 and group 2 of causes, respectively. So that $T_j = \min(X_{1j}, X_{2j})$ be the observed death time of the j^{th} patient. Hence, if the patient is died of prostate cancer, then only X_{1j} is observed, if the patient is died of cerebrovascular accident, then only X_{2j} is observed and if the patient is died of other causes, then X_{1j} and X_{2j} both are observed simultaneously.

In this data set, time of death for patients is recorded in months. We analyze the data with failure times in years by dividing it by 12. Now, before performing any inferential analysis, we first check the nature of its hazard rate by using the total time on test (TTT) plot. It has been shown by Aarset (1987) that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing). The resulting TTT Figure 1 suggests to fit an increasing hazard rate distribution for this data. So we fit MOBWD and MOBRD distributions for the data and obtained the AIC and BIC values viz. 926.8753 and 939.6913 for MOBWD and 950.2248 and 959.8368 for MOBRD, respectively. Hence, we fit MOBWD as its AIC and BIC are minimum and estimates of the parameters of MOBWD are given in Table 12.

6.2. Diabetic retinopathy study data

Diabetic Retinopathy study data belongs to a case study conducted by the National Eye Institute. It was used to assess the effect of laser treatment in delaying the onset of blindness in patients with diabetic retinopathy. For more details about this data set given in Table 13, one can see Xu and Zhou (2017). The data set contains observation on X_1 , time to blindness of the eye under laser treatment and X_2 , time to blindness of the eye with no treatment. That is, we have minimum of time to blindness $T = \min(X_1, X_2)$ whereas there is a possibility that $X_1 = X_2$. In the data set, time of failure is given in days. We analyze the data with failure times in years by dividing it by 365. Now before performing any inferential analysis, we first check the nature of its hazard

Total time on test plot

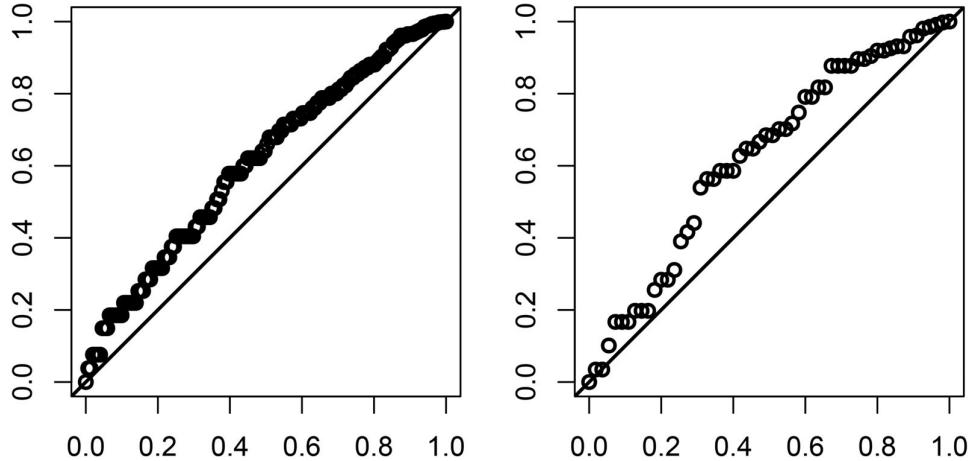


Figure 1. TTT plot for X_1 (time to death of patients in group 1) and X_2 (time to death of patients in group 2).

Table 12. The ML and Bayes estimates (95% CI) of parameters for MOBWD of prostate cancer data.

Method		β (95% CI)	λ_1 (95% CI)	λ_2 (95% CI)	λ_3 (95% CI)
MLE		1.5085 (1.3434, 1.6736)	0.2519 (0.1907, 0.3131)	0.6978 (0.6311, 0.7645)	0.1703 (0.1157, 0.2249)
Bayes	Prior				
	ω_{21}	1.5062 (1.3708, 1.6484)	0.2533 (0.1907, 0.3204)	0.6964 (0.5526, 0.8603)	0.1728 (0.1068, 0.2413)
	ω_{22}	1.5015 (1.3558, 1.6309)	0.2551 (0.1963, 0.3193)	0.6961 (0.5457, 0.8700)	0.1730 (0.1145, 0.2398)
	ω_{23}	1.5087 (1.3707, 1.6408)	0.2523 (0.1877, 0.3229)	0.6952 (0.6243, 0.7558)	0.1720 (0.1185, 0.2400)
	ω_{24}	1.5096 (1.3710, 1.6350)	0.2527 (0.2015, 0.3090)	0.6944 (0.6254, 0.7658)	0.1718 (0.1195, 0.2234)

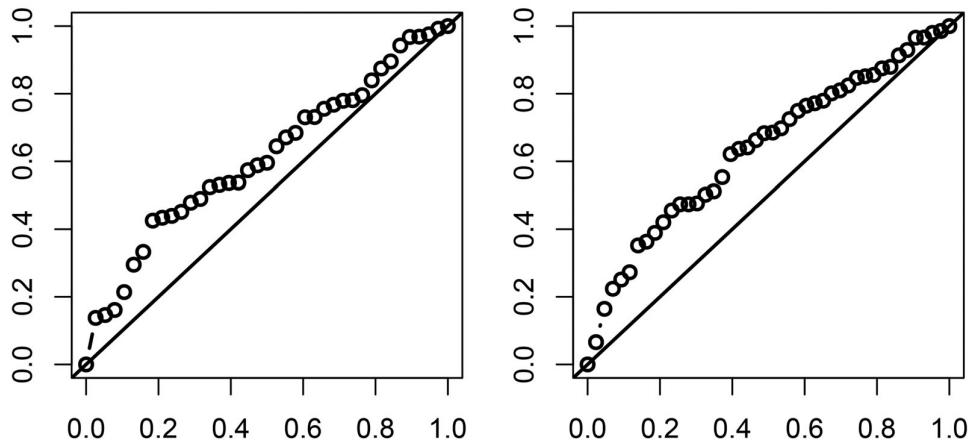
using the total time on test (TTT) plot. It has been shown by Aarset (1987) that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing). The resulting TTT Figure 2 suggest to fit a increasing hazard rate distribution. So, we fit MOBWD and MOBRD distributions to this data and obtained AIC and BIC values viz. 326.235 and 335.286 for MOBWD and 332.848 and 339.636 for MOBRD, respectively. Hence, we fit MOBWD as its AIC and BIC are minimum and estimates of the parameters of MOBWD are given in Table 14.

The dependency between the two competing risks is governed by the parameter θ_0 . So that we can test $H_0 : \theta_0 = 0$ against $H_1 : \theta_0 > 0$ by using the asymptotic normality of ML estimator of θ_0 . The test statistic in this case is $t^* = \hat{\theta}_0 / S.E.(\hat{\theta}_0)$. The value of test statistic comes out to be 2.9576, along with p-value 0.00155. So the independent assumption between the two risks is rejected in light of evidence obtained from observed data at $\alpha = 0.05$ level of significance.

Table 13. Time to blindness in days and causes for 71 patients.

Cause	Time
Eye under Laser Treatment	266, 583, 79, 93, 805, 344, 306, 415, 178, 1484, 315, 1252, 642, 407, 356, 699, 667, 126, 350, 84, 392, 901, 276, 520, 503, 584, 355, 1302
Eye under No Treatment	91, 154, 547, 707, 469, 1313, 790, 125, 777, 307, 637, 577, 517, 287, 717,
Both Eyes	141, 427, 36, 588, 350, 567, 1140, 448, 904, 485, 248, 423, 285, 315, 727, 210, 409, 227, 285, 622, 272, 1137, 1653, 471, 663, 966, 203, 1247

Total time on test plot

**Figure 2.** TTT plot for time to blindness of eye under laser treatment observations, X_1 , and time to blindness of the eye with no treatment observations, X_2 .

In DRS study, one may be further interested in whether the diabetic patient group differs significantly by laser treatment, i.e., one wish to test $H_0 : \lambda_2 = \lambda_3$ against $H_1 : \lambda_2 \neq \lambda_3$ which is equivalent to the testing of symmetry $H_0 : \theta_1 = \theta_2$ against $H_1 : \theta_1 \neq \theta_2$.

The test statistic in this case is given by Hanagal and Kale (1992)

$$T^* = \frac{n(\hat{\theta}_1 - \hat{\theta}_2)^2}{\hat{I}^{22} + \hat{I}^{33} - 2\hat{I}^{23}},$$

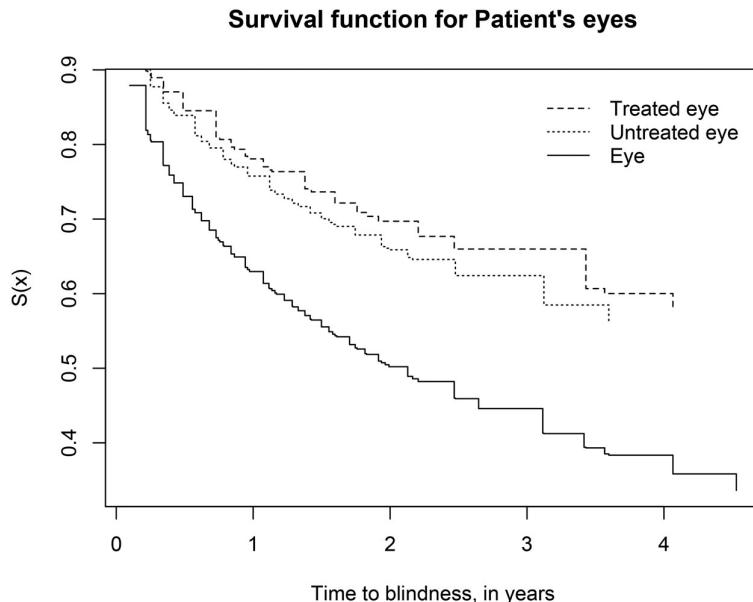
where \hat{I}^{22} , \hat{I}^{33} and \hat{I}^{23} are the elements of inverse of the Fisher information matrix Σ_2 .

The null distribution of T^* is χ^2 with 1 degrees of freedom, i.e., $T^* \sim \chi^2_{(1)}$. The value of test statistic T^* comes out to be 28.9286 along with p-value < 0.00001, so the claim that the two treatment groups differ significantly cannot be rejected.

As given by Table 14, the probability of failure of the eye receiving laser treatment is less than that of the eye with no treatment, i.e., $\lambda_2 < \lambda_3$, we may conclude that laser treatment is quite significant in delaying the onset of blindness. Moreover, from Figure 3, it is clear that the survival probability of the eye receiving laser treatment is greater than that of the eye with no treatment throughout the time. Hence, the survival is more favorable to the eye receiving laser treatment. Since one may be interested in the time to blindness in the eye receiving laser treatment, i.e., in the parameter λ_3 , so for

Table 14. The ML and Bayes estimates(95% CI) of parameters for MOBWD of Diabetic Retinopathy study data.

Method		β (95% CI)	λ_1 (95% CI)	λ_2 (95% CI)	λ_3 (95% CI)
MLE		1.5582 (1.2535, 1.8629)	0.4691 (0.3141, 0.6240)	0.3945 (0.2807, 0.5080)	0.4648 (0.3488, 0.5808)
Bayes	Prior				
	ω_{21}	1.5532 (1.3226, 1.8107)	0.4741 (0.3368, 0.6536)	0.3980 (0.2483, 0.5819)	0.4640 (0.3054, 0.6810)
	ω_{22}	1.5514 (1.2973, 1.8246)	0.4677 (0.3268, 0.6213)	0.4010 (0.2452, 0.5820)	0.4719 (0.3055, 0.7138)
	ω_{23}	1.5527 (1.2964, 1.8011)	0.4778 (0.3416, 0.6492)	0.3891 (0.2871, 0.4948)	0.4622 (0.3519, 0.5635)
	ω_{24}	1.5549 (1.3608, 1.7551)	0.4653 (0.3415, 0.6086)	0.3950 (0.2849, 0.5033)	0.4611 (0.3467, 0.5674)

**Figure 3.** Survival function estimate for DRS data.

Bayesian inference of the unknown parameters, one can choose the reference prior ω_{23} , which gives more weight to the parameter λ_3 .

7. Conclusion

In this article, we have analyzed the dependent competing risks data using Marshall-Olkin set-up for a generalized family of lifetime distribution. We also obtained the bivariate version of the Marshall-Olkin distribution for each member of the family. Bayesian estimation was done using reference prior, an objective Bayesian analysis and classical estimators are compared with their Bayesian counterparts in terms of their absolute bias, root mean square error and coverage probability. Based on these measures, we can choose the suitable reference prior for our parameter of inferential interest.

Although in this paper, the work is done for MOBGLD competing risks data in case of complete sample, it can be further extended for competing risks data under different censoring schemes, such as Type-I censoring, Type-II censoring, random censoring, progressively hybrid censoring, etc and also for incomplete data problem.

Note

1. **MOBE:** Marshall-Olkin Bivariate Exponential; **MOBR:** Marshall-Olkin Bivariate Rayleigh; **MOBP:** Marshall-Olkin Bivariate Pareto; **MOBB:** Marshall-Olkin Bivariate Burr; **MOBW:** Marshall-Olkin Bivariate Weibull.

ORCID

M. S Panwar  <http://orcid.org/0000-0002-3535-2327>

References

- Aarset, M. V. 1987. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability* R-36 (1):106–8. doi:[10.1109/TR.1987.5222310](https://doi.org/10.1109/TR.1987.5222310).
- Chaturvedi, A., and G. S. Khumukcham. 2006. Bayesian estimation procedures for a family of lifetime distributions under squared-error and entropy losses. *Metron* 64 (2):179–98.
- Barlow, R. E., and F. Proschan. 1975. Statistical theory of reliability and life testing: Probability models. Technical report, Florida State Univ., Tallahassee.
- Basu, A., and J. Ghosh. 1978. Identifiability of the multinormal and other distributions under competing risks model. *Journal of Multivariate Analysis* 8 (3):413–29. doi:[10.1016/0047-259X\(78\)90064-7](https://doi.org/10.1016/0047-259X(78)90064-7).
- Berger, J. O., and J. M. Bernardo. 1992. On the development of the reference prior method. *Bayesian Statistics* 4 (4):35–60.
- Byar, D. P., and S. B. Green. 1980. The choice of treatment for cancer patients based on covariate information. *Bulletin du Cancer* 67 (4):477–90.
- Cai, J., Y. Shi, and B. Liu. 2017. Analysis of incomplete data in the presence of dependent competing risks from marshall-olkin bivariate weibull distribution under progressive hybrid censoring. *Communications in Statistics - Theory and Methods* 46 (13):6497–511. doi:[10.1080/03610926.2015.1129420](https://doi.org/10.1080/03610926.2015.1129420).
- Chen, M. H. Q. M. Shao, and J. G. Ibrahim. 2000. *Monte Carlo methods in Bayesian computation*. New York: Springer-Verlag.
- Crowder, M. 1991. On the identifiability crisis in competing risks analysis. *Scandinavian Journal of Statistics* 18 (3):223–33.
- Feizjavadian, S., and R. Hashemi. 2015. Analysis of dependent competing risks in the presence of progressive hybrid censoring using Marshall–Olkin bivariate Weibull distribution. *Computational Statistics & Data Analysis* 82:19–34. doi:[10.1016/j.csda.2014.08.002](https://doi.org/10.1016/j.csda.2014.08.002).
- Ghosh, J., and R. Mukerjee. 1992. Noninformative priors (with discussion). In *Bayesian statistics*, Vol. 4, ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford: Oxford University Press, 195–210.
- Guan, Q., Y. Tang, and A. Xu. 2013. Objective bayesian analysis for bivariate Marshall–Olkin exponential distribution. *Computational Statistics & Data Analysis* 64:299–313. doi:[10.1016/j.csda.2013.03.021](https://doi.org/10.1016/j.csda.2013.03.021).
- Hanagal, D. D., and B. Kale. 1992. Large sample tests for testing symmetry and independence in some bivariate exponential models. *Communications in Statistics - Theory and Methods* 21 (9): 2625–43. doi:[10.1080/03610929208830934](https://doi.org/10.1080/03610929208830934).
- Kundu, D., and A. K. Gupta. 2013. Bayes estimation for the Marshall–Olkin bivariate Weibull distribution. *Computational Statistics & Data Analysis* 57 (1):271–81. doi:[10.1016/j.csda.2012.06.002](https://doi.org/10.1016/j.csda.2012.06.002).
- Kundu, D., N. Kannan, and N. Balakrishnan. 2003. Analysis of progressively censored competing risks data. *Handbook of Statistics* 23:331–48. doi:[10.1016/S0169-7161\(03\)23018-2](https://doi.org/10.1016/S0169-7161(03)23018-2).

- Kuo, L., and T. Y. Yang. 2000. Bayesian reliability modeling for masked system lifetime data. *Statistics & Probability Letters* 47 (3):229–41. doi:[10.1016/S0167-7152\(99\)00160-1](https://doi.org/10.1016/S0167-7152(99)00160-1).
- Lawless, J. F. 2003. *Statistical models and methods for lifetime data*. 2nd ed. New York: John Wiley and Sons Inc.
- Lindqvist, B. H., and G. Skogsrud. 2008. Modeling of dependent competing risks by first passage times of wiener processes. *IIE Transactions* 41 (1):72–80. doi:[10.1080/07408170802322697](https://doi.org/10.1080/07408170802322697).
- Marshall, A. W., and I. Olkin. 1967. A multivariate exponential distribution. *Journal of the American Statistical Association* 62 (317):30–44. doi:[10.1080/01621459.1967.10482885](https://doi.org/10.1080/01621459.1967.10482885).
- Miyakawa, M. 1984. Analysis of incomplete data in competing risks model. *IEEE Transactions on Reliability* R-33 (4):293–6. doi:[10.1109/TR.1984.5221828](https://doi.org/10.1109/TR.1984.5221828).
- Moeschberger, M. 1974. Life tests under dependent competing causes of failure. *Technometrics* 16 (1):39–47. doi:[10.1080/00401706.1974.10489147](https://doi.org/10.1080/00401706.1974.10489147).
- Moore, A. H., and J. E. Bilikam. 1978. Bayesian estimation of parameters of life distributions and reliability from type ii censored samples. *IEEE Transactions on Reliability* R-27 (1):64–7. doi:[10.1109/TR.1978.5220246](https://doi.org/10.1109/TR.1978.5220246).
- Proschan, F., and P. Sullo. 1973. Estimating the parameters of a certain multivariate exponential distribution. Technical report, Dept of Statistics, Florida State University, Tallahassee.
- Shen, Y., and A. Xu. 2018. On the dependent competing risks using Marshall–Olkin bivariate Weibull model: Parameter estimation with different methods. *Communications in Statistics - Theory and Methods* 47 (22):5558–72. doi:[10.1080/03610926.2017.1397170](https://doi.org/10.1080/03610926.2017.1397170).
- Sinha, S. K. 1986. *Reliability and life testing*. New Delhi: Wiley Eastern Ltd.
- Tomer, S. K., A. K. Singh, and M. Panwar. 2014. 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. doi:[10.1007/s13198-013-0191-4](https://doi.org/10.1007/s13198-013-0191-4).
- Tsiatis, A. 1975. A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences of the United States of America* 72 (1):20–2. doi:[10.1073/pnas.72.1.20](https://doi.org/10.1073/pnas.72.1.20).
- Wada, C. Y., P. K. Sen, and S. E. Shimakura. 1996. A bivariate exponential model with covariates in competing risk data. *Calcutta Statistical Association Bulletin* 46 (3-4):197–210. doi:[10.1177/0008068319960305](https://doi.org/10.1177/0008068319960305).
- Wang, C.-P., and M. Ghosh. 2003. Bayesian analysis of bivariate competing risks models with covariates. *Journal of Statistical Planning and Inference* 115 (2):441–59. doi:[10.1016/S0378-3758\(02\)00177-5](https://doi.org/10.1016/S0378-3758(02)00177-5).
- Xu, A., and S. Zhou. 2017. Bayesian analysis of series system with dependent causes of failure. *Statistical Theory and Related Fields* 1 (1):128–40. doi:[10.1080/24754269.2017.1348708](https://doi.org/10.1080/24754269.2017.1348708).

Appendix A

By using the appropriate expressions from (2.6) and (2.7), we have

$$\begin{aligned}
 L(\underline{t}; \Theta) &= \prod_{i=1}^{n_0} \left\{ \frac{\theta_0}{\Delta} f_{GLD}(t_j; \beta, \Delta) \right\} \prod_{i=1}^{n_1} \{S_{GLD}(t_j; \beta, \theta_0 + \theta_2) f_{GLD}(t_j; \beta, \theta_1)\} \\
 &\quad \prod_{i=1}^{n_2} \{S_{GLD}(t_j; \beta, \theta_0 + \theta_1) f_{GLD}(t_j; \beta, \theta_2)\} \\
 &= \theta_0^{n_0} \beta^{n_0} \prod_{i=1}^{n_0} \left\{ [g(t_j)]^{\beta-1} g'(t_j) \right\} \exp \left\{ -\Delta \sum_{i=1}^{n_0} [g(t_j)]^\beta \right\} \\
 &\quad \cdot \exp \left\{ -(\theta_0 + \theta_2) \sum_{i=1}^{n_1} [g(t_j)]^\beta \right\} \beta^{n_1} \theta_1^{n_1} \prod_{i=1}^{n_1} \left\{ [g(t_j)]^{\beta-1} g'(t_j) \right\} \exp \left\{ -\theta_1 \sum_{i=1}^{n_1} [g(t_j)]^\beta \right\} \\
 &\quad \cdot \exp \left\{ -(\theta_0 + \theta_1) \sum_{i=1}^{n_2} [g(t_j)]^\beta \right\} \beta^{n_2} \theta_2^{n_2} \prod_{i=1}^{n_2} \left\{ [g(t_j)]^{\beta-1} g'(t_j) \right\} \exp \left\{ -\theta_2 \sum_{i=1}^{n_2} [g(t_j)]^\beta \right\}.
 \end{aligned}$$

Hence the likelihood is

$$L(t; \Theta) = \beta^n \theta_0^{n_0} \theta_1^{n_1} \theta_2^{n_2} \prod_{i=1}^n \left\{ [g(t_j)]^{\beta-1} g'(t_j) \right\} \exp \left\{ -\Delta \sum_{i=1}^n [g(t_j)]^\beta \right\}. \quad (\text{A1})$$

Appendix B

If assume that $z = \Delta[g(t)]^\beta$ then z follows exponential distribution with mean 1. Now as $[g(t)]^\beta = z/\Delta$ and it implies that $g(t) = (z/\Delta)^{1/\beta}$ or $\ln g(t) = \frac{1}{\beta} \ln(z/\Delta)$. Hence the expectation term can be calculated such as

$$\begin{aligned} E\left[[g(t)]^\beta \ln g(t) \right] &= E\left[\frac{z}{\Delta} \cdot \frac{1}{\beta} \ln \left(\frac{z}{\Delta} \right) \right] \\ &= \frac{1}{\Delta \beta} [E[z \ln z] - \ln \Delta] \\ &= \frac{1}{\Delta \beta} \left[\int_0^\infty z \ln z \exp\{-z\} dz - \ln \Delta \right] \\ &= \frac{1}{\Delta \beta} \left[-|z \ln z \exp\{-z\}|_0^\infty + \int_0^\infty (1 + \ln z) \exp\{-z\} dz - \ln \Delta \right] \\ &= \frac{1}{\Delta \beta} [r_1 + (1 - \ln \Delta)], \end{aligned} \quad (\text{B1})$$

where $r_u = \int_0^\infty (\ln z)^u \exp\{-z\} dz$ and similarly

$$\begin{aligned} E\left[[g(t)]^\beta (\ln g(t))^2 \right] &= E\left[\frac{z}{\Delta} \cdot \frac{1}{\beta^2} \ln^2 \left(\frac{z}{\Delta} \right) \right] \\ &= \frac{1}{\Delta \beta^2} [E[z(\ln z)^2] + (\ln \Delta)^2 - 2(1 + r_1) \ln \Delta] \\ &= \frac{1}{\Delta \beta^2} \left[\int_0^\infty z(\ln z)^2 e^{-z} dz + (\ln \Delta)^2 - 2(1 + r_1) \ln \Delta \right] \\ &= \frac{1}{\Delta \beta^2} \left[2 \int_0^\infty \ln z e^{-z} dz + \int_0^\infty (\ln z)^2 e^{-z} dz + (\ln \Delta)^2 - 2(1 + r_1) \ln \Delta \right] \\ &= \frac{1}{\Delta \beta^2} [r_2 + 2r_1(1 - \ln \Delta) - \ln \Delta(2 - \ln \Delta)]. \end{aligned} \quad (\text{B2})$$

Appendix C

Case (i): When β is known

After re-parametrization, the Jacobian matrix of the transformation and Fisher information matrix are follows:

$$J_1 = \begin{bmatrix} 1 - \lambda_2 - \lambda_3 & -\lambda_1 & -\lambda_1 \\ \lambda_2 & \lambda_1 & 0 \\ \lambda_3 & 0 & \lambda_1 \end{bmatrix}$$

and

$$\Sigma_1 = J'_1 \Sigma_1 J_1 = n \begin{bmatrix} \frac{1}{\lambda_1^2} & 0 & 0 \\ 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}.$$

Case (ii): When β is unknown

The Jacobian matrix of the transformation and Fisher information matrix are follows:

$$J_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \lambda_2 - \lambda_3 & -\lambda_1 & -\lambda_1 \\ 0 & \lambda_2 & \lambda_1 & 0 \\ 0 & \lambda_3 & 0 & \lambda_1 \end{bmatrix}$$

and

$$S_2 = J'_2 \Sigma_2 J_2 = n \begin{bmatrix} \frac{k(\Lambda)}{\beta^2} & \frac{r_1+1-\ln \lambda_1}{\beta \lambda_1} & 0 & 0 \\ \frac{r_1+1-\ln \lambda_1}{\beta \lambda_1} & \frac{1}{\lambda_1^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda_2} + \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{1-\lambda_2-\lambda_3} \\ 0 & 0 & \frac{1}{1-\lambda_2-\lambda_3} & \frac{1}{\lambda_3} + \frac{1}{1-\lambda_2-\lambda_3} \end{bmatrix}$$

Appendix D. Reference prior for MOBGLD

To obtain the reference prior for parameters according to our inferential interest, we proceed as follows. First of all, order a multi-dimensional parameter $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)$ and separate them into m groups of sizes g_1, g_2, \dots, g_m such as $\alpha_{(i)} = (\alpha_{N_{i-1}+1}, \alpha_{N_{i-1}+2}, \dots, \alpha_{N_i})$..., where $N_j = \sum_{i=1}^j g_i$. Also let $\alpha_{[j]} = (\alpha_{(1)}, \alpha_{(2)}, \dots, \alpha_{(j)}) = (\alpha_1, \alpha_2, \dots, \alpha_{N_j})$ and $\alpha_{[\sim j]} = (\alpha_{N_j+1}, \dots, \alpha_k)$, then it is obvious that $\alpha_{[\sim 0]} = \alpha$ and $\alpha_{[0]}$ is vacuous. Let $\Sigma(\alpha)$ and $\Sigma^{-1}(\alpha)$ are Fisher information matrix and variance-covariance matrix respectively. Here Σ^{-1} can be defined as block partition matrix such as

$$\Sigma^{-1} = \begin{pmatrix} A_{11} & A_{21}^t & \dots & A_{m1}^t \\ A_{21} & A_{22} & \dots & A_{m2}^t \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{pmatrix},$$

where A_{ij} is a $n_j \times n_j$ matrix. Let Σ_j^{-1} = Upper Left $(N_j \times N_j)$ corner of Σ^{-1} , with $\Sigma_m^{-1} = \Sigma^{-1}$, and $H_j = \Sigma_j$. Then the matrices h_j = Lower Right $(n_j \times n_j)$ corner of H_j ; $j = 1, 2, \dots, m$, have an important role in deriving the reference priors. In particular, $h_1 = H_1 = A_{11}^{-1}$ and, if S is a block diagonal matrix, then $h_j = A_{jj}^{-1}$, $j = 1, 2, \dots, m$. Finally, the reference prior has been derived for possible orders using the lemma given by Berger and Bernardo (1992).

Lemma D.1. If $|h_j(\alpha)|$ depending only on $\alpha_{[j]}$ holds, for $j = 1, 2, \dots, m$, then the reference prior

$$\pi(\alpha) = \lim_{k \rightarrow \infty} \frac{\pi^k(\alpha)}{\pi^k(\alpha*)}, \quad \text{where } \pi^k(\alpha) = \left(\prod_{j=1}^m \frac{|h_j|^{1/2}}{\int_{\Theta(\alpha_{[j-1]})} |h_j|^{1/2} d\alpha_{(j)}} \right) I_{\Theta}(\alpha),$$

$|h_j(\alpha)|$ is the determinant of h_j , $\alpha*$ is any fixed point in Θ , and Θ is a compact subset.

Here, few reference priors have been derived by considering the parameter β known as well as unknown.

Case (i): when β is known

Let consider an order grouping of parameters such as $\{\lambda_1, (\lambda_2, \lambda_3)\}$ to derive the reference prior, say $\omega_{11}(\Lambda)$. As here in this case, the number of groups is $m=2$ with group sizes $g_1 = 1$ and $g_2 = 2$, given by $\alpha_{(1)} = \lambda_1$ and $\alpha_{(2)} = (\lambda_2, \lambda_3)$. Also $\alpha_{[1]} = \lambda_1$ and $\alpha_{[2]} = (\lambda_1, \lambda_2, \lambda_3)$ with $\alpha_{[\sim 0]} = (\lambda_1, \lambda_2, \lambda_3)$ and $\alpha_{[\sim 1]} = (\lambda_2, \lambda_3)$. $\alpha_{[0]}$ and $\alpha_{[\sim 2]}$ are vacuous. The variance-covariance matrix for $\{\lambda_1, (\lambda_2, \lambda_3)\}$ is

$$\Sigma^{-1} = \begin{bmatrix} \frac{\lambda_1^2}{n} & 0 & 0 \\ 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}.$$

Hence $\Sigma_1^{-1} = \frac{\lambda_1^2}{n}$, $H_1 = \frac{n}{\lambda_1^2}$, $h_1 = H_1$, $\Sigma_2^{-1} = \Sigma^{-1}$,

$$H_2 = \Sigma = \begin{bmatrix} \frac{n}{\lambda_1^2} & 0 & 0 \\ 0 & n \frac{1-\lambda_3}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{1-\lambda_2-\lambda_3} \\ 0 & \frac{n}{1-\lambda_2-\lambda_3} & n \frac{1-\lambda_2}{\lambda_3(1-\lambda_2-\lambda_3)} \end{bmatrix}, \quad \text{and } h_2 = n \begin{bmatrix} \frac{1-\lambda_3}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{1}{1-\lambda_2-\lambda_3} \\ \frac{1}{1-\lambda_2-\lambda_3} & \frac{1-\lambda_2}{\lambda_3(1-\lambda_2-\lambda_3)} \end{bmatrix}.$$

Now let choose the parametric space such as $\Theta_k = \{\lambda_1, (\lambda_2, \lambda_3)\} | a_{1k} < \lambda_1 < b_{1k}, a_{2k} < \lambda_2, a_{3k} < \lambda_3, \lambda_2 + \lambda_3 < d_k\}$, where $a_{1k}, a_{2k}, a_{3k} \rightarrow 0$, $b_{1k} \rightarrow \infty$, $d_k \rightarrow 1$. Since h_1 and h_2 satisfy Lemma D.1, so the reference prior for $\{\lambda_1, (\lambda_2, \lambda_3)\}$ is given by

$$\omega_{11}(\Lambda) = [\lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1}. \quad (\text{D1})$$

Case (ii): when β is unknown

Let first derive the reference prior, say $\omega_{21}(\beta, \Lambda)$, for the ordered grouping $\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$. In this case, the number of groups is $m=3$ given as $\alpha_{(1)} = (\lambda_2, \lambda_3)$, $\alpha_{(2)} = \beta$ and $\alpha_{(3)} = \lambda_1$. Also $\alpha_{[1]} = (\lambda_2, \lambda_3)$, $\alpha_{[2]} = (\lambda_2, \lambda_3, \beta)$, $\alpha_{[3]} = (\lambda_2, \lambda_3, \beta, \lambda_1)$, $\alpha_{[\sim 0]} = (\lambda_2, \lambda_3, \beta, \lambda_1)$; $\alpha_{[\sim 1]} = (\beta, \lambda_1)$ and $\alpha_{[\sim 2]} = \lambda_1$ and $\alpha_{[0]}$ and $\alpha_{[\sim 3]}$ as are vacuous. So the variance-covariance matrix for $\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$ is

$$\Sigma^{-1} = \begin{bmatrix} \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} & 0 & 0 \\ -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} & 0 & 0 \\ 0 & 0 & \frac{\beta^2}{n(r_2-r_1^2)} & -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} \\ 0 & 0 & -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2-r_1^2)} \end{bmatrix}.$$

Then we obtain

$$\Sigma_1^{-1} = \begin{bmatrix} \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}, \quad H_1 = \begin{bmatrix} \frac{n(1-\lambda_3)}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{(1-\lambda_2-\lambda_3)} \\ \frac{n}{(1-\lambda_2-\lambda_3)} & \frac{n(1-\lambda_2)}{\lambda_3(1-\lambda_2-\lambda_3)} \end{bmatrix},$$

$$\Sigma_2^{-1} = \begin{bmatrix} \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} & 0 \\ -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} & 0 \\ 0 & 0 & \frac{\beta^2}{n(r_2-r_1^2)} \end{bmatrix}, \quad H_2 = \begin{bmatrix} \frac{n(1-\lambda_3)}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{(1-\lambda_2-\lambda_3)} & 0 \\ \frac{n}{(1-\lambda_2-\lambda_3)} & \frac{n(1-\lambda_2)}{\lambda_3(1-\lambda_2-\lambda_3)} & 0 \\ 0 & 0 & \frac{n(r_2-r_1^2)}{\beta^2} \end{bmatrix},$$

$$\Sigma_3^{-1} = \Sigma^{-1} \text{ and } H_3 = \Sigma = \begin{bmatrix} \frac{n(1-\lambda_3)}{\lambda_2(1-\lambda_2-\lambda_3)} & \frac{n}{(1-\lambda_2-\lambda_3)} & 0 & 0 \\ \frac{n}{(1-\lambda_2-\lambda_3)} & \frac{n(1-\lambda_2)}{\lambda_3(1-\lambda_2-\lambda_3)} & 0 & 0 \\ 0 & 0 & \frac{nk(\lambda_1)}{\beta^2} & \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} \\ 0 & 0 & \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} & \frac{n}{\lambda_1^2} \end{bmatrix}.$$

Hence, we get $h_1 = H_1$, $h_2 = \frac{n(r_2-r_1^2)}{\beta^2}$ and $h_3 = \frac{n}{\lambda_1^2}$. Now the parametric space $\Theta_k = \{(\lambda_2, \lambda_3), \beta, \lambda_1\} | a_{1k} < \beta < b_{1k}, a_{2k} < \lambda_1 < b_{2k}, a_{3k} < \lambda_2, a_{4k} < \lambda_3, \lambda_2 + \lambda_3 < d_k\}$; where $a_{1k}, a_{2k}, a_{3k}, a_{4k} \rightarrow 0$, $b_{1k}, b_{2k} \rightarrow \infty$, $d_k \rightarrow 1$. Since h_1, h_2 and h_3 satisfy the Lemma D.1, so the reference prior for $\{(\lambda_2, \lambda_3), \beta, \lambda_1\}$ is given by

$$\omega_{21}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2 - \lambda_3)]^{-1}. \quad (\text{D2})$$

Let attempt for one more reference prior, say, $\omega_{24}(\beta, \Lambda)$ for ordered grouping $\{\beta, \lambda_1, \lambda_2, \lambda_3\}$. For this case, the number of groups $m=4$ which is given by $\alpha_{(1)} = \beta, \alpha_{(2)} = \lambda_1, \alpha_{(3)} = \lambda_2$ and $\alpha_{(4)} = \lambda_3$. So we have $\alpha_{[1]} = \beta, \alpha_{[2]} = (\beta, \lambda_1), \alpha_{[3]} = (\beta, \lambda_1, \lambda_2), \alpha_{[4]} = (\beta, \lambda_1, \lambda_2, \lambda_3), \alpha_{[\sim 0]} = (\beta, \lambda_1, \lambda_2, \lambda_3), \alpha_{[\sim 1]} = (\lambda_1, \lambda_2, \lambda_3), \alpha_{[\sim 2]} = (\lambda_2, \lambda_3), \alpha_{[\sim 3]} = \lambda_3$ whereas $\alpha_{[0]}$ and $\alpha_{[\sim 4]}$ are vacuous. The variance-covariance matrix for $\{\beta, \lambda_1, \lambda_2, \lambda_3\}$ is

$$\Sigma^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2-r_1^2)} & -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} & 0 & 0 \\ -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2-r_1^2)} & 0 & 0 \\ 0 & 0 & \frac{\lambda_2(1-\lambda_2)}{n} & -\frac{\lambda_2\lambda_3}{n} \\ 0 & 0 & -\frac{\lambda_2\lambda_3}{n} & \frac{\lambda_3(1-\lambda_3)}{n} \end{bmatrix}$$

Then we get $\Sigma_1^{-1} = \frac{\beta^2}{n(r_2-r_1^2)}$, $H_1 = \frac{n(r_2-r_1^2)}{\beta^2}$,

$$\Sigma_2^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2-r_1^2)} & -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} \\ -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2-r_1^2)} \end{bmatrix}, \quad \Sigma_3^{-1} = \begin{bmatrix} \frac{\beta^2}{n(r_2-r_1^2)} & -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} & 0 \\ -\frac{\beta\lambda_1(1+r_1-\ln\lambda_1)}{n(r_2-r_1^2)} & \frac{\lambda_1^2 k(\lambda_1)}{n(r_2-r_1^2)} & 0 \\ 0 & 0 & \frac{\lambda_2(1-\lambda_2)}{n} \end{bmatrix},$$

$$H_2 = \begin{bmatrix} \frac{nk(\lambda_1)}{\beta^2} & \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} \\ \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} & \frac{n}{\lambda_1^2} \end{bmatrix}, \quad H_3 = \begin{bmatrix} \frac{nk(\lambda_1)}{\beta^2} & \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} & 0 \\ \frac{n(1+r_1-\ln\lambda_1)}{\beta\lambda_1} & \frac{n}{\lambda_1^2} & 0 \\ 0 & 0 & \frac{n}{\lambda_2(1-\lambda_2)} \end{bmatrix},$$

$\Sigma_4^{-1} = \Sigma^{-1}$ and $H_4 = \Sigma$. Hence we obtain

$$h_1 = \frac{n(r_2-r_1^2)}{\beta^2}; \quad h_2 = \frac{n}{\lambda_1^2}; \quad h_3 = \frac{n}{\lambda_2(1-\lambda_2)} \text{ and } h_4 = \frac{1-\lambda_2}{\lambda_3(1-\lambda_2-\lambda_2)}.$$

Now, for the parametric space $\Theta_k = \{(\beta, \lambda_1, \lambda_2, \lambda_3) | a_{1k} < \beta < b_{1k}, a_{2k} < \lambda_1 < b_{2k}, a_{3k} < \lambda_2 < b_{2k}, a_{4k} < \lambda_3, \lambda_2 + \lambda_3 < d_k\}$; where $a_{1k}, a_{2k}, a_{3k}, a_{4k} \rightarrow 0$, $b_{1k}, b_{2k} \rightarrow \infty$, $d_k \rightarrow 1$. Now we have to compute the following integral $\int_{\Theta(x_{[j-1]})} |h_j|^{1/2} dx_{(j)}$ for $j = 1, 2, 3$ and 4. So we obtain

$$\begin{aligned}\int_{\Theta(x_{[0]})} |h_1|^{1/2} d\beta &= \int_{a_{1k}}^{b_{1k}} |h_1|^{1/2} d\beta = \{n(r_2 - r_1^2)\}^{1/2} (\log b_{1k} - \log a_{1k}), \\ \int_{\Theta(x_{[1]})} |h_2|^{1/2} d\lambda_1 &= \int_{a_{2k}}^{b_{2k}} |h_2|^{1/2} d\lambda_1 = n^{1/2} (\log b_{2k} - \log a_{2k}).\end{aligned}$$

For some λ_3° , we get

$$\begin{aligned}\int_{\Theta(x_{[2]})} |h_3|^{1/2} d\lambda_2 &= \int_{a_{3k}}^{d_k - \lambda_3^\circ} |h_3|^{1/2} d\lambda_2 = n^{1/2} [\sin^{-1}(1 - 2a_{3k}) - \sin^{-1}(1 - 2(d_k - \lambda_3^\circ))], \\ \int_{\Theta(x_{[3]})} |h_4|^{1/2} d\lambda_3 &= \int_{a_{4k}}^{d_k - \lambda_2} |h_4|^{1/2} d\lambda_3 \\ &= \sqrt{1 - \lambda_2} \left[\sin^{-1} \left(\frac{1 - \lambda_2 - 2a_{4k}}{1 - \lambda_2} \right) - \sin^{-1} \left(\frac{1 - \lambda_2 - 2(d_k - \lambda_2)}{1 - \lambda_2} \right) \right].\end{aligned}$$

Using lemma D.1, we obtain the reference prior for $\{\beta, \lambda_1, \lambda_2, \lambda_3\}$, given by

$$\omega_{24}(\beta, \Lambda) = [\beta^2 \lambda_1^2 \lambda_2 \lambda_3 (1 - \lambda_2)(1 - \lambda_2 - \lambda_3)]^{-1}. \quad (\text{D3})$$

Vikas Barnwal

Curriculum Vitae

Department of Statistics
Banaras Hindu University
Varanasi (U.P.) - 221005, India
+91-8896538410
vikas.barnwal5@bhu.ac.in



Doctoral Thesis (Submitted)

Title **Statistical Modelling and Analysis of Time-to-Event Data under Bayesian Paradigm**
Supervisor **Dr. Mahaveer Singh Panwar**
University **Banaras Hindu University**, Varanasi-221005, Uttar Pradesh, India

Academic Qualifications

2016–2018 **Masters of Science**, *University of Allahabad*, Prayagraj, U.P., India.
2012–2015 **Bachelor of Science**, *University of Allahabad*, Prayagraj, U.P., India.

Publications

Published **Competing Risks Analysis for Dependent Causes Using Marshall-Olkin Bivariate Generalized Lifetime Family**, Communications in Statistics - Theory and Methods, Taylor & Francis, July 2022.

Published **Objective Bayesian Analysis of Weibull Distribution with Partial Information**, Communications in Statistics - Theory and Methods, Taylor & Francis, June 2023.

Accepted **A Latent Variable Approach for Parametric Modeling of Recall-Based Current Status Data**, Computational Statistics, Springer, December 2023.

Participation in Conferences/Seminar/Workshop

04 **Oral presentations at National/International Conferences/Seminars**
05 **Workshops Attended**

Computer Skills

Softwares R/R-Studio, SPSS, Python
Office Tools MS-Office, L^AT_EX

Personal Details

Nationality **Indian** Religion **Hindu**
Languages **Hindi and English** Date of Birth **December 25, 1995**

Declaration

The information given above is true to the best of my knowledge.

Date:

Place: Varanasi

(Vikas Barnwal)