Statistical Analysis of Load-sharing Systems

by

Shilpi Biswas

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, INDIA

October, 2025

Statistical Analysis of Load-sharing Systems

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Shilpi Biswas

(Roll No. - 196123009)

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
GUWAHATI-781039, INDIA

October, 2025

Dedicated To My Family

Baba (Sanju Biswas)
&
Maa (Madhabi Biswas)
&
Dii (Lucky Biswas)
&
Bhai (Suraj Biswas)

Declaration

I do hereby declare that this thesis entitled **Statistical Analysis of Load-sharing Systems** is a presentation of my original research work done under the supervision of **Dr. Ayon Ganguly**, Assistant Professor, Department of Mathematics, Indian Institute of Technology Guwahati for the award of the degree of Doctor of Philosophy and this work has not been submitted elsewhere for a degree.

October, 2025

Shilpi Biswas
Roll No. 196123009
Department of Mathematics
Indian Institute of Technology Guwahati

Certificate

It is to certify that the work contained in this thesis entitled **Statistical Analysis of Load-sharing Systems** has been carried out by **Shilpi Biswas**, a student in the Department of Mathematics, Indian Institute of Technology Guwahati, under my supervision for the award of the degree of Doctor of Philosophy and this work has not been submitted elsewhere for a degree.

October, 2025

Dr. Ayon Ganguly
Assistant Professor
Department of Mathematics
Indian Institute of Technology Guwahati

Acknowledgements

With utmost reverence and devotion, I dedicate this thesis to **Baba Shani Dev** and **Maa Rakhha Kali**, the eternal sources of my strength, guidance, and inspiration. The divine blessings have given me the courage to endure hardships, the wisdom to walk the right path, and the resilience to rise above every challenge. Whatever I have achieved is by their grace, and I bow in gratitude for their constant protection and unconditional love.

The journey toward completing my thesis would not have been possible without the support and encouragement of several special individuals. I would like to take this opportunity to sincerely acknowledge their invaluable contributions throughout my time at IIT Guwahati.

First and foremost, I am deeply grateful to my supervisor, Dr. Ayon Ganguly, for his unwavering support, guidance, and inspiration throughout my research tenure. I extend my heartfelt appreciation to him for his patience, understanding, and constant belief in my abilities. His vast knowledge, rich experience, and insightful feedback have been crucial in helping me navigate the challenges of my research. His continuous encouragement has been a source of strength and motivation, both academically and personally. It is his persistent support and belief in the quality of my work that made the completion of this thesis possible.

I would also like to express my sincere thanks to my doctoral committee members, Prof. N. Selvaraju, Dr. Palash Ghosh, and Dr. Buddhadipta Hazra for their valuable suggestions, support, and encouragement throughout the course of this work.

I would like to extend my sincere thanks to my collaborator, Dr. Debanjan Mitra from IIM Udaipur, for his valuable suggestions and insightful comments throughout my research journey. His expert advice played a significant role in refining my research ideas and contributed meaningfully to the quality of this work.

I am especially thankful to the Head of the Department of Mathematics at IIT Guwahati for providing the necessary infrastructure and facilities for my research. I extend my gratitude to all the faculty members of the Department of Mathematics for their guidance and cooperation. My thanks also go to the staff members of the department for their consistent assistance and support in various ways during my research journey.

I would like to express my sincere gratitude to Dr. Subhamay Saha for his constant motivation and valuable suggestions during my progress seminars. His insightful advice has greatly supported and enriched my research. I also want to thank Dr. Chandan Pal, Dr. Arup Chattopadhyay, and Dr. Farha Sultana for their encouragement during my PhD journey.

I am deeply grateful to the Ministry of Human Resource Development, Government of India, for providing the financial assistance that supported my research over the years. I also sincerely acknowledge the Indian Institute of Technology Guwahati for offering an intellectually stimulating academic environment, exceptional resources, and unwavering institutional support throughout my time here.

Words cannot fully express my deepest gratitude to my parents: Sanju Biswas (Baba), Madhabi Biswas (Maa), elder sister: Lucky Biswas (Dii), and younger brother: Suraj Biswas (Baby), whose unconditional love, constant care, and boundless encouragement have shaped every step of my journey. Their unwavering belief in my abilities has been a source of strength, resilience, and hope, guiding me through challenges and celebrating my milestones. I owe every achievement to the values they instilled in me and to the endless support they continue to provide, which has been the greatest blessing in my life. It would not have been possible for me to reach this far without their sacrifices and love for me.

I would like to express my heartfelt gratitude to some very special places on our campus—my cabin(RS-E1-16), Core 1, Core 2, the old badminton court, the mess, and the canteens. These spaces have been more than just physical settings; they have offered me hope, strength, and fresh motivation to rise from difficult times and face challenges with renewed courage. Equally important are the wonderful people with whom I have shared countless moments at those places—those who listened to my struggles, stood by me, and encouraged me. They hold a very special place in my heart, and I sincerely acknowledge their support and companionship.

I would like to thank my batchmate Khyodeno, who has become my friend and gradually my soul sister, and two of my juniors, Sourav and Amit, who have become like my little brothers for their constant encouragement, support, and companionship throughout my research journey. Also, I would like to thank my friends and batchmates, Sandip, Sagar, Mijanur, Anirudda Da, and Abhijit. I would also like to extend thanks to my juniors and seniors for the many cherished moments we shared in the lab.

Last but not least, I would like to thank everyone who contributed to the completion of the thesis, and I apologize for not being able to mention them all individually.

October, 2025

Shilpi Biswas
Department of Mathematics
Indian Institute of Technology Guwahati

ABSTRACT

Many real-life systems are composed of different components connected in different fashions. The failure of one component in such a system, in general, increases the workload of the other surviving components. This increase in load changes the lifetime behavior of the surviving components. These systems are known as load-sharing systems. Analysis of load-sharing systems is challenging due to their complex dependency between components and the changes in failure rates after a failure. This thesis addresses three problems related to the analysis of load-sharing systems. In Chapter 1, we provide basic definitions and background of the problems that are discussed in this thesis. We describe three datasets to motivate the readers, and a comprehensive literature review is performed in the same chapter. In Chapter 2, we propose a Bayesian estimation framework for the Generalized Freund Bivariate (GFB) distribution to model the reliability of a two-component load-sharing system. This approach offers an alternative to the classical estimation method introduced by Franco et al. [22]. Assuming independent gamma priors, we derive the posterior distribution of model parameters. Due to the complicated expression of the posterior probability density function, we propose to use Markov chain Monte Carlo based method to perform posterior analysis. The performance of the proposed model fitting method is observed to be quite satisfactory through a Monte Carlo simulation study. The analyses of two load-sharing datasets, one about the lifetimes of two-motor load-sharing systems and another related to a nuclear power plant, are provided as illustrative examples. In Chapter 3, for two-component load-sharing systems, a doubly-flexible model is developed where the GFB distribution is used for the baseline of the component lifetimes, and the generalized gamma (GG) family of distributions is used to incorporate a shared frailty that captures dependence between the component lifetimes. We call the model as GFB-GG model. The proposed model structure results in a very general two-way class of models that enables a researcher to choose an appropriate model for a given two-component load-sharing data within the respective families of distributions. The GFB-GG model structure provides a better fit to two-component

load-sharing systems compared to existing models. Fitting methods for the proposed model are discussed. Through Monte Carlo simulations, the effectiveness of the fitting methods is demonstrated. Also, through simulations, it is shown that the proposed model serves the intended purpose of model choice for a given two-component load-sharing data. A simulation case and analysis of a real dataset are presented to illustrate the strength of the proposed model. We also estimate some important reliability characteristics such as reliability at a mission time (RMT), mean time to failure (MTTF), and mean residual time (MRT) under GFB-GG model. In Chapter 4, a flexible model for analysing load-sharing data is developed by approximating the cumulative hazard functions of component lifetimes by piecewise linear functions. The proposed model is data-driven and does not depend on restrictive parametric assumptions on underlying component lifetimes. Maximum likelihood estimation and construction of confidence intervals for model parameters are discussed. Estimates of reliability characteristics such as RMT, quantile function, MTTF, and MRT for load-sharing systems are developed in this setting. As the proposed model is capable of providing a good fit to load-sharing data, it also results in better estimation of these important reliability characteristics. The performance of the proposed model is observed to be quite satisfactory through a detailed Monte Carlo simulation study. The analyses of two load-sharing datasets, one about the lifetimes of two-motor load-sharing systems and another related to basketball games, are provided as illustrative examples. Finally, the thesis is concluded in Chapter 5, and we also mention some related open problems for future directions.

NOMENCLATURE

Abbreviations

ABE	Average Bayes Estimate
AE	Average Estimate
AFT	Accelerated Failure Time
AIC	Akaike Information Criterion
AICc	Corrected Akaike Information Criterion
AIE	Absolute Integrated Error
AL	Average Length
BC	Bridge Criterion
BE	Bayes Estimate
BIC	Bayesian Information Criterion
BUE	Best Unbiased Estimator
CDF	Cumulative Distribution Function
CHF	Cumulative Hazard function
CI	Confidence Interval
CP	Coverage Percentage
EFB	Extended Freund's Bivariate
EM	Expectation-Maximization
ESOS	Extended Sequential Order Statistics
FBE	Freund's Bivariate Exponential
GFB	Generalized Freund's Bivariate
GG	Generalized Gamma
HF	Hazard Function
iid	Independent and Identically Distributed
lppd	Log-Pointwise Predictive Density
MCMC	Markov Chain Monte Carlo
MGF	Moment Generating Function

MLE	Maximum Likelihood Estimate
MRT	Mean Residual Time
MSE	Mean Squared Error
MTTF	Mean Time To Failure
OS	Order Statistics
PDF	Probability Density Function
PFR	Proportional Failure Rate
PH	Proportional Hazard
PLA	Piecewise Linear Approximation
RMT	Reliability at a Mission Time
SD	Standard Deviation
SE	Standard Error
SF	Survival Function
SOS	Sequential Order Statistics
Q-Q	Quantile-Quantile
WAIC	Watanabe-Akaike Information Criterion

LIST OF FIGURES

LIST OF TABLES

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Basic Definitions and Preliminaries

To enhance system reliability, one of the common practices is to use redundancy techniques in reliability engineering. In reliability theory, a redundancy technique is the practice of adding backup components or systems so that if one fails, others can take over and keep the system operating. Most of the time, to investigate the redundancy assumption, independence is considered over the components of the system. If a component fails in a system, it is assumed that it will not affect the failure rates of the remaining working components. However, this assumption of independence is not truly justified in the real cases.

If a component fails in a load-sharing system, then the workload of the system is redistributed among the remaining surviving components. Therefore, the load on each of the working components is increased. This leads to an increase in failure rates of the surviving components in most cases. Therefore, reliability models that include stochastic dependencies in the system's component lifetimes should be developed. Examples of load-sharing systems include cables in a suspension bridge, the central processing unit of a multiprocessor computer, valves or pumps in a hydraulic system, generators in a power plant, kidneys in the human body, etc.

The stochastic dependency models mainly have two broad areas, which are named as shock models and load-sharing system models, also known as dynamic system models. In shock models, a random amount of damage occurs due to exposure to shocks in the system. Moreover, it can be seen that in the load-sharing model, whenever a component fails, the performance of the other components changes.

To understand the reliability of a load-sharing system, it is important to study how the load is applied to the system across time, known as the load pattern. The load-sharing rule describes how the total system load is divided among individual components at any moment. Each time the load pattern changes, the load-sharing rule determines how the new load is distributed. By carefully considering the load-sharing system, the load pattern, and the specific load-sharing rule in use, we can gain a better understanding of how the system's reliability changes over its life. To understand more precisely, some basic definitions and preliminaries are provided below.

1.1.1 k-out-of-n Load-sharing System

Consider a system that is made up of n connected components. The system is in a working state if at least k $(\leq n)$ components work. Once $(n-k+1)$ components fail, the system fails. This kind of system is called a k-out-of-n system. Now, assume that the total load on a k-out-of-n system does change over time. Initially, when all n components work, the total load gets distributed among all components. When a component fails, the total load gets redistributed among the surviving components. Such a k-out-of-n system is called a k-out-of-n load-sharing system.

When $k=1$, the system is called a parallel system. In parallel systems, if at least one component works, the system works, and if all the n components fail, then the system fails, see Figure 1.1. When $k=n$, the system is called a series system. In a series system, if one of the components fails, the system fails; for reference, see Figure 1.2. The assumption

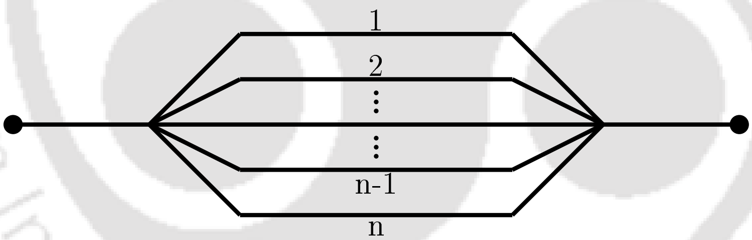


Figure 1.1: n-component parallel system

Figure 1.2: n-component series system

of a constant total load is applicable to a broad range of practical systems, and many works in the literature adopt this setting. In this thesis, we focus on the constant total load assumption for all the models.

1.1.2 Load-sharing Rule

An important element of the load-sharing model is the rule that governs how failure rates change after some components in the system fail. This rule depends on the reliability application and how the components within the system interact with each other. Some of the commonly used load-sharing rules are as follows

- **Equal load-sharing rule:** It implies that the extra load caused by the failed component is shared equally among the surviving ones.
- **Local load-sharing rule:** It dictates that a failed component's load is transferred to adjacent components; the proportion of the load distributed among the surviving components depends on their distance to the failed component.
- **Monotone load-sharing rule:** It states that the load on the surviving components is non-decreasing with respect to the failure of other components in the system. The above two load-sharing rules are special cases of a monotone load-sharing rule.

1.2 Motivating Data

In this section, we describe three datasets that are analysed throughout the thesis. Among the datasets, two are examples of load-sharing data, and one of them is bivariate dependent data.

1.2.1 Two-motor Data

Recently, a dataset on a load-sharing system has been analysed by several authors, including Sutar et al. [56], Asha et al. [5], Franco et al. [22]. The data consist of information on component lifetimes of 18 two-component load-sharing systems. Each system is a parallel combination of two identical motors - "A" and "B". When both motors A and B are in working condition, the total load on the system is equally shared between them. When one of the motors fails, the total load shifts to the operational motor. There is no information available on concomitant variables. Also, there is no censoring in the data, as failures of all the motors are observed. We refer to this data as the two-motor data. The average and standard deviation of lifetimes of the motors that fail first are 178.61 and 62.75, respectively, while those of the lifetime between first and second component failures are 49.72 and 29.45, respectively. The dataset is reported in Table 1.1.

1.2.2 Three-player Basketball Data

Kvam and Peña [32] presented an interesting analysis of sports data with the help of a load-sharing model. The data are from basketball games played by the Basketball Association franchise Boston Celtics. The time period of the used data is the second half of the sports season of 2001 - 2002. For the games played during the above mentioned period by Boston Celtics, Kvam and Peña [32] perceived a three-component load-sharing system

Table 1.1: Time to failure (in days) data set for two motors in a load-sharing configuration

System	Time to failure of motor A	Time to failure of motor B	Event description
1	102	65	B failed first
2	84	148	A failed first
3	88	202	A failed first
4	156	121	B failed first
5	148	123	B failed first
6	139	150	A failed first
7	245	156	B failed first
8	235	172	B failed first
9	220	192	B failed first
10	207	214	A failed first
11	250	212	B failed first
12	212	220	A failed first
13	213	265	A failed first
14	220	275	A failed first
15	243	300	A failed first
16	300	248	B failed first
17	257	330	A failed first
18	263	350	A failed first

consisting of the three star players of Boston Celtics as the components, and defined a component failure as “the event when a player fouled out or was removed from the game due to accumulating a high number of fouls”. Relevant here is that a player is usually temporarily removed from the game for a very brief time period when they accumulate two fouls; this brief time period acts as a cool-down period for the player. For each of the 28 games in the dataset, the three-player basketball data gives elapsed game times at which each of the three players commits their second personal foul in that game. For all these games, all three players started the game and committed at least two fouls by the end of the game. The idea is that once a player commits two fouls and returns to the game after a brief cool-down period, the foul rate of all three star players changes, as follows. The player who reached the two-foul limit plays conservatively, not to commit more fouls. The rate of fouls of the other two star players, however, depends on the strategy of the team for that game: it may go down if the team plays conservatively on the whole, or may go up if the team plays aggressively, where these two star players take more responsibility. The average and standard deviation of lifetimes of the components that fail first are 18.37 and 8.73, respectively, while those of the lifetime between the first and second component failures are 9.97 and 7.19, respectively, and for the lifetime between the second and third component failures are 9.04 and 6.50, respectively. The dataset is provided in Table 1.2.

Table 1.2: Time until second foul for the three-star players

Game	Player 1	Player 2	Player 3
1	21.02	30.22	43.43
2	24.25	45.54	17.19
3	6.56	19.47	23.28
4	15.35	16.37	25.4
5	39.08	30.32	43.53
6	16.2	4.16	39.52
7	34.59	46.44	16.33
8	19.1	38.4	20.17
9	28.22	37.43	25.41
10	32	45.52	39.11
11	11.25	19.09	11.59
12	17.39	25.43	22.51
13	28.47	31.15	2.41
14	23.42	31.28	40.03
15	42.06	23.21	45.36
16	28.51	33.59	16.2
17	34.56	32.53	40.44
18	40.33	15.35	28.33
19	27.56	46.21	28.05
20	9.54	36.21	28.12
21	27.09	11.11	23.33
22	40.36	33.21	17.04
23	41.44	36.28	19.13
24	32.23	8.17	41.27
25	7.53	37.31	13.43
26	28.34	35.58	41.48
27	26.32	28.02	29.33
28	30.47	40.4	42.13

1.2.3 Nuclear Reactor Data

The dataset consists of 30 paired observations of failure times and corresponding repair times for nuclear reactors. These data were obtained from the Operational Performance Monitoring System used in nuclear power plants across South Korea, as reported by Park and Kim [44]. The sample comprises data from 20 operational nuclear power units, which are reactors that were fully functional and generating electricity during the observation period. Their analysis revealed a significant positive correlation between the two variables, indicating that longer failure times are often accompanied by longer repair times, and vice versa. This dependence is important when modeling warranty-related costs, as it violates the assumption of independence often used in simpler models. The average and standard deviation of failure times are 277.05 and 296.06, respectively, while those of the warranty servicing times are 3.29 and 3.05, respectively. The dataset is presented here in Table 1.3;

see the work by Park and Kim [44] for more details on the dataset.

Table 1.3: Failure times(days) and warranty servicing times(days) for Nuclear Power Plants reported in Park and Kim [44]

System	Failure times	Warranty servicing times
1	353.04	4.37
2	334.72	1.91
3	80.04	2.04
4	6.49	1.72
5	1.34	0.29
6	467.19	1.93
7	0.35	1.82
8	398.86	1.77
9	1048.23	9.61
10	829.39	3.80
11	227.20	2.86
12	260.14	0.31
13	14.00	0.85
14	14.15	2.04
15	38.96	2.73
16	30.27	3.63
17	117.37	2.73
18	126.27	2.55
19	56.45	0.72
20	45.28	3.69
21	267.31	0.36
22	615.64	10.63
23	115.37	11.24
24	359.76	9.70
25	412.30	3.31
26	276.69	4.96
27	601.04	2.99
28	1021.01	2.36
29	192.17	1.63
30	0.36	0.26

1.3 Models for Load-sharing Systems

1.3.1 Freund's Bivariate Exponential Model

Consider a two-component parallel system with the components C_1 and C_2 (say), where neither component is replaced or repaired upon failure. Let T_1 and T_2 represent the lifetimes of components C_1 and C_2, respectively. The observed lifetime of one component may change depending on which fails first. Assume that T_1 and T_2 are independent, but not necessarily identically distributed. Specifically, let T_1 and T_2 follow exponential distributions, with respective failure rates θ_1 and θ_2. Thus, the reliability functions for T_1 and T_2, respectively, are given by

$$R_1(t) = P(T_1 > t) = e^{-\theta_1 t},\ t > 0,\ \text{and}\ R_2(t) = P(T_2 > t) = e^{-\theta_2 t},\ t > 0.$$

In load-sharing systems, when a component fails, the failure risk for the surviving component changes. Two possible scenarios are considered, depending on the sequence of failures:

Case-1: If C_1 fails before C_2 (i.e., $T_1 < T_2$), the lifetime of C_2 becomes T_2^*, which is assumed to follow an exponential distribution with rate θ_2^*.

Case-2: If C_2 fails before C_1 (i.e., $T_2 < T_1$), the lifetime of C_1 changes to T_1^*, which is assumed to follow an exponential distribution with rate θ_1^*.

Thus, the FBE model assumes that the initial failure rate of component i, when both components are operating, is

$$\lambda_{i0}(y) = \theta_i,\ y > 0,\ i = 1, 2.$$

After the first failure, the survivor's failure rate changes. If the j-th component fails at time x, the remaining life of the i-th component is exponential with rate

$$\lambda_{ij}(y|x) = \theta_i^*,\ \text{if}\ y > x > 0,\ i = 1, 2;\ j = 1, 2;\ i \neq j.$$

Let the observed lifetime of the components be denoted by (Y_1, Y_2). Under FBE model, the joint PDF of (Y_1, Y_2) is expressed as

$$f_{(Y_1,Y_2)}(y_1, y_2) = \begin{cases} \theta_2\theta_1^*, \exp\left[-\theta_1^* y_1 - (\theta_1 + \theta_2 - \theta_1^*)y_2\right], & \text{if}\ 0 < y_2 < y_1, \\ \theta_1\theta_2^*, \exp\left[-\theta_2^* y_2 - (\theta_1 + \theta_2 - \theta_2^*)y_1\right], & \text{if}\ 0 < y_1 < y_2. \end{cases}$$

Please see Freund [23] for detailed derivations. The distribution specified by the above PDF is referred to as the FBE distribution.

1.3.2 Extended Freund's Bivariate Model

A class of distributions is called a PFR class if its reliability function is given by $(R_B(y;\theta_B))^\alpha$, $\alpha > 0$, with $R_B(\cdot;\theta_B)$ as the baseline reliability function involving parameter θ_B. Here, α is called the PFR parameter. Examples of distributions in the PFR class are exponential, Weibull, gamma, etc. EFB model assumes that T_1 and T_2 have reliability functions from a PFR class with baseline reliability function $R_B(\cdot, \theta_B)$ and power parameters

$\theta_i(>0),\ i=1,2$, respectively, that is

$$R_1(t) = P(T_1 > t) = [R_B(t,\theta_B)]^{\theta_1} \text{ and } R_2(t) = P(T_2 > t) = [R_B(t,\theta_B)]^{\theta_2}.$$

The model also assumes that the distribution of T_1^* and T_2^* belongs to the PFR family with the same baseline reliability function, but with different PFR parameters $\theta_1^*(>0)$ and $\theta_2^*(>0)$, respectively. Therefore, the reliability function of T_1^* and T_2^*, respectively are

$$R_1^*(t) = P(T_1^* > t) = [R_B(t,\theta_B)]^{\theta_1^*} \text{ and } R_2^*(t) = P(T_2^* > t) = [R_B(t,\theta_B)]^{\theta_2^*}.$$

Let $\lambda_{i0}(\cdot)$ denote the failure rate of the i-th component at time y when no component has failed. Then, according to the EFB model,

$$\lambda_{i0}(y) = \theta_i r_B(y,\theta_B),\ y>0,\ i=1,2,$$

where $r_B(\cdot,\theta_B)$ denotes the baseline HF of the lifetimes of the components before the occurrence of the first failure. Let the j-th component fail first. Then, the failure rate of the i-th component at time y, when the j-th component has failed at time x, denoted by $\lambda_{ij}(y|x)$, is given by

$$\lambda_{ij}(y|x) = \theta_i^* r_B(y,\theta_B), \quad \text{if } y > x > 0,\ i=1,2;\ j=1,2;\ i \neq j.$$

The joint PDF of observed lifetimes (Y_1, Y_2) is derived by Asha et al. [4] and given by

$$f_{(Y_1,Y_2)}(y_1,y_2) = \begin{cases} \theta_1^*\theta_2 r_B(y_1,\theta_B) r_B(y_2,\theta_B)\left[\frac{R_B(y_1,\theta_B)}{R_B(y_2,\theta_B)}\right]^{\theta_1^*}[R_B(y_2,\theta_B)]^{(\theta_1+\theta_2)}, \\ \qquad\qquad\qquad \text{if } y_1 > y_2 > 0, \\ \theta_1\theta_2^* r_B(y_1,\theta_B) r_B(y_2,\theta_B)\left[\frac{R_B(y_2,\theta_B)}{R_B(y_1,\theta_B)}\right]^{\theta_2^*}[R_B(y_1,\theta_B)]^{(\theta_1+\theta_2)}, \\ \qquad\qquad\qquad \text{if } y_2 > y_1 > 0. \end{cases}$$

In particular, when $R_B(t,\theta_B) = e^{-t}$ then EFB model becomes the FBE model. Thus, EFB model can be considered as a generalization of FBE model.

1.3.3 Generalized Freund's Bivariate Model

The GFB model assumes that T_1 and T_2 have reliability functions from a PFR class with baseline reliability function $R_B(\cdot;\theta_B)$, and PFR parameters $\theta_i(>0)$, $i=1,2$, respectively. That is,

$$R_1(t) = P(T_1 > t) = [R_B(t,\theta_B)]^{\theta_1} \text{ and } R_2(t) = P(T_2 > t) = [R_B(t,\theta_B)]^{\theta_2}.$$

The model also assumes that the distribution of T_1^* and T_2^* belongs to the PFR family with baseline reliability function $R_B^*(\cdot,\theta_B^*)$ and with different PFR parameters $\theta_1^*(>0)$ and $\theta_2^*(>0)$, respectively. Therefore, the reliability function of T_1^* and T_2^*, respectively, are

$$R_1^*(t)=P(T_1^*>t)=[R_B^*(t,\theta_B^*)]^{\theta_1^*} \text{ and } R_2^*(t)=P(T_2^*>t)=[R_B^*(t,\theta_B^*)]^{\theta_2^*}.$$

Let $\lambda_{i0}(\cdot)$ denote the failure rate of the i-th component at time y when no component has failed. Then, according to the GFB model,

$$\lambda_{i0}(y)=\theta_i r_B(y,\theta_B),\ y>0,\ i=1,2,$$

where $r_B(\cdot,\theta_B)$ denotes the baseline HF of the lifetimes of the components before the occurrence of the first failure. Let the j-th component fail first. Then, the failure rate of the i-th component at time y, when the j-th component has failed at time x, denoted by $\lambda_{ij}(y|x)$, is given by

$$\lambda_{ij}(y|x)=\theta_i^* r_B^*(y,\theta_B^*),\quad \text{if } y>x>0,\ i=1,2;\ j=1,2;\ i\neq j.$$

The PDF of (Y_1,Y_2) is derived by Franco et al. [22] and given by

$$f_{(Y_1,Y_2)}(y_1,y_2)=\begin{cases}\theta_1^*\theta_2 r_B^*(y_1,\theta_B^*)r_B(y_2,\theta_B)\left[\frac{R_B^*(y_1,\theta_B^*)}{R_B^*(y_2,\theta_B^*)}\right]^{\theta_1^*}[R_B(y_2,\theta_B)]^{(\theta_1+\theta_2)}, & \\ \qquad\qquad \text{if } y_1>y_2>0, \\ \theta_1\theta_2^* r_B(y_1,\theta_B)r_B^*(y_2,\theta_B^*)\left[\frac{R_B^*(y_2,\theta_B^*)}{R_B^*(y_1,\theta_B^*)}\right]^{\theta_2^*}[R_B(y_1,\theta_B)]^{(\theta_1+\theta_2)}, & \\ \qquad\qquad \text{if } y_2>y_1>0.\end{cases}$$

The distribution specified by the above PDF is called GFB distribution. We denote it by

$$(Y_1,Y_2)\sim GFB(R_B,R_B^*,\theta_1,\theta_2,\theta_1^*,\theta_2^*,\theta_B,\theta_B^*),$$

with PFR parameters $\theta_i,\theta_i^*>0$ for $i=1,2$, and baseline parameters θ_B,θ_B^*. The GFB family of distributions can be used to generate new bivariate lifetime models by combining different baseline reliability functions R_B and R_B^*. Note that for $R_B=R_B^*$, GFB model resembles the EFB model of Asha et al. [4].

1.3.4 Sequential Order Statistics Model

Consider a load-sharing system consisting of n components. Let the n components be labeled consecutively from 1 to n, and introduce indicator random variables

$$C_i,\quad i\in\{1,2,\ldots,n\},$$

where C_i denotes the label of the component that fails at the ith failure. The random vector

$$\mathbf{\Pi}_k = (C_k, C_{k-1}, \ldots, C_1), \quad 1 \leq k \leq n,$$

represents the sequence of failed components up to the kth failure, listed in reverse order. For a specific realization $\pi_k = (c_k, c_{k-1}, \ldots, c_1)$, the set of surviving components after the kth failure is

$$B_{\pi_k} = \{1, 2, \ldots, n\} \setminus \{c_1, c_2, \ldots, c_k\}.$$

Stage 1: At time $x_0 = 0$, all components are operational. Their initial lifetimes are modeled by n iid random variables

$$Y_1^{(1)}, Y_2^{(2)}, \ldots, Y_n^{(1)}, \quad Y_i^{(1)} \overset{\text{iid}}{\sim} F_1, \quad i \in \{1, 2, \ldots, n\},$$

where F_1 is the common CDF at stage 1. The failure time of the component that fails first is

$$X_*^{(1)} = \min\{Y_1^{(1)}, Y_2^{(1)}, \ldots, Y_n^{(1)}\}.$$

Suppose component c_1 fails first at $X_*^{(1)} = x_1$.

Stage 2: After the first failure, the lifetimes of the surviving components are reassigned. For each $i \in B_{\pi_1}$,

$$Y_i^{(2)} \mid \pi_1 \overset{\text{iid}}{\sim} F_{2|\pi_1},$$

where $Y_i^{(2)}$ denotes the lifetime of component i in stage 2, and $F_{2|\pi_1}$ is the conditional CDF governing the distribution of surviving components given that the first failure sequence is π_1. From these, we define the second failure time as

$$X_*^{(2)} = \min\{Y_i^{(2)} \mid \pi_1 : i \in B_{\pi_1}\}.$$

This procedure is repeated until the nth failure is observed.

The random variables $X_*^{(1)}, X_*^{(2)}, \ldots, X_*^{(n)}$ with corresponding failure sources $C_1, C_2, \ldots, C_n$ are called the SOS in a system of size n if their joint density is

$$f_{X_*^{(n)}, C_n, \ldots, X_*^{(1)}, C_1}(x_n, c_n, \ldots, x_1, c_1) = \prod_{k=1}^{n} \frac{f_{k|\pi_{k-1}}(x_k) \left(\bar{F}_{k|\pi_{k-1}}(x_k)\right)^{n-k}}{\left(\bar{F}_{k|\pi_{k-1}}(x_{k-1})\right)^{n-k+1}},$$

where $\pi_0 = \emptyset$. For more details, see Kamps [30] and Cramer and Kamps [15].

1.3.5 Extended Sequential Order Statistics Model

Stage 1: At time $x_0 = 0$, all components are operational. Their initial lifetimes are modeled by n independent random variables

$$Y_1^{(1)}, Y_2^{(1)}, \dots, Y_n^{(1)}, \quad Y_i^{(1)} \sim F_i, \quad i \in \{1, 2, \dots, n\},$$

having independent CDFs F_i, $1 \leq i \leq n$ at stage 1. The failure time of the component that fails first is

$$X_*^{(1)} = \min\{Y_1^{(1)}, Y_2^{(1)}, \dots, Y_n^{(1)}\}.$$

Suppose component c_1 fails first at $X_*^{(1)} = x_1$.

Stage 2: After the first failure, the lifetimes of the surviving components are reassigned. Let $Y_i^{(2)}$ denote the lifetime of component i in Stage 2, and $F_{i|\pi_1}$ is the continuous conditional CDF governing the distribution of surviving components given that the first failure sequence is π_1. Then,

$$Y_i^{(2)} \mid \pi_1 \sim F_{i|\pi_1}, \quad i \in B_{\pi_1},$$

where $F_{i|\pi_1} = F_i(\cdot \mid \pi_1)$ satisfies the technical restriction

$$F_{i|\pi_1}^{-1}(1) \leq F_{l|\pi_2}^{-1}(1), \quad i \in B_{\pi_1}, \ l \in B_{\pi_2}.$$

From each $Y_i^{(2)} \mid \pi_1$, construct the truncated lifetimes $X_i^{(2)} \mid \pi_1$ to represent the time beyond x_1, and define the second failure time as

$$X_*^{(2)} = \min\{X_i^{(2)} \mid \pi_1 : i \in B_{\pi_1}\}.$$

This procedure is repeated until the nth failure is observed. Formally, the random variables $X_*^{(1)}, X_*^{(2)}, \dots, X_*^{(n)}$ with corresponding failure sources $C_1, C_2, \dots, C_n$ are called the ESOS in a system of size n if their joint density is

$$f_{X_*^{(n)}, C_n, \dots, X_*^{(1)}, C_1}(x_n, c_n, \dots, x_1, c_1) = \prod_{k=1}^{n} \left(\frac{f_{c_k|\pi_{k-1}}(x_k)}{\bar{F}_{c_k|\pi_{k-1}}(x_{k-1})} \prod_{j \in B_{\pi_k}} \frac{\bar{F}_{j|\pi_{k-1}}(x_k)}{\bar{F}_{j|\pi_{k-1}}(x_{k-1})} \right), \tag{1.1}$$

where $x_0 < x_1 < \cdots < x_n$ and $\bar{F}_j(x_0) = 1$ for all $j \in \{1, 2, \dots, n\}$ denotes the SF of component j at time x_0.

Note that, in general, the number of parameters involved in the model is quite large. Therefore, the complexity of the model and the estimation procedure grow quickly. To reduce the number of parameters, one needs additional assumptions, such as the history independence assumption or the conditional proportional hazard rate assumption. In the history independence assumption, the load change in surviving components depends on the number of failed components only, i.e.,

$$\alpha_{j|\pi_{k-1}} = \alpha_{j,k}, \quad j \in B_{\pi_{k-1}}, \quad k \in \{1, 2, \ldots, n\}.$$

In the conditional proportional hazard rate assumption, all component CDFs are expressed using unknown, absolutely continuous, and strictly increasing baseline distributions F_j^*, i.e.,

$$F_{j,k} = 1 - \left(1 - F_j^*\right)^{\alpha_{j,k}},$$

with positive real numbers $\alpha_{j,k}$ for component $j \in \{1, 2, \ldots, n\}$.

Under the conditional proportional hazard rate assumption, hazard rate corresponding to the CDF $F_{j,k}$ is proportional to that of the baseline CDF F_j^*. From Eq. (1.1), the joint PDF of ESOS becomes

$$f_{X_*^{(n)},C_n,\ldots,X_*^{(1)},C_1}(x_n, c_n, \ldots, x_1, c_1) = \prod_{k=1}^{n} \left(\alpha_{c_k,k} \frac{f_{c_k}^*(x_k)}{\bar{F}_{c_k}^*(x_k)} \prod_{j \in B_{\pi_{k-1}}} \left(\frac{\bar{F}_j^*(x_k)}{\bar{F}_j^*(x_{k-1})} \right)^{\alpha_{j,k}} \right).$$

For more details, see Pesch et al. [45].

1.3.6 Accelerated Failure Time Models

According to an AFT model, two HFs $r_0(\cdot)$ and $r_1(\cdot)$ are related by

$$r_1(t) = \frac{1}{\beta} r_0 \left(\frac{t}{\beta} \right), \quad t > 0,$$

where $\beta > 0$ is known as the acceleration factor, and $r_0(t)$ is the baseline HF. The corresponding reliability functions $S_0(t)$ and $S_1(t)$ are related by

$$S_1(t) = S_0 \left(\frac{t}{\beta} \right), \quad t > 0.$$

In this model, if $\beta > 1$, it implies that $S_1(t) \geq S_0(t)$, indicating that the reliability of the system increases under S_1 compared to S_0, and if $\beta < 1$, then $S_1(t) \leq S_0(t)$.

Consider a k-out-of-m system with $1 \leq k \leq m$. Let $X_{(i)}$ denote the failure time of i-th failed component in the system, for $i = 1, 2, \ldots, (m - k + 1)$. That is, $X_{(i)}$ represents the minimum of the failure times among the remaining $(m - i + 1)$ components, and the system fails at time $X_{(m-k+1)}$. Assume that the lifetimes of the m components are iid with common CDF $F(\cdot)$, reliability function $\bar{F}(\cdot)$, PDF $f(\cdot)$, and HF $r(\cdot)$. Then, the relationship between the HF of $X_{(j+1)}$ given that $X_{(j)} = s$ and the HF $r(t)$ is given by

$$r_{X_{(j+1)}|X_{(j)}=s}(t) = (m-j)r(t),\ t \geq s,\ j = 1, 2, \ldots, (m-k).$$

To capture the load-sharing effect, Sutar and Naik-Nimbalkar [56] introduce the following AFT models.

If the HF $r(t)$ is increasing in t, then

$$r_{X_{(j+1)}|X_{(j)}=s}(t) = \frac{(m-j)}{\beta} r(\frac{t}{\beta}),\quad t \geq s,\ j = 1, 2, \ldots, (m-k),\ \beta > 0. \tag{1.2}$$

If $r(t)$ is increasing in t, then

$$r_{X_{(j+1)}|X_{(j)}=s}(t) = \beta(m-j) r(\frac{t}{\beta}),\quad t \geq s,\ j = 1, 2, \ldots, (m-k),\ \beta > 0. \tag{1.3}$$

For the conditional HF given in Eq. (1.2), the joint PDF of $\left(X_{(1)}, X_{(2)}, \ldots, X_{(m-k+1)}\right)$ is derived by Sutar and Naik-Nimbalkar [56] and is given by

$$g\left(x_1, x_2, \ldots, x_{m-k+1}\right) = \gamma_1 B(x_1) \left\{ \bar{F}\left(\frac{x_{m-k+1}}{\beta}\right)\right\}^{(k-1)} f(x_1) \prod_{i=1}^{m-k} f\left(\frac{x_{i+1}}{\beta}\right),$$
$$0 < x_1 < x_2 < \ldots < x_{m-k+1}, \tag{1.4}$$

where $\beta > 0$, $\gamma_1 = \frac{m!}{(k-1)!\beta^{(m-k)}}$, $B(x_1) = \left\{\frac{\bar{F}(x_1)}{\bar{F}(\frac{x_1}{\beta})}\right\}^{(m-1)}$.
For the conditional HF given in Eq. (1.3), the joint PDF of $\left(X_{(1)}, X_{(2)}, \ldots, X_{(m-k+1)}\right)$ is given by

$$g\left(x_1, x_2, \ldots, x_{m-k+1}\right) = \gamma_2 \times f(x_1) \left[\bar{F}(x_1)\right]^{(m-1)} f\left(\frac{x_{m-k+1}}{\beta}\right) \frac{\left\{\bar{F}\left(\frac{x_{m-k+1}}{\beta}\right)\right\}^{(k\beta^2-1)}}{\left\{\bar{F}\left(\frac{x_1}{\beta}\right)\right\}^{(m-1)\beta^2}}$$
$$\times \prod_{i=1}^{m-k-1} \frac{f\left(\frac{x_{i+1}}{\beta}\right)}{\left\{\bar{F}\left(\frac{x_{i+1}}{\beta}\right)\right\}^{(1-\beta^2)}},\ 0 < x_1 < x_2 < \cdots < x_{m-k+1}, \tag{1.5}$$

where $\beta > 0$, $\gamma_2 = \frac{\beta^{(m-k)} m!}{(k-1)!}$.
The joint density functions of the ordered random variables given in Eqs. (1.4) and (1.5) are special cases of the joint density functions of sequential order statistics, as introduced by Kamps [30], and Cramer and Kamps [16].

1.3.7 Park's Model

Consider a J-component system connected in parallel. Assume that the lifetimes of the working components are iid at each failure step. At the start, when all components are initially working, the parameter vector of J components associated with the underlying lifetime distribution of each component is labeled as $\boldsymbol{\theta}^{(0)}$. After j failures, there are $J-j$ surviving components remaining in the system and the parameter vector of the remaining surviving components changes from $\boldsymbol{\theta}^{(j-1)}$ to $\boldsymbol{\theta}^{(j)}$ for $j = 0, 1, \ldots, J-1$. Denote the hypothetical latent lifetimes of the $J-j$ surviving components by $U_1, U_2, \ldots, U_{J-j}$, where $j = 0, 1, \ldots, J-1$. The next failure (i.e., the $(j+1)$-st failure) occurs when the weakest component fails. Thus, denote the shortest lifetime among all the $J-j$ surviving components by $Y^{(j)}$. The random variable $Y^{(j)}$ in the system is realized after j components have failed. For notational purposes, assume that the component indexed by J failed first, and the one indexed by $J-1$ next, and so on. Therefore, using the traditional latent-variable approach, we can write

$$Y^{(j)} = \min\{U_1^{(j)}, U_2^{(j)}, \ldots, U_{J-j}^{(j)}\}. \tag{1.6}$$

Note that the lifetime of the system can be expressed as $\sum_{j=0}^{J-1} Y^{(j)}$. It follows from Eq. (1.6) that

$$P(Y^{(j)} > y) = P\left[\min\{U_1^{(j)}, U_2^{(j)}, \ldots, U_{J-j}^{(j)}\} > y\right] = \prod_{\ell=1}^{J-j} S_{U_\ell}^{(j)}(y), \tag{1.7}$$

where $S_{U_\ell}^{(j)}(y) = P(U_\ell^{(j)} > y)$ denotes the SF of each hypothetical latent random variable with its corresponding PDF $f_{U_\ell}^{(j)}(\cdot)$, for $\ell = 1, 2, \ldots, J-j$; $j = 0, 1, \ldots, J-1$. It immediately follows from Eq. (1.7) that the CDF of $Y^{(j)}$ is

$$F_Y^{(j)}(y) = P(Y^{(j)} \le y) = 1 - \prod_{\ell=1}^{J-j} S_{U_\ell}^{(j)}(y). \tag{1.8}$$

Differentiating the CDF in Eq. (1.8) with respect to y, we get the PDF of $Y^{(j)}$ as

$$f_Y^{(j)}(y) = \prod_{\ell=1}^{J-j} S_{U_\ell}^{(j)}(y) \times \sum_{\ell=1}^{J-j} h_{U_\ell}^{(j)}(y),$$

where $h_{U_\ell}^{(j)}(y) = \frac{f_{U_\ell}^{(j)}(y)}{S_{U_\ell}^{(j)}(y)}$ is the corresponding HF.

1.4 Methods of Inference

1.4.1 Maximum Likelihood Estimation

Let $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ be a sample of n independent observations from a PDF $f(\boldsymbol{x} \mid \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the parameter vector of interest. The likelihood function is defined as

$$L(\boldsymbol{\theta} \mid \boldsymbol{x}) = \prod_{i=1}^{n} f(x_i \mid \boldsymbol{\theta}).$$

Therefore, the log-likelihood function is given by

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{x}) = \log L(\boldsymbol{\theta} \mid \boldsymbol{x}) = \sum_{i=1}^{n} \log f(x_i \mid \boldsymbol{\theta}).$$

The MLE of $\boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\theta}}$, is the value of $\boldsymbol{\theta}$ that maximizes the likelihood (or log-likelihood) function given by

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta} \mid \boldsymbol{x}) = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta} \mid \boldsymbol{x}).$$

Expectation-Maximization Algorithm

The EM algorithm is a general iterative approach for computing the MLE of parametric models when closed-form solutions are not available and the data are incomplete. Introduced by Dempster et al [18], the EM algorithm elegantly addresses the complications arising from missing or partially observed data by decomposing a difficult problem into two simpler, alternating steps.

Suppose that we are interested in estimating a p-dimensional parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_p)$ based on the observed data $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$. Let $\boldsymbol{z} = (z_{m+1}, z_{m+2}, \cdots, z_n)$ be the missing unobserved data vector and $\boldsymbol{x} = (\boldsymbol{y}, \boldsymbol{z})$ denote the complete data vector. The complete-data likelihood function of $\boldsymbol{\theta}$ is the likelihood function of $\boldsymbol{\theta}$ based on complete data $\boldsymbol{x}$ and can be expressed as

$$L_c(\boldsymbol{\theta}|\boldsymbol{x}) = f(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{\theta})g(\boldsymbol{z}|\boldsymbol{\theta}),$$

where $f(\cdot|\boldsymbol{z}, \boldsymbol{\theta})$ is the conditional PDF of $\boldsymbol{y}$ given $\boldsymbol{z}$ and $g(\cdot|\boldsymbol{\theta})$ is the marginal PDF of $\boldsymbol{z}$.

The EM algorithm is composed of two steps, called E-step and M-step, that run iteratively until convergence is obtained. Let $\boldsymbol{\theta}_s$ denote the parameter estimate at s-th iteration. In E-step, of $(s + 1)$-st iteration one need to find $Q(\boldsymbol{\theta}|\boldsymbol{\theta}_s)$, the conditional expectation of logarithm of the complete-data likelihood with respect to missing data $\boldsymbol{z}$

given observed data $\boldsymbol{y}$ and current value of parameter vector $\boldsymbol{\theta}_s$. Thus,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}_s) = \mathbb{E}_{\boldsymbol{z}|\boldsymbol{y},\boldsymbol{\theta}_s}\Big[\log L_c(\boldsymbol{\theta}|\boldsymbol{x})\Big] = \int [\log L_c(\boldsymbol{\theta}|\boldsymbol{x})]\, h(\boldsymbol{z}|\boldsymbol{y},\boldsymbol{\theta}_s)\, d\boldsymbol{z},$$

where $h(\cdot|\boldsymbol{y},\boldsymbol{\theta}_s)$ is the conditional PDF of $\boldsymbol{z}$ given $\boldsymbol{y}$ at $\boldsymbol{\theta} = \boldsymbol{\theta}_s$. In M-step of $(s+1)$-st iteration, $Q(\boldsymbol{\theta}|\boldsymbol{\theta}_s)$ is maximized with respect to $\boldsymbol{\theta}$ to obtain updated value of parameter vector $\boldsymbol{\theta}_{s+1}$. Therefore,

$$\boldsymbol{\theta}_{s+1} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}_s).$$

These two steps are iterated until convergence, usually determined by the relative change in parameter estimates or log-likelihood values falling below a pre-specified threshold. In the context of the present study, the EM algorithm has been adapted to handle the incomplete data structure inherent in the load-sharing system model. The general framework above provides an intuitive basis before moving to the model-specific derivations. For more details on EM algorithm, the readers are referred to the excellent monograph by McLachlan et al. [36]

1.4.2 Confidence Interval Based on the Likelihood Theory

Under suitable regularity conditions and for a large sample size n, the distribution of MLE $\hat{\boldsymbol{\theta}}$ can be approximated by a multivariate normal distribution with mean $\boldsymbol{\theta_0}$, the true value of $\boldsymbol{\theta}$ and a variance-covariance matrix that can be estimated using the observed Fisher information matrix as follows.

$$V = \widehat{\mathrm{Var}}\left(\hat{\boldsymbol{\theta}}\right) = \left[I_n\left(\hat{\boldsymbol{\theta}}\right)\right]^{-1}, \text{ where } I_n\left(\hat{\boldsymbol{\theta}}\right) = -\frac{\partial^2}{\partial \boldsymbol{\theta}^2}\ell(\boldsymbol{\theta} \mid \boldsymbol{x})\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}},$$

where V is the observed Fisher information matrix evaluated at the MLE $\hat{\boldsymbol{\theta}}$.

An asymptotic $100(1-\alpha)\%$ CI for θ_j, the j-th component of $\boldsymbol{\theta}$, is then given by

$$\hat{\theta}_j \pm z_{1-\alpha/2}\,\mathrm{SE}\left(\hat{\theta}_j\right),$$

where $\mathrm{SE}\left(\hat{\theta}_j\right) = \sqrt{V_{jj}}$, V_{jj} is the j-th diagonal of V, and $z_{1-\alpha/2}$ is the $(1-\frac{\alpha}{2})$ quantile of the standard normal distribution.

Louis' Missing Information Principle

Louis' missing information principle provided by Louis [34] expresses the observed Fisher information using derivatives and conditional expectations with respect to the complete-data log-likelihood. The observed Fisher information matrix, expressed using conditional expectations and variances (Louis' principle), is

$$\begin{aligned}\mathcal{I}_{\text{obs}}(\boldsymbol{\theta}) = \;& \mathbb{E}\left[-\frac{\partial^2 \ell_c(\boldsymbol{\theta} \mid \boldsymbol{x})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \mid \boldsymbol{y}, \boldsymbol{\theta}\right] \\ & - \left(\mathbb{E}\left[\mathbf{S}_c(\boldsymbol{\theta})\mathbf{S}_c(\boldsymbol{\theta})^\top \mid \boldsymbol{y}, \boldsymbol{\theta}\right] - \mathbb{E}\left[\mathbf{S}_c(\boldsymbol{\theta}) \mid \boldsymbol{y}, \boldsymbol{\theta}\right] \mathbb{E}\left[\mathbf{S}_c(\boldsymbol{\theta}) \mid \boldsymbol{y}, \boldsymbol{\theta}\right]^\top\right), \end{aligned} \tag{1.9}$$

where $\ell_c(\boldsymbol{\theta} \mid \boldsymbol{x})$ is the complete-data log-likelihood function of $\boldsymbol{\theta}$ based on complete data $\boldsymbol{x}$, $\mathbb{E}[\cdot \mid \boldsymbol{y}, \boldsymbol{\theta}]$ denotes conditional expectation with respect to the conditional distribution of $\boldsymbol{z}$ given $\boldsymbol{y}$ under parameter $\boldsymbol{\theta}$, $\mathbf{S}_c(\boldsymbol{\theta}) = \frac{\partial \ell_c(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \boldsymbol{\theta}}$ is the score vector (gradient) and $\mathbf{H}_c(\boldsymbol{\theta}) = \frac{\partial^2 \ell_c(\boldsymbol{\theta}|\boldsymbol{x})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^\top}$ is the Hessian matrix (second derivative).

The Eq. (1.9) interprets the observed information as the complete-data information minus the missing information owing to the unobserved data. Importantly, this formulation allows one to compute standard errors for MLEs achieved via the EM algorithm using only complete-data derivatives, without directly differentiating the observed-data likelihood.

The estimated asymptotic covariance matrix is $V\hat{a}r(\hat{\boldsymbol{\theta}}) \approx \mathcal{I}_{\text{obs}}^{-1}(\hat{\boldsymbol{\theta}})$, where $\hat{\boldsymbol{\theta}}$ is EM estimate of $\boldsymbol{\theta}$. The SE for each component $\hat{\theta}_j$ is the square root of the j-th diagonal element of $V\hat{a}r(\hat{\boldsymbol{\theta}})$, $j = 1, 2, \ldots, p$. An approximate $100(1-\alpha)\%$ CI for θ_j is

$$\left[\hat{\theta}_j - z_{\alpha/2}\, \text{SE}(\hat{\theta}_j), \quad \hat{\theta}_j + z_{\alpha/2}\, \text{SE}(\hat{\theta}_j)\right],$$

where $z_{\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution and $\text{SE}(\hat{\theta}_j)$ is the standard error.

1.4.3 Parametric Bootstrap Confidence Interval

The bootstrap is a simulation-based method for assessing the uncertainty of parameter estimates under the assumption that the specified statistical model is correct. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ denote the observed data, and let $\hat{\boldsymbol{\theta}}$ be the MLE of the p-dimensional parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_p)$ for the assumed parametric model $f(\boldsymbol{x} \mid \boldsymbol{\theta})$. The procedure begins by fitting the model to the observed data to obtain $\hat{\boldsymbol{\theta}}$. Using this fitted model, B bootstrap samples $\boldsymbol{x}_1^*, \boldsymbol{x}_2^*, \ldots, \boldsymbol{x}_B^*$, each of size n, are generated by simulating from $f(\boldsymbol{x} \mid \hat{\boldsymbol{\theta}})$. For each bootstrap sample $\boldsymbol{x}_s^*$, the parameter estimate $\hat{\boldsymbol{\theta}}_s^*$, for $s = 1, 2, \ldots, B$, is computed using the same estimation method as for the original data. The resulting set of bootstrap estimates $\{\hat{\boldsymbol{\theta}}_s^*\}_{s=1}^B$ approximates the sampling distribution of the estimator.

A $100(1-\alpha)\%$ percentile parametric bootstrap CI for the jth component of θ_j is obtained as

$$\left[Q_{\alpha/2}(\hat{\theta}_{j1}^*, \ldots, \hat{\theta}_{jB}^*), \quad Q_{1-\alpha/2}(\hat{\theta}_{j1}^*, \ldots, \hat{\theta}_{jB}^*)\right],$$

where $Q_q(\cdot)$ denotes the empirical q-th quantile. Alternatively, a normal-approximation parametric bootstrap CI can be constructed using the SE of the bootstrap estimate, which is defined as

$$\mathrm{SE_b}(\hat{\theta}_j) = \sqrt{\frac{1}{B-1}\sum_{s=1}^{B}\left(\hat{\theta}^*_{js} - \bar{\hat{\theta}}^*_j\right)^2},$$

and the bootstrap bias is defined as

$$bias_b(\widehat{\theta}_j) = \bar{\hat{\theta}}^*_j - \widehat{\theta}_j,$$

where

$$\bar{\hat{\theta}}^*_j = \frac{1}{B}\sum_{s=1}^{B}\hat{\theta}^*_{js}, \ j = 1, 2, \ldots, p.$$

The normal-based $100(1-\alpha)\%$ bootstrap CI for θ_j is then given by

$$\left[\hat{\theta}_j - z_{\alpha/2}\,\mathrm{SE_b}(\hat{\theta}_j), \quad \hat{\theta}_j + z_{\alpha/2}\,\mathrm{SE_b}(\hat{\theta}_j)\right],$$

and $100(1-\alpha)\%$ bootstrap CI after correcting the bias for θ_j is then given by

$$\left[\hat{\theta}_j - bias_b(\widehat{\theta}_j) - z_{\alpha/2}\,\mathrm{SE_b}(\hat{\theta}_j), \quad \hat{\theta}_j - bias_b(\widehat{\theta}_j) + z_{\alpha/2}\,\mathrm{SE_b}(\hat{\theta}_j)\right],$$

where $\hat{\theta}_j$ is the estimate from the original data and $z_{\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

1.4.4 Bayesian Inference

In a Bayesian framework, statistical inference is treated as a process of learning from data. We begin with a prior belief about the model parameters, which represents our knowledge (or uncertainty) before observing the data. Once the data are observed, these beliefs are updated through Bayes' theorem, producing a posterior distribution that combines both prior information and evidence from the data. Thus, the posterior may be viewed as a compromise between what was believed before and what is suggested by the data. From a practical point of view, the posterior distribution contains all the information required for decision making and uncertainty quantification. Instead of providing only a single point estimate, Bayesian analysis describes the entire range of plausible parameter values and their probabilities. This is particularly meaningful in reliability analysis of load-sharing systems, where the dependence between components and the uncertainty of model parameters play a crucial role. For instance, when a component fails and the remaining component carries additional load, the exact strength or lifetime parameters cannot be known with certainty. The Bayesian approach allows this uncertainty to be naturally incorporated and updated as more evidence becomes available.

Suppose the observed data are denoted by $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ and the parameter vector by $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_p)$, where p is the number of parameters. The prior distribution for $\boldsymbol{\theta}$ is given by $\pi(\boldsymbol{\theta})$, which encodes prior beliefs about the parameters before observing the data. The likelihood function is $L(\boldsymbol{\theta} \mid \boldsymbol{x})$.

The posterior distribution of $\boldsymbol{\theta}$ given the data $\boldsymbol{x}$ is then

$$\pi(\boldsymbol{\theta} \mid \boldsymbol{x}) = \frac{L(\boldsymbol{\theta} \mid \boldsymbol{x})\,\pi(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta}^* \mid \boldsymbol{x})\pi(\boldsymbol{\theta}^*)\,d\boldsymbol{\theta}^*}.$$

The BE of $\boldsymbol{\theta}$ under squared error loss is the posterior mean is given by

$$\hat{\boldsymbol{\theta}}_{\text{Bayes}} = \mathbb{E}[\boldsymbol{\theta} \mid \boldsymbol{x}] = \int \boldsymbol{\theta}\,\pi(\boldsymbol{\theta} \mid \boldsymbol{x})\,d\boldsymbol{\theta},$$

and the posterior variance quantifies uncertainty is defined by

$$\text{Var}(\boldsymbol{\theta} \mid \boldsymbol{x}) = \int \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{Bayes}}\right)^2 \pi(\boldsymbol{\theta} \mid \boldsymbol{x})\,d\boldsymbol{\theta}.$$

A $100(1-\alpha)\%$ credible interval for the jth component θ_j of the parameter vector $\boldsymbol{\theta}$ is an interval $[l_j(\boldsymbol{x}), u_j(\boldsymbol{x})]$ such that

$$\Pr\left(l_j(\boldsymbol{x}) \le \theta_j \le u_j(\boldsymbol{x}) \mid \boldsymbol{x}\right) \ge 1 - \alpha,$$

where the probability is taken with respect to the marginal posterior distribution $\pi(\theta_j \mid \boldsymbol{x})$. In general, it's difficult to find the closed form of the BE, so computation using MCMC is an alternative method to find the estimates.

1.5 Literature Review

The load-sharing system has been studied by numerous authors since 1940s. Daniels [17] considered the load-sharing model for studying the reliability of a fiber composite material. Afterwards, Coleman [12, 13] and Birnbaum and Saunders [9] adopted this model for analyzing the strength behavior of fiber bundles. Rosen [49] and Phoenix [47] analyzed strength and fatigue in composite materials, expanding applications to aerospace and civil engineering. Harlow and Phoenix [27, 28] developed the chain-of-bundles probability model for fibrous materials. Ross [50], Durham et al. [21], and Schechner [51] extended these concepts to more general statistical inference and dynamic load-sharing contexts.

Parametric models rely on specified distributions for component lifetimes. Kim and Kvam [31] considered parametric inference in a multi-component reliability system. Their work considered the situation where the underlying lifetime distribution of each of the components in the system is exponential and an unknown load-sharing rule is applied. Accelerated testing models are also used to draw inferences about load-sharing parameters by Mettas and Vassiliou [37], Amari and Bergman [1], and Amari and Pham [3]. In 2005, Park and Padgett [42] studied the strength distribution of multi-modal failures. They provided a parameter estimation procedure for Weibull, lognormal, and inverse Gaussian distributions using the EM algorithm. They have also given some examples to illustrate their results. Amari et al. [2] provided an overview of load-sharing systems and the relationship between the load and the failure behavior of a component. Singh et al. [54] executed the classical and Bayesian analysis of the estimation problem of parameters of a k-components load-sharing parallel system in which some of the components follow a constant failure-rate and the remaining follow a linearly increasing failure-rate.

In 2010, Park [40] considered a multi-component parallel load-sharing system. A closed-form MLE and BUE are provided under a general load-sharing rule when the underlying lifetime distribution of the components in the system is exponential or Weibull. In 2012, Singh and Gupta [52] in their study dealt with the classical and Bayesian estimation of the parameters of a k-component load-sharing parallel system model in which each component's lifetime follows a Lindley distribution. In 2013, Park [41] considered a multiple-component parallel load-sharing system and modeled the system with the assumption of the existence of underlying hypothetical latent random variables. The underlying lifetime distribution of the components was taken to be lognormal or normal. A methodology is given to obtain the MLEs using the EM algorithm. Sutar and Naik-Nimbalker [56] proposed accelerated failure time models for load-sharing systems. Zhao et al. [59] provided reliability modeling and analysis of load-sharing systems with continuously degrading components.

Generally, the models discussed above do not consider how long a surviving component has worked before the redistribution of the workload; these types of load-sharing models are referred to as memoryless. Wang et al. [58] proposed a general framework for load-sharing models that account for the work history. Park et al. [43] introduced a methodology that integrates the conventional procedure under the assumption of the load-sharing system being made up of fundamental hypothetical latent random variables. They developed an EM algorithm for performing the maximum likelihood estimation of the system with Lindley-distributed component lifetimes. Asha et al. [5] presented frailty-based models capturing dependence among components. Franco et al. [22] proposed the GFB model for load-sharing, increasing flexibility for heterogeneous systems. Recently, Sutar et al. [57] proposed a reliability analysis of load-sharing systems under an unequal load-sharing rule. Models based on sequential order statistics were considered for load-sharing problems by Kamps [30], Cramer and Kamps [15], and Burkschat et al. [10], among others. Recently, Pesch et al. [46] modeled failure risks in a load-sharing system with heterogeneous components. Pesch et al. [45] utilized ESOS to model component lifetimes of a load-sharing system and used a link function to reduce the dimension of the parametric space.

Non-parametric approaches do not specify a form for component or system lifetimes. Kvam and Peña [33] introduced a distribution-free method for estimating load-sharing system reliability. Miyakawa [39] developed nonparametric estimators under incomplete and competing risks data.

Semi-parametric techniques balance between parametric and nonparametric. Balakrishnan et al. [6] applied order statistics from heterogeneous populations to load-sharing, while Mies and Bedbur [38] developed exact inference and model selection for these systems.

Recent research on Bayesian analysis of load-sharing systems has seen significant advances. In a Bayesian parametric context, Singh and Gupta [53] provided Bayesian estimation strategies under censoring. Singh et al. [54] executed Bayesian estimation for k-component load-sharing systems, showcasing Bayesian advantages in uncertainty quantification. Singh and Gupta [52] deal with the Bayesian estimation of the parameters of a k-component load-sharing parallel system model in which each component's lifetime follows a Lindley distribution. Recently, Maurya et al. [35] provided detailed classical and Bayesian approaches for analyzing k-out-of-n load-sharing systems under progressive censoring. These studies illustrate the growing application and scope of Bayesian inference in reliability analysis of load-sharing systems.

1.6 Use of R Software

All simulations, data analyses, and graphical outputs presented in this thesis are conducted using the R software; for details, please see R Core Team [48]. R is a free and open-source platform that provides extensive capabilities for statistical modeling, simulation, and data analysis. For each problem across all chapters, R scripts are developed to generate simulated datasets, fit models, and compute relevant statistical estimates. The software is also utilised to produce graphs, including diagnostic plots, trace plots.

The `rstan` [55] package serves as an interface between R and `Stan`, a state-of-the-art platform for statistical modeling and high-performance Bayesian inference. `Stan` implements advanced Markov Chain Monte Carlo algorithms, such as the No-U-Turn sampler, which is an adaptive form of Hamiltonian Monte Carlo. This enables efficient sampling even in models with high-dimensional and correlated parameter spaces. The use of `rstan` facilitates automatic differentiation, robust convergence diagnostics, and flexible model specification through a probabilistic programming framework. In this thesis, `rstan` is used for Bayesian computation because it can efficiently handle complex hierarchical models, provides automatic differentiation for likelihood computations, and supports diagnostics and convergence checks. Several R packages are also used.

1.7 Organization of the Thesis

In Chapter 2, we discuss Bayesian inferential methods for the GFB model, proposed by Franco et al. [22], assuming independent gamma priors on model parameters. The GFB model is a family of models that includes many well-known parametric models as its special cases. Therefore, the selection of the best model for a given dataset is of interest. We study the selection of the best model in the GFB family using WAIC. Our simulation results suggest that the performance of WAIC for model selection is quite satisfactory. We analyzed two real data sets for illustration - one corresponding to a two-motor load-sharing system and the other related to nuclear reactors, both representing two-component parallel systems.

In Chapter 3, we develop a doubly-flexible model for two-component load-sharing systems, when an unobserved shared random effect is present between two components of a system. We then discuss the statistical analysis of a dataset obtained from load-sharing systems using the proposed model. In the proposed model, we use the GFB distribution for the baseline of the component lifetimes, and the GG family of distributions to incorporate an unobserved random effect that captures dependence between the component lifetimes. Note that this model is doubly-flexible, as both for the baseline as well as for the shared frailty, we use general families of distributions, GFB and GG, respectively. We call the model as GFB-GG model. Thus, the GFB-GG model structure leads to a two-way general class of models for the two-component load-sharing systems. For any given data on a two-component load-sharing system, one can fit the proposed model and then choose suitable distributions for the baseline as well as the shared frailty, within the respective families of distributions used for the baseline and the shared frailty. Fitting methods for the proposed model, based on direct optimization and an EM-type algorithm, are discussed in this chapter. Through simulations, the effectiveness of the model fitting and selection methods is demonstrated. A simulation case and analysis of a real dataset are presented to illustrate the strength of the proposed model. We also estimate some important quantities, such as RMT, MTTF, and MRT of the underlying lifetime distribution of load-sharing systems under GFB-GG model. These quantities are of practical importance for making various strategies and plans.

In Chapter 4, we develop a flexible model for analyzing data on multi-component load-sharing systems, without using restrictive assumptions on the distribution of component lifetimes. More specifically, we develop a model involving PLAs of the CHFs, capturing the unknown load-share rule at successive stages of component failures. Usually, parametric analyses of load-sharing systems use a specific lifetime distribution to model component lifetimes for the entire range. In contrast, the proposed PLA-based modeling approach does not require such strong parametric assumptions for component lifetime distributions. The fitting of the proposed PLA-based model is discussed using a likelihood-based method. To implement a PLA-based model, one needs to specify a set of cut-points. An objective method of selecting the position and number of cut-points is discussed. We perform detailed simulation studies to judge the efficacy of proposed model fitting and selection methods. We also discuss the estimation of some important quantities such as RMT, quantile function, MTTF, and MRT of the underlying lifetime distribution of load-sharing systems.

In Chapter 5, we conclude the thesis by summarizing the work done. Also, we briefly discuss some future directions for research in this chapter.

CHAPTER 2

BAYESIAN INFERENCE FOR GFB MODEL

One of the useful parametric models for component lifetimes of a two-component load-sharing system is the FBE model proposed by Freund [23], for details please see Section 1.3.1. In Asha et al. [4], a generalization of Freund's model was proposed. They consider the baseline reliability functions before and after failure belong to a common PFR class with different PFR parameters, please see Section 1.3.2. In Franco et al. [22], a further generalization has been done, where they have considered that the baseline reliability function changes after failure of one of the components due to the extra load on the surviving component. They named the model as GFB model, please see Section 1.3.3 for more details.

Bayesian inference offers several important advantages in the analysis of load-sharing systems. Bayesian methods yield full posterior distributions for each parameter. Moreover, Bayesian estimation performs well in small sample sizes, which are common in load-sharing data due to practical constraints on data collection. By incorporating prior information, Bayesian analysis provides a flexible framework for improving estimation stability. Also, posterior predictive distributions offer a natural way to assess future reliability behavior based on observed data. In this chapter, we consider Bayesian analysis of GFB model.

The main contributions of this chapter are as follows.

- We derive the joint posterior distribution of the model parameters considering independent gamma priors for the GFB model of the lifetimes of a two-component load-sharing system.
- We discuss the model selection procedure using WAIC within the GFB model.
- A simulation study is conducted to check the efficacy of the fitting and model selection method. The proposed methods are found to be quite satisfactory.

- Two real bivariate data sets are analyzed for illustrative purposes.

The rest of this chapter is organized as follows. In Section 2.1, we discuss the load-sharing model and Bayesian inference for the GFB distribution. We provide model selection criteria in the Section 2.2. In Section 2.3, a simulation study is presented for checking the accuracy of the fitting method and the model selection criteria. Also, two real datasets on two-component parallel systems are analyzed in Section 2.4. Finally, the chapter is concluded in Section 2.5.

2.1 Model Description

Consider n two-component load-sharing systems, with observed component lifetimes as $Data = \{(y_{1i}, y_{2i}),\ i = 1, 2, \ldots, n\}$. Let $I_1 = \{i : y_{1i} > y_{2i}\}$, $|I_1| = n_1$ and $I_2 = \{i : y_{1i} < y_{2i}\}$, $|I_2| = n_2$. Note that $n_1 + n_2 = n$. Let random vector (Y_{1i}, Y_{2i}) be the lifetime of i-th two-component load-sharing system. Here, we assume that

$$(Y_{1i}, Y_{2i}) \overset{\text{iid}}{\sim} GFB(R_B, R_B^*, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6).$$

2.1.1 Prior Assumptions

We assume independent gamma priors for each parameter θ_i with shape parameter $a_i > 0$ and scale parameter $b_i > 0$. Therefore, the joint prior density on $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$ is given by

$$\pi_1(\Theta|a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5, a_6, b_6) = \prod_{i=1}^{6} \frac{1}{b_i^{a_i}\Gamma(a_i)} \theta_i^{a_i-1} e^{-\frac{\theta_i}{b_i}},\ \theta_i > 0, \quad \text{for all } i = 1, 2, \ldots, 6. \tag{2.1}$$

2.1.2 Posterior Analysis

The likelihood function of parameter vector Θ for the GFB model, based on observed data, is given by

$$\mathbb{L}(\Theta|\text{Data}) = \prod_{i\in I_1} \theta_2\theta_3 r_B^*(y_{1i}, \theta_6) r_B(y_{2i}, \theta_5) \left[\frac{R_B^*(y_{1i}, \theta_6)}{R_B^*(y_{2i}, \theta_6)}\right]^{\theta_3} [R_B(y_{2i}, \theta_5)]^{(\theta_1+\theta_2)} \times \prod_{i\in I_2} \theta_1\theta_4 r_B(y_{1i}, \theta_5) r_B^*(y_{2i}, \theta_6) \left[\frac{R_B^*(y_{2i}, \theta_6)}{R_B^*(y_{1i}, \theta_6)}\right]^{\theta_4} [R_B(y_{1i}, \theta_5)]^{(\theta_1+\theta_2)}. \tag{2.2}$$

From the prior assumption Eq. (2.1) and the likelihood function Eq. (2.2), we can obtain the posterior distribution of Θ as

$$\tilde{\pi}(\Theta|\text{Data}) \propto L(\Theta|\text{Data})\pi_1(\Theta|a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5, a_6, b_6).$$

The BE of some parametric function $g(\Theta)$ with respect to squared error loss is given by

$$\begin{aligned}\hat{g}_{\text{Bayes}}(\Theta) &= \int_{\Theta} g(\Theta)\tilde{\pi}\,(\Theta|\text{Data})\,d\Theta \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty g(\Theta)\tilde{\pi}\,(\Theta|\text{Data})\,d\theta_1 d\theta_2 d\theta_3 d\theta_4 d\theta_5 d\theta_6. \qquad (2.3)\end{aligned}$$

Note that, in general, the integration in Eq. (2.3) is difficult to obtain analytically. Therefore, we perform an MCMC-based technique to obtain BEs and credible intervals for the parametric functions. We obtain a sequence of N posterior samples $\{\Theta^{(k)}\}_{k=1}^N$ using MCMC methods. The BEs of each parameter θ_i, $i = 1, 2, \ldots, 6$ with respect to squared error loss, after discarding the initial b burn-in samples, are given by

$$\hat{\theta}_{i(B)} = \frac{1}{N-b}\sum_{k=b+1}^{N} \theta_i^{(k)},$$

and the posterior variances are given by

$$V_{\text{post}}(\theta_i) = \frac{1}{N-b}\sum_{k=b+1}^{N} \left(\theta_i^{(k)} - \hat{\theta}_{i(B)}\right)^2,$$

where $\theta_i^{(k)}$ is the sampled value of θ_i at iteration k. For each parameter $i = 1, 2, \ldots, 6$, the $100(1-\alpha)\%$ credible interval is

$$\left[Q_{\alpha/2}(\theta_i^{(b+1)}, \ldots, \theta_i^{(N)}), \quad Q_{1-\alpha/2}(\theta_i^{(b+1)}, \ldots, \theta_i^{(N)})\right],$$

where $Q_q(\cdot)$ denotes the qth empirical quantile of the sampled values.

In Bayesian computation using `rstan` described in Section 1.6, two widely used diagnostics are the potential scale reduction factor (*Rhat*) and the effective sample size (ESS), both of which are automatically computed after sampling. The *Rhat* statistic measures the ratio of the variance between multiple Markov chains to the variance within each chain, thereby assessing convergence. A value of *Rhat* close to 1 indicates that the chains have mixed well and converged to the target posterior distribution and considered to be satisfactory for convergence. The effective sample size (ESS) quantifies the amount of independent information contained in the correlated MCMC draws by accounting for the autocorrelation between samples. A higher ESS implies better sampling efficiency and more reliable posterior summaries. These diagnostics, computed automatically by `rstan`, were used throughout the thesis to assess and confirm convergence of all Bayesian models.

2.2 Model Selection

For a given two-component load-sharing dataset, it is important to find the best model within the GFB family, i.e., the best choice of baseline distributions in that framework. Since many different baseline distributions are possible, there will be several candidate models. After fitting each candidate model to the data, the best one can be chosen using WAIC, a fully Bayesian model selection criterion. WAIC estimates the out-of-sample expectation by adding lppd and a penalty term for the effective number of parameters. The addition of this penalty term guards against overfitting.

The lppd is defined by

$$\text{lppd} = \log\left(\prod_{i=1}^{n}\int p(\boldsymbol{y}_i \mid \Theta)p_{\text{post}}(\Theta)d\Theta\right) = \sum_{i=1}^{n}\log\left(\int p(\boldsymbol{y}_i \mid \Theta)p_{\text{post}}(\Theta)d\Theta\right),$$

where $p(\boldsymbol{y}_i \mid \Theta)$ is the conditional PDF of $\boldsymbol{y}_i = (y_{1i}, y_{2i})$ given Θ, and $p_{\text{post}}(\Theta)$ is the posterior PDF of Θ. The lppd can be estimated using MCMC draws as follows.

$$\widehat{\text{lppd}} = \sum_{i=1}^{n}\log\left(\frac{1}{N-b}\sum_{s=b+1}^{N}p(\boldsymbol{y}_i \mid \Theta^{(s)})\right).$$

after discarding the initial b burn-in samples. Denoting the posterior variance by $\text{Var}_{\text{post}}(\cdot)$, a measure of the penalty term is given by

$$p_{\text{WAIC}} = \sum_{i=1}^{n}\text{Var}_{\text{post}}(\log p(\boldsymbol{y}_i \mid \Theta)),$$

which can be estimated by

$$\hat{p}_{\text{WAIC}} = \sum_{i=1}^{n}V_{s=1}^{N}\left(\log p(\boldsymbol{y}_i \mid \Theta^{(s)})\right)$$

where $V_{s=1}^{N}(a_s) = \frac{1}{N-1}\sum_{s=1}^{N}(a_s - \bar{a})^2$, and $\bar{a} = \frac{1}{N}\sum_{s=b+1}^{N}a_s$.

Therefore, the estimated WAIC is given by

$$\text{WAIC} = -2\left(\widehat{\text{lppd}} - \hat{p}_{\text{WAIC}}\right).$$

For a given dataset, one can fit all candidate models and estimate WAIC values for each model. The lower value of WAIC is preferred, as a lower value indicates better out-of-sample predictive fit. Thus, the model with the lowest value of WAIC is chosen as the best model among all candidate models. The interested readers are referred to Gelman et al. [25] for more details of different model selection criteria, including WAIC.

2.3 Simulation Study

The simulation study serves two main purposes:

1. It checks the performance of the described method in Section 2.1 with respect to the ABE, MSE, AL, and CP for all parameters;

2. It investigates the performance of the model selection criteria discussed in Section 2.2 for choosing the best model within GFB class of models for a given two-component load-sharing data.

For the simulation study, we take nine different models, taking different combinations of baseline distributions. Here, we indicate them according to their baseline distributions before and after the first failure. We provide the details in Table 2.1.

2.3.1 Demonstration of the Model Fitting Method

For each of the 9 baseline combinations, the following steps are performed across 1000 replications:

1. Simulate two-component load-sharing data using specified model and simulation parameters (all parameters are selected without loss of generality);

2. Estimate the model parameters via the fitting procedure described in Section 2.1, treating the parent model as the true model;

3. Compute the ABE and MSE for the parameter estimates;

4. For each parameter, from the credible intervals, compute both the CP and AL of these intervals.

Four different sample sizes n = 25, 50, 75, 100 are used separately in the simulation studies. First, the prior means are shifted from the true parameter values to assess the robustness of the posterior estimates under mild prior misspecification and in this scenario we take Gamma(0.5, 0.5) priors for all the parameters for all nine models. The outcomes are presented in Tables 2.2, 2.3, 2.4, and 2.5. From these tables, we can see that the average estimates of all the parameters are very close to the true values with small MSE values. The values of AL and CP are also quite satisfactory. The results indicate that the fitting method performs well for the proposed model. We have examined the case where the prior means are centered at the true parameter values keeping the same variance as the shifted mean scenario for Model B1 to evaluate the impact of a well-specified prior. In this case we take priors Gamma(0.5, 2), Gamma(0.5, 1), Gamma(0.5, 0.67) and Gamma(0.5, 0.5) for θ_1, θ_2, θ_3 and θ_4. The outcomes are presented in Tables 2.6. The results indicate that centering the prior improves estimation precision and convergence stability.

2.3.2 Study of Model Selection

For model selection in this context, our objective is to evaluate whether the criteria described in Section 2.2 can accurately identify the parent models corresponding to various combinations of baseline distributions within the GFB model framework. The procedure for conducting the model selection study is as follows:

1. Generate a two-component load-sharing dataset of size n from a parent model, where the parent model is one of the models $B1$ to $B9$, as detailed in Table 2.1;

2. Fit all candidate models $B1$ through $B9$ to the generated data;

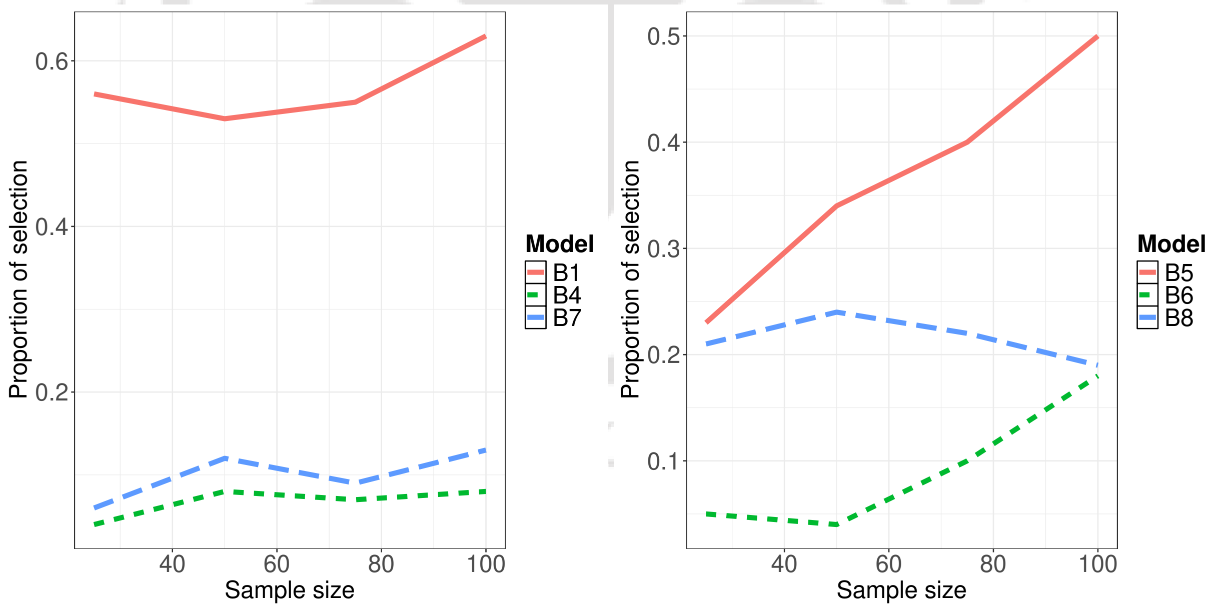
3. Compute the value of WAIC for each fitted candidate model using the estimated parameters;

4. Select the model with the lowest value according to the WAIC as the best-fit model for the given data among the 9 candidates.

For model selection, we generate 100 samples from two different models, B1 and B5, as parent models for different sample sizes $n = 25, 50, 75, 100$ separately. For each generated data point, we fit all the models from B1 to B9. We select those models among the nine models whose WAIC is minimum and count the total number out of 100 generated samples. we take Gamma(0.5, 0.5) priors for all the parameters for all nine models. In Figure 2.1a, we plot the proportion of times models B1, B4, and B7 are selected when the parent model is B1. Note that B1, B4, and B7 are the models that are selected the most number of times, when the parent model is B1. In Figure 2.1b, the proportion of selection of B5, B6, and B8 is shown when the parent model is B5. From the figures, one can notice that if the sample size increases from 25 to 100, the percentage of selection of the parent model increases (when B1 is the parent model, the percentage decreases initially; this may be attributed to sampling error), and for the other two models it decreases. That shows the credibility of our proposed method.

Table 2.1: Model indicators according to their baseline distributions before and after 1st failure

Model Name	Baseline before failure	Baseline after failure
B1	Exp(1)	Exp(1)
B2	Exp(1)	Weibull(θ_6, 1)
B3	Exp(1)	Gamma(θ_6, 1)
B4	Weibull(θ_5, 1)	Exp(1)
B5	Weibull(θ_5, 1)	Weibull(θ_6, 1)
B6	Weibull(θ_5, 1)	Gamma(θ_6, 1)
B7	Gamma(θ_5, 1)	Exp(1)
B8	Gamma(θ_5, 1)	Weibull(θ_6, 1)
B9	Gamma(θ_5, 1)	Gamma(θ_6, 1)



(a) Selection by WAIC for parent model B1 (b) Selection by WAIC for parent model B5

Figure 2.1: Proportion of times the top three models get selected

Table 2.2: ABE, MSE, AL, CP of the parameters based on 1000 simulated samples for size $n = 25$ according to load-sharing systems from different models with GFB parameters $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$

	Sample size			$n = 25$			
Model	Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Name	True Value	0.250	0.500	0.750	1.000	2.000	1.500
	ABE	0.269	0.535	0.805	1.132	-	-
B1	MSE	0.009	0.019	0.046	0.214	-	-
	AL	0.350	0.500	0.760	1.542	-	-
	CP	94.7	95.5	95.3	96.0	-	-
	ABE	0.269	0.535	0.879	1.205	-	1.545
B2	MSE	0.009	0.018	0.144	0.346	-	0.096
	AL	0.350	0.500	1.448	2.161	-	1.181
	CP	95.0	95.4	95.9	95.8	-	94.6
	ABE	0.269	0.535	0.795	1.109	-	1.223
B3	MSE	0.009	0.018	0.106	0.288	-	0.465
	AL	0.350	0.499	1.207	2.087	-	2.436
	CP	94.8	95.2	95.4	95.7	-	93.5
	ABE	0.267	0.530	0.805	1.132	2.068	-
B4	MSE	0.010	0.021	0.046	0.214	0.117	-
	AL	0.364	0.538	0.760	1.542	1.268	-
	CP	95.3	95.5	95.2	96.1	94.2	-
	ABE	0.267	0.530	0.940	1.258	2.068	1.580
B5	MSE	0.010	0.021	0.218	0.414	0.117	0.174
	AL	0.363	0.538	1.947	2.613	1.267	1.643
	CP	95.2	95.6	96.9	96.7	94.2	96.3
	ABE	0.266	0.530	0.778	1.076	2.067	1.102
B6	MSE	0.010	0.021	0.109	0.258	0.117	0.522
	AL	0.364	0.538	1.312	2.178	1.268	2.624
	CP	95.3	95.8	96.0	96.1	94.2	95.8
	ABE	0.307	0.612	0.805	1.132	2.092	-
B7	MSE	0.026	0.078	0.046	0.214	0.227	-
	AL	0.535	0.897	0.760	1.541	1.796	-
	CP	95.5	94.4	95.2	95.9	94.3	-
	ABE	0.307	0.612	0.984	1.314	2.093	1.527
B8	MSE	0.025	0.078	0.275	0.514	0.228	0.087
	AL	0.534	0.897	2.124	2.901	1.793	1.248
	CP	95.5	94.3	97.2	96.9	94.1	97.2
	ABE	0.307	0.612	0.779	1.088	2.093	1.083
B9	MSE	0.025	0.078	0.081	0.244	0.228	0.596
	AL	0.533	0.897	1.149	1.976	1.793	2.838
	CP	95.5	94.4	96.6	96.4	94.7	96.2

Table 2.3: ABE, MSE, AL, CP of the parameters based on 1000 simulated samples for size $n = 50$ according to load-sharing systems from different models with GFB parameters $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$

	Sample size			$n = 50$			
Model	Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Name	True Value	0.250	0.500	0.750	1.000	2.000	1.500
B1	ABE	0.261	0.513	0.775	1.058	-	-
	MSE	0.004	0.008	0.021	0.080	-	-
	AL	0.244	0.345	0.523	1.010	-	
	CP	93.5	94.8	94.2	94.4	-	-
B2	ABE	0.261	0.513	0.809	1.101	-	1.518
	MSE	0.004	0.008	0.054	0.150	-	0.041
	AL	0.244	0.344	0.954	1.424	-	0.829
	CP	93.7	94.5	96.0	94.8	-	96.5
B3	ABE	0.261	0.513	0.759	1.038	-	1.304
	MSE	0.004	0.008	0.049	0.139	-	0.286
	AL	0.244	0.344	0.829	1.404	-	2.016
	CP	93.9	94.6	94.6	95.0	-	93.8
B4	ABE	0.261	0.511	0.775	1.058	2.028	-
	MSE	0.005	0.010	0.021	0.080	0.049	-
	AL	0.256	0.375	0.523	1.010	0.879	-
	CP	94.2	95.0	94.2	94.7	96.2	-
B5	ABE	0.260	0.511	0.855	1.151	2.028	1.527
	MSE	0.005	0.010	0.094	0.210	0.050	0.078
	AL	0.256	0.375	1.307	1.791	0.878	1.185
	CP	94.0	95.0	96.4	96.5	96.4	97.5
B6	ABE	0.260	0.511	0.736	1.006	2.028	1.151
	MSE	0.005	0.010	0.056	0.153	0.049	0.420
	AL	0.255	0.375	0.924	1.534	0.877	2.306
	CP	93.9	94.8	94.4	94.5	96.2	94.5
B7	ABE	0.277	0.543	0.775	1.059	2.040	-
	MSE	0.008	0.022	0.021	0.080	0.105	-
	AL	0.337	0.554	0.523	1.010	1.278	-
	CP	94.3	96.1	94.2	94.4	95.3	-
B8	ABE	0.277	0.543	0.885	1.196	2.041	1.501
	MSE	0.008	0.022	0.119	0.272	0.105	0.044
	AL	0.534	0.897	2.124	2.901	1.793	1.248
	CP	94.4	96.2	97.0	96.4	95.5	97.1
B9	ABE	0.277	0.543	0.730	0.996	2.040	1.089
	MSE	0.008	0.022	0.041	0.115	0.106	0.507
	AL	0.336	0.553	0.796	1.343	1.276	2.470
	CP	94.3	95.8	95.3	94.9	95.4	94.9

Table 2.4: ABE, MSE, AL, CP of the parameters based on 1000 simulated samples for size $n = 75$ according to load-sharing systems from different models with GFB parameters $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$

	Sample size	$n = 75$					
Model	Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Name	True Value	0.250	0.500	0.750	1.000	2.000	1.500
B1	ABE	0.258	0.513	0.768	1.047	-	-
	MSE	0.002	0.006	0.013	0.053	-	-
	AL	0.199	0.282	0.423	0.818	-	-
	CP	96.0	94.1	95.2	94.3	-	-
B2	ABE	0.258	0.513	0.785	1.064	-	1.522
	MSE	0.002	0.006	0.039	0.089	-	0.030
	AL	0.199	0.282	0.755	1.128	-	0.677
	CP	96.2	94.0	94.9	94.5	-	94.6
B3	ABE	0.258	0.513	0.764	1.049	-	1.390
	MSE	0.002	0.006	0.031	0.103	-	0.195
	AL	0.199	0.282	0.680	1.160	-	1.702
	CP	96.0	94.3	94.3	94.3	-	94.9
B4	ABE	0.257	0.512	0.768	1.047	2.020	-
	MSE	0.003	0.006	0.013	0.053	0.033	-
	AL	0.208	0.307	0.423	0.819	0.713	-
	CP	95.8	94.7	95.3	94.6	95.4	-
B5	ABE	0.257	0.512	0.818	1.102	2.020	1.529
	MSE	0.003	0.006	0.067	0.131	0.033	0.059
	AL	0.208	0.307	1.017	1.407	0.713	0.973
	CP	95.5	94.5	95.7	95.2	95.6	95.9
B6	ABE	0.257	0.512	0.741	1.016	2.020	1.238
	MSE	0.003	0.006	0.040	0.119	0.033	0.342
	AL	0.208	0.307	0.781	1.292	0.713	2.102
	CP	95.9	94.7	94.3	93.9	95.4	93.8
B7	ABE	0.269	0.534	0.768	1.047	2.029	-
	MSE	0.005	0.015	0.013	0.053	0.071	-
	AL	0.267	0.443	0.423	0.819	1.041	-
	CP	95.2	94.7	95.1	94.5	95.0	-
B8	ABE	0.269	0.534	0.838	1.129	2.028	1.511
	MSE	0.005	0.015	0.086	0.172	0.071	0.035
	AL	0.267	0.443	1.124	1.581	1.042	0.756
	CP	95.0	94.9	95.2	95.2	95.5	94.7
B9	ABE	0.269	0.535	0.734	1.003	2.029	1.179
	MSE	0.005	0.015	0.029	0.089	0.070	0.426
	AL	0.266	0.443	0.678	1.135	1.040	2.289
	CP	95.1	94.8	94.3	93.7	95.3	92.6

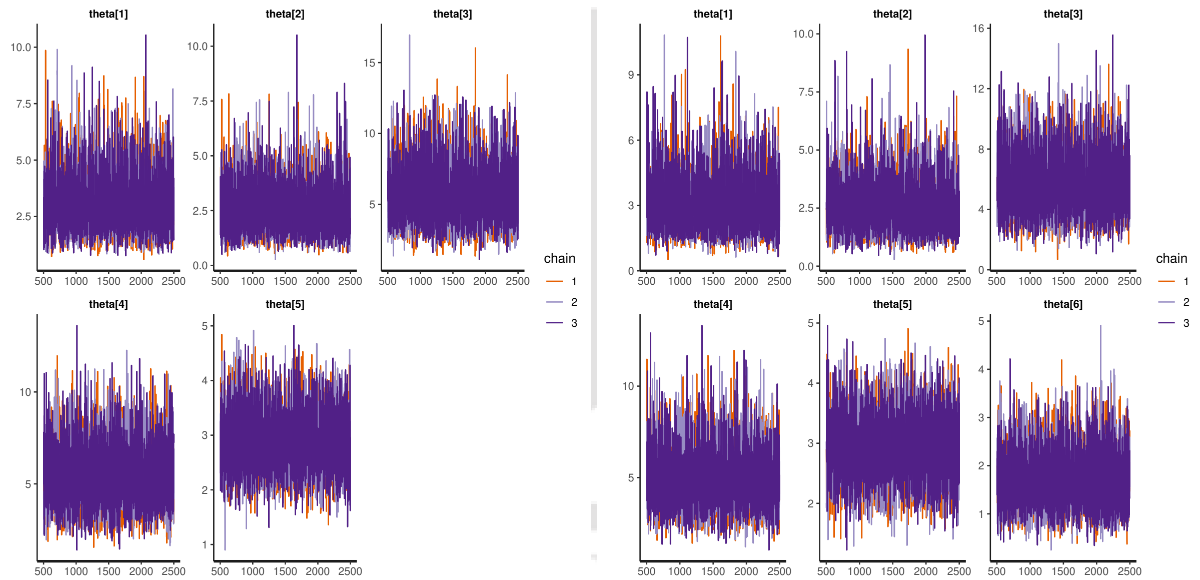
Table 2.5: ABE, MSE, AL, CP of the parameters based on 1000 simulated samples for size $n = 100$ according to load-sharing systems from different models with GFB parameters $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$

	Sample size	$n = 100$					
Model	Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Name	True Value	0.250	0.500	0.750	1.000	2.000	1.500
B1	ABE	0.258	0.513	0.768	1.047	-	-
	MSE	0.002	0.006	0.013	0.053	-	-
	AL	0.199	0.282	0.423	0.818	-	-
	CP	96.0	94.1	95.2	94.3	-	-
B2	ABE	0.258	0.513	0.785	1.064	-	1.522
	MSE	0.002	0.006	0.039	0.089	-	0.030
	AL	0.199	0.282	0.755	1.128	-	0.677
	CP	96.2	94.0	94.9	94.5	-	94.6
B3	ABE	0.258	0.513	0.764	1.049	-	1.390
	MSE	0.002	0.006	0.031	0.103	-	0.195
	AL	0.199	0.282	0.680	1.160	-	1.702
	CP	96.0	94.3	94.3	94.3	-	94.9
B4	ABE	0.257	0.512	0.768	1.047	2.020	-
	MSE	0.003	0.006	0.013	0.053	0.033	-
	AL	0.208	0.307	0.423	0.819	0.713	-
	CP	95.8	94.7	95.3	94.6	95.4	-
B5	ABE	0.257	0.512	0.818	1.102	2.020	1.529
	MSE	0.003	0.006	0.067	0.131	0.033	0.059
	AL	0.208	0.307	1.017	1.407	0.713	0.973
	CP	95.5	94.5	95.7	95.2	95.6	95.9
B6	ABE	0.257	0.512	0.741	1.016	2.020	1.238
	MSE	0.003	0.006	0.040	0.119	0.033	0.342
	AL	0.208	0.307	0.781	1.292	0.713	2.102
	CP	95.9	94.7	94.3	93.9	95.4	93.8
B7	ABE	0.269	0.534	0.768	1.047	2.029	-
	MSE	0.005	0.015	0.013	0.053	0.071	-
	AL	0.267	0.443	0.423	0.819	1.041	-
	CP	95.7	95.1	95.1	95.2	94.5	-
B8	ABE	0.267	0.522	0.822	1.097	2.025	1.501
	MSE	0.004	0.010	0.058	0.118	0.053	0.024
	AL	0.228	0.374	0.960	1.335	0.904	0.653
	CP	95.5	94.6	96.2	96.2	94.9	96.0
B9	ABE	0.266	0.522	0.733	0.983	2.025	1.218
	MSE	0.004	0.010	0.026	0.064	0.053	0.368
	AL	0.228	0.373	0.596	0.969	0.903	2.132
	CP	95.6	94.4	93.8	94.4	94.7	92.8

2.4 Data Analysis

2.4.1 Analysis of Two-motor Data

We present an analysis of the dataset on two-motor data introduced in Subsection 1.2.1, by using the proposed GFB model. All nine models with various baseline distribution combinations are fitted to the data, and model selection within the family has been performed similarly as explained above. The failure times of the motors are scaled before analysis by dividing them by 365. The scale parameter of the baseline distributions is taken to be one for the identifiability of models. The priors are considered to be Gamma(0.5, 0.5). The top two models for the two-motor data, in order of lowest WAIC values, are B5 and B4, and the corresponding parameter estimates are presented in Table 2.7, along with the standard error, and 95% credible intervals. The trace plot of MCMC realizations for models B4 and B5 are given in Figure 2.2. From the figure, it is clear that the chains are mixing well and converge. Moreover, Rhat values for all the parameters are close to 1, indicating convergence of the corresponding chains. The best model, according to WAIC, is the one with a Weibull-Weibull baseline with a WAIC value of -15.8134.

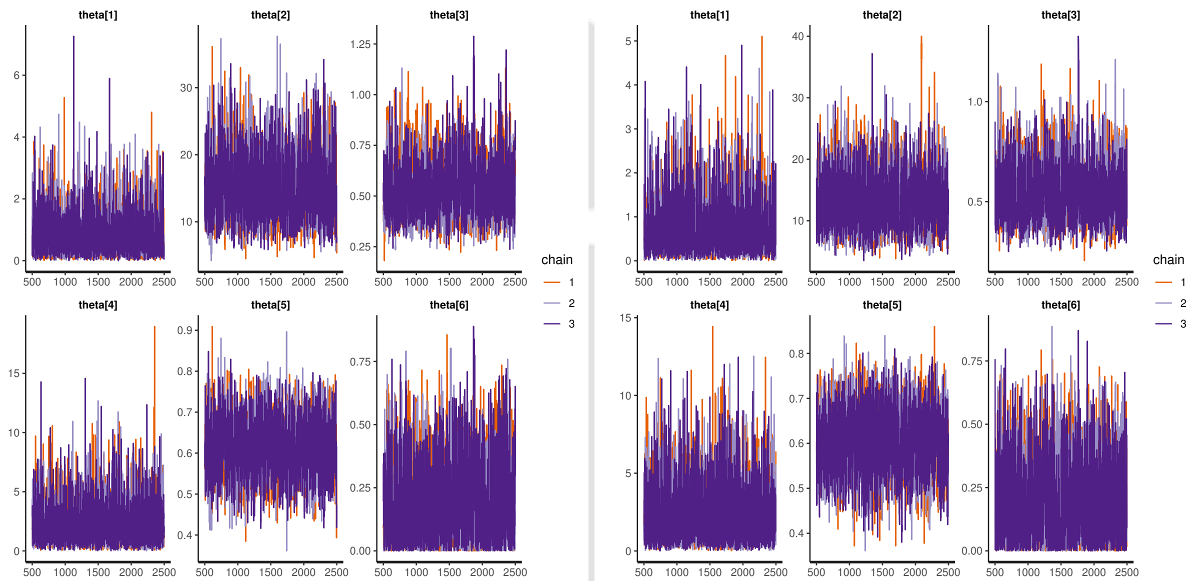


(a) Trace plot of parameters of Model B4 (b) Trace plot of parameters of Model B5

Figure 2.2: Trace plots of parameters of the best two models for two-motor data

2.4.2 Analysis of Nuclear Reactor Data

We conduct an analysis of the nuclear reactor dataset introduced in Subsection 1.2.3, employing the proposed GFB model. Each of the nine models, constructed from different combinations of baseline distributions, was fitted to the data, and model selection within this family was carried out as previously described. Before analysis, the failure times for the nuclear plants were normalized by dividing each value by 365. The scale parameter for each baseline distribution was set to one. We take the prior as Gamma(0.5, 0.5) for all the parameters. Table 2.8 presents the two best-fitting models ranked by the lowest WAIC values for the nuclear reactor data along with the parameter estimates, the standard error of those estimates, and their 95% credible intervals. Figure 2.3 is the trace plot of MCMC realizations of the models B6 and B9. It is clear from the figure that the chains are well-mixed and converge. Also, we see that the Rhat values are close to 1, indicating the satisfactory convergence of the chains. According to the WAIC criterion, the model with the gamma-gamma baseline yields the best fit, achieving the lowest WAIC value among the nine models at -163.6349.



(a) Trace plot of parameters of Model B6 (b) Trace plot of parameters of Model B9

Figure 2.3: Trace plots of parameters of the best two models for nuclear reactor data

2.5 Conclusions

In this chapter, we discuss a Bayesian estimation method for the GFB model of the lifetimes of a two-component load-sharing system. We derive the joint posterior distribution of the model parameters considering independent gamma priors, and implement an MCMC algorithm using `rstan` to obtain Bayesian point estimates and credible intervals. We discuss model selection within the GFB family using WAIC. A simulation study shows that the BEs are quite accurate. Through simulation, we show that WAIC performs quite well in selecting the parent model within the GFB family. Two real load-sharing data sets are analyzed for illustrative purposes. The best-fitted model using WAIC is obtained for both datasets.

Table 2.6: ABE, MSE, AL, CP of the parameters based on 1000 simulated samples for size n = 25, 50, 75, 100 according to load-sharing systems from Model B1 with GFB parameters $\Theta = (\theta_1, \theta_2, \theta_3, \theta_4)$

Sample	Parameters	θ_1	θ_2	θ_3	θ_4
size	True Value	0.250	0.500	0.750	1.000
$n = 25$	ABE	0.256	0.526	0.798	1.132
	MSE	0.008	0.017	0.044	0.214
	AL	0.334	0.492	0.753	1.540
	CP	94.7	95.3	95.4	96.0
$n = 50$	ABE	0.255	0.509	0.772	1.059
	MSE	0.004	0.008	0.020	0.080
	AL	0.239	0.342	0.521	1.010
	CP	93.8	94.6	94.2	94.5
$n = 75$	ABE	0.254	0.510	0.766	1.047
	MSE	0.002	0.005	0.013	0.053
	AL	0.196	0.280	0.422	0.818
	CP	95.9	94.3	95.4	94.5
$n = 100$	ABE	0.255	0.502	0.761	1.024
	MSE	0.002	0.004	0.009	0.033
	AL	0.170	0.240	0.365	0.689
	CP	95.7	94.8	95.2	95.1

Table 2.7: Estimates of the parameters of the top two models based on the motor data

Model	Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	WAIC
B5	BE	3.0606	2.4908	5.5537	4.8612	2.8967	1.7181	-15.8134
	SD	1.2117	1.0511	1.8935	1.5666	0.5257	1.5752	
	LL	1.2792	0.9639	2.5314	2.3019	1.9370	0.7507	
	UL	5.5636	4.9777	9.8845	8.4278	3.9832	2.9565	
	Rhat	1.0000	0.9999	0.9997	0.9997	1.0003	0.9997	
B4	BE	3.0417	2.4605	5.6925	5.3968	2.8822	-	-15.1930
	SD	1.2116	1.0574	1.9507	1.6425	0.5324	-	
	LL	1.2628	0.9318	2.6104	2.6604	1.9279	-	
	UL	5.9856	5.0028	10.1679	9.0828	4.0161	-	
	Rhat	0.9999	0.9998	1.0003	1.0010	0.9997	-	

Table 2.8: Estimates of the parameters of the top two models based on the nuclear reactor data

Model	Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	WAIC
B9	BE	0.6762	12.9490	0.5497	2.2065	0.6027	0.1858	-163.6349
	SD	0.5913	4.3405	0.1467	1.8676	0.0711	0.1642	
	LL	0.0464	6.0768	0.3334	0.1338	0.4609	0.0005	
	UL	2.2676	22.6648	0.8881	7.1268	0.7416	0.5625	
	Rhat	0.9999	0.1000	1.0009	0.9997	1.0002	1.0014	
B6	BE	0.7645	15.0388	0.5415	2.1695	0.6135	0.1790	-162.8100
	SD	0.6590	4.4577	0.1463	1.8143	0.0672	0.1678	
	LL	0.0511	7.8936	0.3195	0.1587	0.4835	0.0002	
	UL	2.5110	25.1077	0.8807	6.9252	0.7460	0.5812	
	Rhat	1.0008	1.0001	1.0058	1.0013	1.0003	1.0037	

CHAPTER 3

A DOUBLY-FLEXIBLE MODEL BASED ON GG FRAILTY

3.1 Introduction

A shared frailty is a random factor incorporated into a model to account for unobserved random effects that may be present in the study units due to various factors, such as commonalities among them. For more details, see the book by Hougaard [29]. In reliability and survival studies, the literature on shared frailty models is well developed; the interested readers are referred to Hanagal [26]. For load-sharing systems, a shared frailty can conveniently incorporate dependence between component lifetimes; refer to the work of Asha et al. [5]. The shared frailty model can also capture the unit-to-unit variation of the study units. However, since a shared frailty is attributed to unobservable quantities, goodness-of-fit for such models might be difficult to assess.

Aiming at improving the quality of model fit, Balakrishnan and Peng [8] proposed the GG frailty model. The GG family of distributions includes exponential, gamma, Weibull, etc., as special members of the family. Therefore, a GG distribution provides a very convenient way to model frailty, as it is possible to choose an appropriate model for the frailty within the GG family for any given data, thus resulting in a better model fit compared to the scenario where only a specific frailty distribution is fitted to the data.

In this work, we present a doubly-flexible model for two-component load-sharing systems. In the proposed model, we use the GFB distribution for the baseline of the component lifetimes, and the GG family of distributions to incorporate a shared frailty that captures dependence between the component lifetimes. We call the model as GFB-GG throughout the chapter. This model is doubly-flexible, as both for the baseline as well as for the shared random effect, we use general families of distributions, GFB and GG, respectively. Thus, the GFB-GG model structure leads to a two-way general class of models for two-component load-sharing systems. For any given data on a two-component load-sharing system, one can fit the proposed model and then choose suitable distributions for the baseline as well as the shared frailty, within the respective families of distributions used for the baseline and the shared frailty. This way, this GFB-GG model structure provides a better fit to two-component load-sharing systems compared to existing models. In summary, among parametric models considered for two-component load-sharing systems in the literature, the proposed model presents a very general case that accommodates load-sharing data with a wide variety of baseline component lifetime distributions and dependence structures.

The main contributions of this chapter are as follows:

- We present a very general class of parametric models with the GFB and GG family of distributions for the baseline component lifetimes and frailty distributions, respectively, for a two-component load-sharing system.

- We describe a model fitting method based on an EM-type algorithm in detail, and extensive simulations are carried out.

- We develop estimation procedures for important reliability characteristics such as MTTF, RMT, and MRT estimates of two-component load-sharing systems under the GFB-GG model.

The rest of the chapter is organized as follows. In Section 3.2, we give details about the GG distribution. Section 3.3 presents the proposed GFB-GG model for load-sharing data, and the fitting method in detail is given in Section 3.4. Results of an elaborate Monte Carlo simulation study, demonstrating the performance of the fitting method and model selection method in this setting, are presented in Section 3.5. Numerical examples are presented in Section 3.6, wherein a simulation case demonstrates the benefit of using the GFB-GG model structure for a two-component load-sharing system. Analysis of a real dataset and MTTF, MRT RMT are also presented in this section for illustrative purposes. Finally, the chapter is concluded in Section 3.7 with some remarks.

3.2 GG Distribution

The GG distribution, denoted by $GG(\theta, k, \beta)$, has the PDF

$$g_Z(z) = \frac{k}{\theta^{k\beta}\Gamma(\beta)} z^{k\beta-1} e^{-(\frac{z}{\theta})^k}, \; z > 0, \tag{3.1}$$

with $\theta > 0$ as the scale parameter, and $k > 0, \beta > 0$ as shape parameters. The mean and variance, respectively, of the distribution are

$$E(Z) = \theta \cdot \frac{\Gamma\left(\beta + \frac{1}{k}\right)}{\Gamma(\beta)} \quad \text{and} \quad Var(Z) = \theta^2 \cdot \left[\frac{\Gamma\left(\beta + \frac{2}{k}\right)}{\Gamma(\beta)} - \left\{ \frac{\Gamma\left(\beta + \frac{1}{k}\right)}{\Gamma(\beta)} \right\}^2 \right].$$

The GG distribution is a family of distributions that includes some well-known distributions, such as the exponential, gamma, Weibull, etc., as special members of the family, and the lognormal distribution as a limiting case. Table 3.1 gives the choices of parameters for the special members of the GG family of distributions. Balakrishnan and Peng [8]

Table 3.1: Special members of the GG family of distributions

Distribution	Parameter choice
Exponential	$k = \beta = 1$
Gamma	$k = 1$
Weibull	$\beta = 1$
Lognormal	$\beta \to 0$

advocated for the use of the GG frailty model, highlighting its flexibility derived from its shape parameters. As the PDF of the GG family of distributions takes various forms depending on the parameter values, a GG frailty model is capable of describing various types of dependence between the units that share the frailty. For more details on the GG frailty model, refer to Balakrishnan and Peng [8].

3.3 Model Description

Suppose that Y_1 and Y_2 are the lifetimes of the components of a two-component load-sharing parallel system. Assume that conditional on the unobserved random variable Z, which we refer to as the shared frailty, the distribution of (Y_1, Y_2) is the GFB distribution, and Z follows the GG distribution. The proposed model can be described in terms of the failure rates of components. Conditional on the unobserved random factor Z, the failure rates of the components when both of them are operational are given by

$$\lambda_{10}(y|z) = z\theta_1 r_B(y, \theta_B), \;\; y > 0,$$

and

$$\lambda_{20}(y|z) = z\theta_2 r_B(y, \theta_B), \;\; y > 0,$$

where $\lambda_{i0}(\cdot|z)$ is the failure rate of the i-th component.

Now, suppose the j-th component fails first at time y_j. Then, conditional on the shared frailty Z and the failure time y_j, the failure rate of the i-th $(i \neq j)$ component is given by

$$\lambda_{ij}(y_i|y_j, z) = z\theta_i^* r_B^*(y_i, \theta_B^*),\ i = 1, 2;\ j = 1, 2.$$

Without loss of generality, the baseline parameters θ_B and θ_B^* is omitted to denote $R_B(\cdot, \theta_B)$, $R_B^*(\cdot, \theta_B^*)$, $r_B(\cdot, \theta_B)$ and $r_B^*(\cdot, \theta_B^*)$ by $R_B(\cdot)$, $R_B^*(\cdot)$, $r_B(\cdot)$ and $r_B^*(\cdot)$ respectively.

Correspondingly, the conditional joint PDF of Y_1 and Y_2 given Z is (see Franco et al. [22] for further details)

$$f((y_1, y_2)|z) = \begin{cases} z^2\theta_1^*\theta_2 r_B^*(y_1) r_B(y_2) \left[\frac{R_B^*(y_1)}{R_B^*(y_2)}\right]^{z\theta_1^*} [R_B(y_2)]^{z(\theta_1+\theta_2)}, & \text{if } y_1 > y_2 > 0 \\ z^2\theta_1\theta_2^* r_B(y_1) r_B^*(y_2) \left[\frac{R_B^*(y_2)}{R_B^*(y_1)}\right]^{z\theta_2^*} [R_B(y_1)]^{z(\theta_1+\theta_2)}, & \text{if } y_2 > y_1 > 0. \end{cases}$$

The hierarchical representation of the proposed model for component lifetimes of a two-component load-sharing system is

$$(Y_1, Y_2)|(Z = z) \sim GFB(R_B, R_B^*, z\theta_1, z\theta_2, z\theta_1^*, z\theta_2^*, \theta_B, \theta_B^*),$$
$$Z \sim GG(\theta, k, \beta).$$

Clearly, the marginal PDF of the two-dimensional lifetime (Y_1, Y_2) can be obtained as

$$\begin{aligned} f(y_1, y_2) &= \int_{z=0}^{\infty} f((y_1, y_2)|z) g_Z(z) dz \\ &= \begin{cases} \theta_1^*\theta_2 r_B^*(y_1) r_B(y_2) \displaystyle\int_{z=0}^{\infty} z^2 \left[\frac{R_B^*(y_1)}{R_B^*(y_2)}\right]^{z\theta_1^*} [R_B(y_2)]^{z(\theta_1+\theta_2)} g_Z(z) dz, & \text{if } y_1 > y_2 > 0 \\ \theta_1\theta_2^* r_B(y_1) r_B^*(y_2) \displaystyle\int_{z=0}^{\infty} z^2 \left[\frac{R_B^*(y_2)}{R_B^*(y_1)}\right]^{z\theta_2^*} [R_B(y_1)]^{z(\theta_1+\theta_2)} g_Z(z) dz, & \text{if } y_2 > y_1 > 0. \end{cases} \end{aligned}$$

It is a requirement for shared frailty models that the mean of the frailty distribution is fixed at one, for the parameters of the frailty distribution to be identifiable. Using this condition here, we have

$$E(Z) = 1 \implies \theta = \frac{\Gamma(\beta)}{\Gamma(\beta + \frac{1}{k})}. \tag{3.2}$$

Using Eq. (3.2) in Eq. (3.1), the PDF of the frailty distribution becomes

$$g_Z(z) = \frac{ka^\beta}{\Gamma(\beta)} z^{k\beta-1} e^{-az^k},\ z > 0,$$

where $a = \left(\frac{\Gamma(\beta+\frac{1}{k})}{\Gamma(\beta)}\right)^k > 0$.

The lognormal distribution is a limiting case of the GG distribution when $\beta \to 0$. For lognormal frailty, after reparameterization using the identifiability condition of setting the mean at one, the PDF of the frailty Z becomes

$$g_Z(z) = \frac{1}{z\sigma\sqrt{2\pi}} e^{-\frac{(\log(z)+\frac{\sigma^2}{2})^2}{2\sigma^2}}, \; z > 0, \; \sigma > 0.$$

3.4 Model Fitting Method

Consider n two-component load-sharing systems, with observed component lifetimes as Data $= \{(y_{1i}, y_{2i}), \; i = 1, 2, \ldots, n\}$. The unconditional likelihood function for the proposed model, based on the observed data, is given by

$$\mathbb{L}(\Theta|\text{Data}) = \prod_{i=1}^{n} \int_0^\infty f((y_{1i}, y_{2i})|z_i) g_Z(z_i) dz_i,$$

where $\Theta = (\Theta_1, \Theta_2)$, with $\Theta_1 = (\theta_1, \theta_2, \theta_1^*, \theta_2^*, \theta_B, \theta_B^*)$ and $\Theta_2 = (k, \beta)$, represents the vector of model parameters to be estimated. The corresponding log-likelihood function is

$$\begin{aligned} \log \mathbb{L}(\Theta|\text{Data}) &= \sum_{i=1}^{n} \log \left(\int_0^\infty f((y_{1i}, y_{2i})|z_i) g_Z(z_i) dz_i \right) \\ &= \sum_{i \in I_1} \log \left(\theta_1^* \theta_2 r_B^*(y_{1i}) r_B(y_{2i}) \right) + \sum_{i \in I_2} \log \left(\theta_1 \theta_2^* r_B(y_{1i}) r_B^*(y_{2i}) \right) \\ &\quad + \sum_{i \in I_1} \log \left(\int_0^\infty z_i^2 \left[\frac{R_B^*(y_{1i})}{R_B^*(y_{2i})} \right]^{z_i \theta_1^*} [R_B(y_{2i})]^{z_i(\theta_1+\theta_2)} g_Z(z_i) dz_i \right) \\ &\quad + \sum_{i \in I_2} \log \left(\int_0^\infty z_i^2 \left[\frac{R_B^*(y_{2i})}{R_B^*(y_{1i})} \right]^{z_i \theta_2^*} [R_B(y_{1i})]^{z_i(\theta_1+\theta_2)} g_Z(z_i) dz_i \right), \end{aligned} \tag{3.3}$$

where $I_1 = \{i : y_{1i} > y_{2i}\}$, and $I_2 = \{i : y_{1i} < y_{2i}\}$.

The MLEs can be found by maximizing the log-likelihood function given in Eq. (3.3). It can be done using any statistical software. For example, in `R`, one can use the `optim` function. One can use the bootstrap method for the construction of CIs as discussed in Section 1.4.3.

In general, finding the MLEs of the model parameters is a challenging task as it involves an eight-dimensional numerical optimization of a function without an explicit expression. In such cases, the EM algorithm is a very useful tool that offers a reliable method for estimation. Here, we apply an EM-type algorithm wherein we approximate the required conditional expectations by their Monte Carlo estimates, generating samples from suitable conditional distributions. We also discuss the construction of CI using Louis' missing information principle as described in Section 1.4.2.

3.4.1 Implementation of the EM-type Algorithm

For a sample of n two-component load-sharing systems, the hypothetical complete data would consist of the observed data $\{(y_{1i}, y_{2i}), i = 1, 2, \ldots, n\}$ along with shared frailty $\{z_i : i = 1, 2, \ldots, n\}$. In reality, of course, $\{z_i : i = 1, 2, \ldots, n\}$ are not observed, and thus can be considered as the missing data. Therefore, augmenting $\{z_i : i = 1, 2, \ldots, n\}$ with the observed data, the pseudo-complete data would be of the form

$$\text{Data} = \{(y_{1i}, y_{2i}, z_i), i = 1, 2, \ldots, n\}.$$

Based on the pseudo-complete data, the pseudo-complete likelihood function for the proposed model is

$$\mathbb{L}_c(\Theta|\text{Data}) = \prod_{i=1}^{n} f((y_{1i}, y_{2i})|z_i) g_Z(z_i)$$

with the corresponding pseudo-complete log-likelihood function

$$\begin{aligned}
\log \mathbb{L}_c(\Theta|\text{Data}) &= \sum_{i=1}^{n} \log(f((y_{1i}, y_{2i})|z_i) g_Z(z_i)) \\
&= \sum_{i=1}^{n} \log(f((y_{1i}, y_{2i})|z_i)) + \sum_{i=1}^{n} \log(g_Z(z_i)) \\
&\quad + \theta_1^* \sum_{i \in I_1} z_i \log\left(\frac{R_B^*(y_{1i})}{R_B^*(y_{2i})}\right) + (\theta_1 + \theta_2) \sum_{i \in I_1} z_i \log(R_B(y_{2i})) + \sum_{i \in I_2} \log(r_B(y_{1i})) \\
&\quad + \sum_{i \in I_2} \log(r_B^*(y_{2i})) + \theta_2^* \sum_{i \in I_2} z_i \log\left(\frac{R_B^*(y_{2i})}{R_B^*(y_{1i})}\right) + (\theta_1 + \theta_2) \sum_{i \in I_2} z_i \log(R_B(y_{1i})) \\
&\quad + n \log\left(\frac{k a^{\beta}}{\Gamma(\beta)}\right) + (k\beta - 1) \sum_{i=1}^{n} \log(z_i) - a \sum_{i=1}^{n} z_i^k. \qquad (3.4)
\end{aligned}$$

Here, the expected log-likelihood function is

$$E\left(\log \mathbb{L}_c(\Theta|\text{Data})\right) = \int_{z=0}^{\infty} \log \mathbb{L}_c(\Theta|\text{Data}) f_{Z|(Y_1,Y_2)}(\boldsymbol{z}|(\boldsymbol{y_1}, \boldsymbol{y_2})) dz,$$

where $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$, $\boldsymbol{y_1} = (y_{11}, y_{12}, \ldots, y_{1n})$ and $\boldsymbol{y_2} = (y_{21}, y_{22}, \ldots, y_{2n})$.

The conditional expectations to be calculated here are $E(Z|(Y_1, Y_2))$, $E(\log(Z)|(Y_1, Y_2))$, and $E(Z^k|(Y_1, Y_2))$, which cannot be analytically evaluated. Therefore, we use a method wherein we generate samples from the conditional distribution of $Z|(Y_1, Y_2)$, and using the samples, calculate Monte Carlo estimates of the conditional expectations, as described below.

We generate N samples of $\{z_i : i = 1, 2, \ldots, N\}$, from the conditional distribution of $Z|(Y_1, Y_2)$. The conditional PDF $f(z|(y_1, y_2))$ is given by

$$f_{Z|(Y_1,Y_2)}(z|(y_1,y_2)) = \begin{cases} B_1 z^{k\beta+1} e^{-az^k - b_1 z}, & \text{if } y_1 > y_2 > 0 \\ B_2 z^{k\beta+1} e^{-az^k - b_2 z}, & \text{if } y_2 > y_1 > 0, \end{cases}$$

where

$$\begin{aligned} B_1 &= \frac{ka^\beta}{D_1}, \\ B_2 &= \frac{ka^\beta}{D_2}, \\ A_1 &= \left[\frac{R_B^*(y_1)}{R_B^*(y_2)}\right]^{\theta_1^*} [R_B(y_2)]^{(\theta_1+\theta_2)}, \\ A_2 &= \left[\frac{R_B^*(y_2)}{R_B^*(y_1)}\right]^{\theta_2^*} [R_B(y_1)]^{(\theta_1+\theta_2)}, \\ D_1 &= \int_0^\infty z^2 A_1^z \cdot g_Z(z)dz, \\ D_2 &= \int_0^\infty z^2 A_2^z \cdot g_Z(z)dz, \\ b_1 &= -\log(A_1) > 0 \text{ and } b_2 = -\log(A_2) > 0. \end{aligned}$$

For $a,\ \beta,\ z > 0$, we have $e^{-az^k} \leq 1$. Therefore,

$$\begin{aligned} f_{Z|(Y_1,Y_2)}(z|(y_1,y_2)) &\leq \begin{cases} B_1 z^{k\beta+1} e^{-b_1 z}, & \text{if } y_1 > y_2 > 0 \\ B_2 z^{k\beta+1} e^{-b_2 z}, & \text{if } y_2 > y_1 > 0 \end{cases} \\ &\leq \begin{cases} C_1 \cdot \frac{b_1^{k\beta+2}}{\Gamma(k\beta+2)} z^{(k\beta+2)-1} e^{-b_1 z}, & \text{if } y_1 > y_2 > 0 \\ C_2 \cdot \frac{b_2^{k\beta+2}}{\Gamma(k\beta+2)} z^{(k\beta+2)-1} e^{-b_2 z}, & \text{if } y_2 > y_1 > 0 \end{cases} \\ &\leq \begin{cases} C_1 \cdot h_1(z), & \text{if } y_1 > y_2 > 0 \\ C_2 \cdot h_2(z), & \text{if } y_2 > y_1 > 0, \end{cases} \end{aligned}$$

where $C_\ell = \frac{B_\ell \Gamma(k\beta+2)}{b_\ell^{k\beta+2}}$ and $h_\ell(z)$ is the PDF of a Gamma random variable with shape parameter $(k\beta + 2)$ and rate parameter b_ℓ for $\ell = 1,\ 2$. Finally, using the acceptance-rejection method (see Devroy [20]), we can generate samples $z_1, z_2, \cdots, z_N$ from $f_{Z|(Y_1,Y_2)}$. Using the generated sample, the Monte Carlo estimate of the expectation of any function

$h(\cdot)$ of Z can be approximated as

$$E(h(Z)) = \frac{1}{N}\sum_{i=1}^{N} h(z_i).$$

In the M-step, it is convenient to obtain expressions for the PFR parameters $\theta_1, \theta_2, \theta_1^*$, and θ_2^*, by differentiating Eq. (3.4) with respect to the parameters, and equating them to zero:

$$\begin{aligned}
\hat{\theta}_1 &= \frac{-n_2}{\sum_{i\in I_1} E(Z_i|(Y_{1i},Y_{2i}))\log(R_B(y_{2i})) + \sum_{i\in I_2} E(Z_i|(Y_{1i},Y_{2i}))\log(R_B(y_{1i}))},\\
\hat{\theta}_2 &= \frac{-n_1}{\sum_{i\in I_1} E(Z_i|(Y_{1i},Y_{2i}))\log(R_B(y_{2i})) + \sum_{i\in I_2} E(Z_i|(Y_{1i},Y_{2i}))\log(R_B(y_{1i}))},\\
\hat{\theta}_1^* &= \frac{-n_1}{\sum_{i\in I_1} E(Z_i|(Y_{1i},Y_{2i}))\log\left(\frac{R_B(y_{1i})}{R_B(y_{2i})}\right)},\\
\hat{\theta}_2^* &= \frac{-n_2}{\sum_{i\in I_2} E(Z_i|(Y_{1i},Y_{2i}))\log\left(\frac{R_B(y_{2i})}{R_B(y_{1i})}\right)},
\end{aligned} \tag{3.5}$$

where $|I_1|= n_1,\ |I_2|= n_2,$ and $|I|= n_1+n_2 = n$. Note that, $\hat{\theta}_1,\ \hat{\theta}_2,\ \hat{\theta}_1^*,\ \hat{\theta}_2^*$ are functions of θ_B and θ_B^*. Using Eq. (3.5) in Eq. (3.4) gives the profile log-likelihood function (up to an additive constant) of $\tilde{\Theta}_1 = (\theta_B, \theta_B^*)$ and Θ_2 as

$$\begin{aligned}
p(\tilde{\Theta}_1, \Theta_2) &= n_1\log(\hat{\theta}_1^*\hat{\theta}_2) + \sum_{i\in I_1}\log(r_B^*(y_{1i})) + \sum_{i\in I_1}\log(r_B(y_{2i}))\\
&\quad + n_2\log(\hat{\theta}_1\hat{\theta}_2^*) + \sum_{i\in I_2}\log(r_B(y_{1i})) + \sum_{i\in I_2}\log(r_B^*(y_{2i}))\\
&\quad + n\log\left(\frac{ka^\beta}{\Gamma(\beta)}\right) + (k\beta-1)\sum_{i=1}^{n} E(\log(Z_i|(Y_{1i},Y_{2i}))) - a\sum_{i=1}^{n} E((Z_i|(Y_{1i},Y_{2i}))^k)\\
&= H_1(\tilde{\Theta}_1) + H_2(\Theta_2),
\end{aligned}$$

where

$$\begin{aligned}
H_1(\tilde{\Theta}_1) &= n_1\log(\hat{\theta}_1^*\hat{\theta}_2) + \sum_{i\in I_1}\log(r_B^*(y_{1i})) + \sum_{i\in I_1}\log(r_B(y_{2i}))\\
&\quad + n_2\log(\hat{\theta}_1\hat{\theta}_2^*) + \sum_{i\in I_2}\log(r_B(y_{1i})) + \sum_{i\in I_2}\log(r_B^*(y_{2i})),\\
H_2(\Theta_2) &= n\log\left(\frac{ka^\beta}{\Gamma(\beta)}\right) + (k\beta-1)\sum_{i=1}^{n} E(\log(Z_i|(Y_{1i},Y_{2i}))) - a\sum_{i=1}^{n} E((Z_i|(Y_{1i},Y_{2i}))^k).
\end{aligned}$$

Maximization of $p(\tilde{\Theta}_1, \Theta_2)$ is relatively simpler, as it is the sum of two two-dimensional functions. Therefore, $H_1(\tilde{\Theta}_1)$ and $H_2(\Theta_2)$ can be separately maximized to obtain MLEs $\hat{\theta}_B, \hat{\theta}^*_B$, and $\hat{k}, \hat{\beta}$, respectively, using which in Eq. (3.5) finally gives the MLEs $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}^*_i$, and $\hat{\theta}^*_2$.

3.4.2 Model Selection

For a given two-component load-sharing data, it is of prime importance to select the best model within the GFB-GG structure, i.e., the best combination of the baseline and frailty distributions within the GFB and GG families, respectively. Clearly, there is a large number of candidates for the best baseline-frailty combination within the GFB-GG model structure. When all the candidate models are fitted to the given data, the best model can be selected with the help of one of the standard model selection criteria.

Consider model M with any particular combination of baseline and frailty distribution within the GFB-GG structure, involving d_M number of parameters to be estimated. Let n denote the size of the available sample for model fitting. Suppose $\hat{L}_M$ is the maximized value of the likelihood function for model M. Then, the AIC, BIC, AICc, and BC for the model M are the following.

$$\begin{aligned}
\text{AIC}_M &= 2d_M - 2\ln(\widehat{L}_M),\\
\text{BIC}_M &= d_M \cdot \log(n) - 2\ln(\widehat{L}_M),\\
\text{AICc}_M &= \text{AIC} + \frac{2d_M(d_M+1)}{n-d_M-1},\\
\text{BC}_M &= n^{2/3}\sum_{m=1}^{d_M}\frac{1}{k} - 2\ln(\widehat{L}_M).
\end{aligned}$$

Among all the candidate models, the best model for the given data is the one with the minimum AIC (or BIC, AICc, BC) value. Note that BC is developed to combine the strengths of AIC and BIC. In the parametric scenario, BC attains the properties of BIC, and in the nonparametric scenario, it attains the same as AIC adaptively. For details on statistical model selection, see Claeskens [11].

3.5 Simulation Study

The Monte Carlo simulations conducted in this study serve two main purposes:

1. It demonstrates the performance of the model fitting method as described in Section 3.4, with respect to fitting characteristics such as AE, MSE;

2. It examines, for a given two-component load-sharing data, the performance of different model selection criteria for choosing the best model with a baseline-frailty

combination within the GFB-GG model structure.

Within the GFB-GG model structure, we consider 27 different combinations of baseline and frailty distributions for the simulations; details of these combinations are presented in Table 3.2.

3.5.1 Demonstration of the Model Fitting Method

For each of the 27 baseline-frailty combinations, the broad steps as listed below are followed, based on 500 Monte Carlo runs:

1. Generate two-component load-sharing data with a specific set of model parameters and simulation parameters (where all parameters are chosen without any loss of generality);
2. Estimate the model parameters by using the fitting method of Section 3.4, assuming the parent model (i.e., the baseline-frailty combination used for data generation) as the true model;
3. Compute AE and MSE for the parameter estimates;
4. For each parameter, construct CIs and compute CP and AL of the CIs.

The sample size used for simulations is 100. The results are furnished in Tables 3.3, 3.4, and 3.5, for different frailty distributions. It is clear that the fitting method works quite well for the proposed model, as the AE, MSE, CP, and AL for all the model parameters are satisfactory. Clearly, it can be expected that for larger sample sizes, the performance of the fitting method would improve. The convergence of the stochastic EM algorithm was monitored through the stability of parameter estimates across iterations. Specifically, a large number of iterations (here, 1000) was performed, and convergence was visually assessed by plotting the parameter trajectories. As the estimates stabilized after a few iterations, the final estimates were obtained by averaging the values over the last 500 iterations (from the 501st to the 1000th). Alternatively, one may compute a running average of the estimates after a burn-in period and declare convergence when the change between two consecutive iterations falls below a small pre-specified threshold.

3.5.2 Study of Model Selection

For model selection, in this setting, our interest is to examine whether the standard criteria can accurately detect the parent models with various combinations of baseline-frailty distributions from the GFB-GG model structure. As for the model selection criteria, we use the AIC, BIC, AICc, and BC. The steps in which we conduct the model selection study are as follows.

Table 3.2: Model indicators according to their frailty distributions and baseline distributions before and after 1st failure

Model Name	Frailty distribution	Baseline before failure	Baseline after failure
M1	Exp(1)	Exp(1)	Exp(1)
M2	Exp(1)	Exp(1)	Weibull(θ_B^*, 1)
M3	Exp(1)	Exp(1)	Gamma(θ_B^*, 1)
M4	Exp(1)	Weibull(θ_B, 1)	Exp(1)
M5	Exp(1)	Weibull(θ_B, 1)	Weibull(θ_B^*, 1)
M6	Exp(1)	Weibull(θ_B, 1)	Gamma(θ_B^*, 1)
M7	Exp(1)	Gamma(θ_B, 1)	Exp(1)
M8	Exp(1)	Gamma(θ_B, 1)	Weibull(θ_B^*, 1)
M9	Exp(1)	Gamma(θ_B, 1)	Gamma(θ_B^*, 1)
M10	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Exp(1)	Exp(1)
M11	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Exp(1)	Weibull(θ_B^*, 1)
M12	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Exp(1)	Gamma(θ_B^*, 1)
M13	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Weibull(θ_B, 1)	Exp(1)
M14	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Weibull(θ_B, 1)	Weibull(θ_B^*, 1)
M15	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Weibull(θ_B, 1)	Gamma(θ_B^*, 1)
M16	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Gamma(θ_B, 1)	Exp(1)
M17	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Gamma(θ_B, 1)	Weibull(θ_B^*, 1)
M18	Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$)	Gamma(θ_B, 1)	Gamma(θ_B^*, 1)
M19	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Exp(1)	Exp(1)
M20	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Exp(1)	Weibull(θ_B^*, 1)
M21	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Exp(1)	Gamma(θ_B^*, 1)
M22	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Weibull(θ_B, 1)	Exp(1)
M23	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Weibull(θ_B, 1)	Weibull(θ_B^*, 1)
M24	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Weibull(θ_B, 1)	Gamma(θ_B^*, 1)
M25	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Gamma(θ_B, 1)	Exp(1)
M26	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Gamma(θ_B, 1)	Weibull(θ_B^*, 1)
M27	Gamma(shape=β, scale=$\frac{1}{\beta}$)	Gamma(θ_B, 1)	Gamma(θ_B^*, 1)

Table 3.3: AE, MSE, AL, CP of the parameters based on 500 simulated samples for size $n = 100$, according to load-sharing systems from different models with GFB parameters $\Theta_1 = (\theta_1, \theta_2, \theta_1^*, \theta_2^*, \theta_B, \theta_B^*)$ having Exponential(rate=1) frailty distribution

Model	Parameters	θ_1	θ_2	θ_1^*	θ_2^*	θ_B	θ_B^*
Number	True Value	0.3	0.4	0.5	1.0	2.0	1.5
	AE	0.306	0.409	0.514	1.035	-	-
M1	MSE	0.004	0.006	0.011	0.058	-	-
	AL	0.243	0.299	0.414	0.949	-	-
	CP	95.8	94.0	94.6	95.8	-	-
	AE	0.307	0.411	0.515	1.033	-	1.510
M2	MSE	0.004	0.006	0.018	0.077	-	0.012
	AL	0.253	0.313	0.521	1.096	-	0.431
	CP	95.8	94.6	94.6	95.0	-	95.2
	AE	0.308	0.411	0.535	1.084	-	1.553
M3	MSE	0.004	0.006	0.021	0.109	-	0.148
	AL	0.250	0.309	0.540	1.271	-	1.405
	CP	95.8	94.2	95.0	95.8	-	92.8
	AE	0.305	0.407	0.515	1.037	2.021	-
M4	MSE	0.004	0.006	0.011	0.058	0.022	-
	AL	0.249	0.308	0.419	0.960	0.570	-
	CP	95.6	94.8	95.0	95.6	93.2	-
	AE	0.306	0.409	0.510	1.025	2.029	1.536
M5	MSE	0.004	0.007	0.025	0.098	0.029	0.037
	AL	0.256	0.318	0.601	1.197	0.640	0.724
	CP	96.0	95.4	92.8	93.4	94.0	93.4
	AE	0.306	0.409	0.550	1.128	2.024	1.577
M6	MSE	0.004	0.007	0.029	0.178	0.027	0.285
	AL	0.254	0.315	0.643	1.543	0.596	1.875
	CP	96.0	94.8	94.8	95.8	93.4	92.8
	AE	0.312	0.416	0.512	1.031	2.021	-
M7	MSE	0.007	0.012	0.011	0.058	0.059	-
	AL	0.314	0.401	0.415	0.945	0.963	-
	CP	94.0	93.6	94.6	95.2	94.2	-
	AE	0.320	0.427	0.505	1.013	2.037	1.532
M8	MSE	0.010	0.015	0.030	0.120	0.069	0.023
	AL	0.360	0.463	0.636	1.302	1.029	0.551
	CP	92.8	93.4	90.6	91.4	94.4	93.4
	AE	0.316	0.422	0.537	1.093	2.029	1.564
M9	MSE	0.009	0.014	0.024	0.133	0.065	0.413
	AL	0.339	0.433	0.588	1.358	0.990	2.200
	CP	92.4	93.8	94.4	94.8	94.0	94.2

Table 3.4: AE, MSE, AL, CP of the parameters based on 500 simulated samples for size $n = 100$, according to load-sharing systems from different models with GFB parameters $\Theta_1 = (\theta_1, \theta_2, \theta_1^*, \theta_2^*, \theta_B, \theta_B^*)$ having Weibull(shape=k, scale=$\frac{1}{\Gamma(1+\frac{1}{k})}$) frailty distribution with $\Theta_2 = k$

Model Number	Parameters	θ_1	θ_2	θ_1^*	θ_2^*	θ_B	θ_B^*	k
	True Value	0.3	0.4	0.5	1.0	2.0	1.5	1.5
M10	AE	0.304	0.405	0.498	1.041	-	-	1.563
	MSE	0.004	0.005	0.008	0.047	-	-	0.065
	AL	0.229	0.280	0.371	0.877	-	-	0.981
	CP	92.6	94.2	93.8	95.2	-	-	95.8
M11	AE	0.307	0.409	0.511	1.059	-	1.504	1.598
	MSE	0.005	0.007	0.019	0.080	-	0.026	0.224
	AL	0.274	0.341	0.535	1.087	-	0.701	1.421
	CP	93.4	93.0	93.8	95.8	-	94.2	94.2
M12	AE	0.309	0.411	0.526	1.120	-	1.561	1.548
	MSE	0.004	0.006	0.021	0.152	-	0.182	0.083
	AL	0.227	0.341	0.664	1.646	-	1.945	1.451
	CP	94.4	95.8	95.0	97.4	-	95.0	95.6
M13	AE	0.303	0.405	0.500	1.047	2.007	-	1.568
	MSE	0.004	0.005	0.010	0.053	0.030	-	0.073
	AL	0.236	0.290	0.399	0.946	0.728	-	1.218
	CP	94.4	94.4	93.4	95.0	95.4	-	95.8
M14	AE	0.316	0.416	0.502	1.034	2.064	1.588	1.620
	MSE	0.006	0.008	0.042	0.137	0.071	0.134	0.524
	AL	0.283	0.352	0.754	1.381	1.007	1.328	1.906
	CP	93.4	92.4	87.0	91.2	89.8	86.0	84.2
M15	AE	0.305	0.411	0.562	1.223	2.035	1.603	1.562
	MSE	0.005	0.006	0.049	0.397	0.045	0.428	0.152
	AL	0.260	0.325	0.923	2.345	0.880	2.625	1.609
	CP	92.2	94.4	93.0	94.8	94.0	94.4	93.8
M16	AE	0.312	0.416	0.500	1.046	2.016	-	1.559
	MSE	0.008	0.011	0.009	0.052	0.067	-	0.066
	AL	0.340	0.435	0.384	0.908	1.044	-	1.084
	CP	92.2	95.4	93.0	95.6	95.4	-	96.2
M17	AE	0.334	0.441	0.534	1.096	2.036	1.510	1.632
	MSE	0.024	0.035	0.046	0.205	0.105	0.057	0.242
	AL	0.541	0.696	0.853	1.699	1.413	1.045	1.824
	CP	92.6	92.4	92.8	93.6	93.6	91.0	91.6
M18	AE	0.326	0.434	0.550	1.159	2.038	1.598	1.552
	MSE	0.014	0.020	0.040	0.245	0.083	0.637	0.114
	AL	0.460	0.596	0.828	1.984	1.255	0.637	0.114
	CP	92.8	95.4	93.8	94.4	96.0	93.0	93.6

Table 3.5: AE, MSE, AL, CP of the parameters based on 500 simulated samples for size $n = 100$, according to load-sharing systems from different models with GFB parameters $\Theta_1 = (\theta_1, \theta_2, \theta_1^*, \theta_2^*, \theta_B, \theta_B^*)$ having Gamma(shape=β, scale=$\frac{1}{\beta}$) frailty distribution with $\Theta_2 = \beta$

Model	Parameters	θ_1	θ_2	θ_1^*	θ_2^*	θ_B	θ_B^*	β
Number	True Value	0.3	0.4	0.5	1.0	2.0	1.5	1.5
M19	AE	0.303	0.404	0.508	1.028	-	-	1.632
	MSE	0.004	0.005	0.011	0.052	-	-	0.166
	AL	0.233	0.287	0.392	0.900	-	-	1.369
	CP	93.6	95.8	90.4	95.4	-	-	97.0
M20	AE	0.304	0.406	0.524	1.055	-	1.492	1.661
	MSE	0.005	0.006	0.020	0.075	-	0.021	0.263
	AL	0.273	0.343	0.549	1.116	-	0.650	1.946
	CP	92.2	94.0	94.4	95.4	-	96.0	96.8
M21	AE	0.303	0.406	0.520	1.059	-	1.502	1.646
	MSE	0.004	0.006	0.022	0.111	-	0.159	0.228
	AL	0.251	0.312	0.587	1.387	-	1.668	1.661
	CP	92.4	94.4	92.8	94.6	-	95.8	98.0
M22	AE	0.301	0.403	0.510	1.033	2.021	-	1.635
	MSE	0.004	0.005	0.012	0.053	0.027	-	0.214
	AL	0.237	0.293	0.413	0.940	0.682	-	1.593
	CP	93.2	95.0	92.2	95.2	96.2	-	96.4
M23	AE	0.312	0.416	0.518	1.034	2.048	1.557	1.756
	MSE	0.005	0.008	0.037	0.122	0.065	0.108	0.852
	AL	0.291	0.360	0.772	1.408	1.007	1.300	2.921
	CP	94.6	94.0	90.4	92.2	91.6	89.4	91.2
M24	AE	0.304	0.403	0.564	1.164	2.040	1.558	1.652
	MSE	0.004	0.006	0.062	0.342	0.041	0.419	0.425
	AL	0.262	0.328	0.911	2.179	0.873	2.614	2.452
	CP	94.0	94.2	93.4	95.0	94.8	93.8	95.4
M25	AE	0.315	0.422	0.510	1.032	2.037	-	1.628
	MSE	0.007	0.013	0.011	0.053	0.064	-	0.193
	AL	0.340	0.437	0.403	0.931	1.032	-	1.500
	CP	94.8	95.8	92.6	96.2	95.6	-	97.8
M26	AE	0.329	0.438	0.543	1.097	2.045	1.506	1.698
	MSE	0.018	0.032	0.055	0.189	0.103	0.049	0.651
	AL	0.516	0.661	0.858	1.702	1.362	0.949	2.494
	CP	91.6	91.6	92.8	93.6	93.2	93.0	94.0
M27	AE	0.328	0.439	0.544	1.120	2.052	1.537	1.623
	MSE	0.015	0.030	0.041	0.203	0.087	0.532	0.245
	AL	0.444	0.582	0.781	1.792	1.208	2.846	2.052
	CP	94.6	94.8	93.4	95.4	95.2	96.2	95.6

1. A two-component load-sharing data of size n from a parent model is generated, the parent model being one of the models $M1 - M27$, as described in Table 3.2;

2. To the generated data, all the candidate models $M1 - M27$ are fitted;

3. Using the estimated model parameters, the model selection criteria are calculated for each of the fitted candidate models;

4. By any of the model selection criteria, the model with the lowest value is chosen as the best model for the given data among the 27 candidate models.

As the model selection study is quite elaborate in nature, in this chapter, we demonstrate a part of the study for illustrative purposes. We generate two-component load-sharing samples from two different parent models, $M1$ and $M14$, for sizes n =50, 100, 150, 200. Then, to each generated dataset, we fit all the 27 candidate models $M1$ to $M27$ and the best model is selected. Figures 3.1 and 3.2 present the proportions of the top three candidate models selected by different model selection criteria. From Figure 3.1, it is observed that $M19$ is selected more frequently than the parent model $M1$. This is a plausible case, since $M19$ represents a more general model, with the exponential distribution for the baseline of component lifetimes and the gamma distribution for the frailty, compared to $M1$, which has the exponential distribution for both the component's baselines and the frailty.

From Figure 3.2, observe that the parent model $M14$ is selected the highest number of times, as one would expect. Also, as the sample size increases, the probability of correct selection steeply increases. These observations clearly establish that the standard model selection criteria can quite accurately identify the parent baseline-frailty combinations for the GFB-GG model structure.

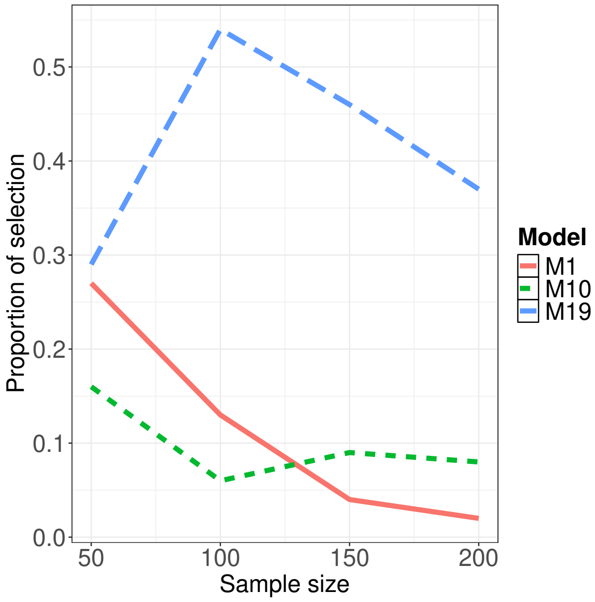
3.5.3 Discussions

The results of this elaborate simulation study clearly indicate that for two-component load-sharing data,

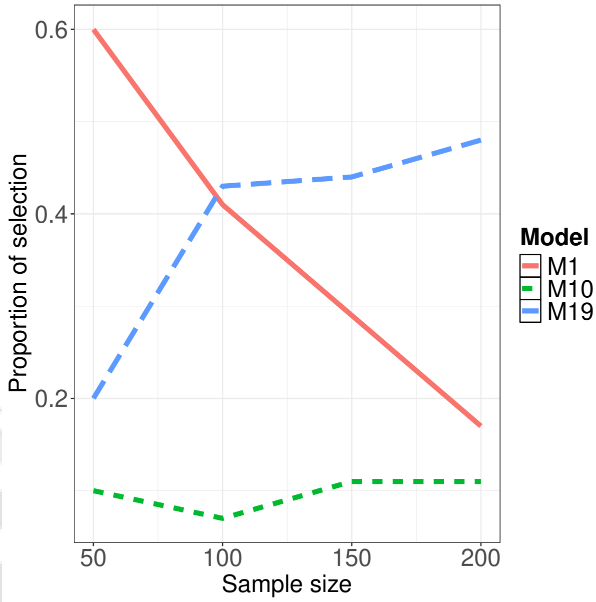
1. The proposed fitting method can satisfactorily estimate the model parameters of the baseline-frailty distribution combinations within the GFB-GG model structure.

2. The standard model selection criteria can accurately identify, even for moderately small sample sizes, the parent distribution with any baseline-frailty combination within the GFB-GG model structure.

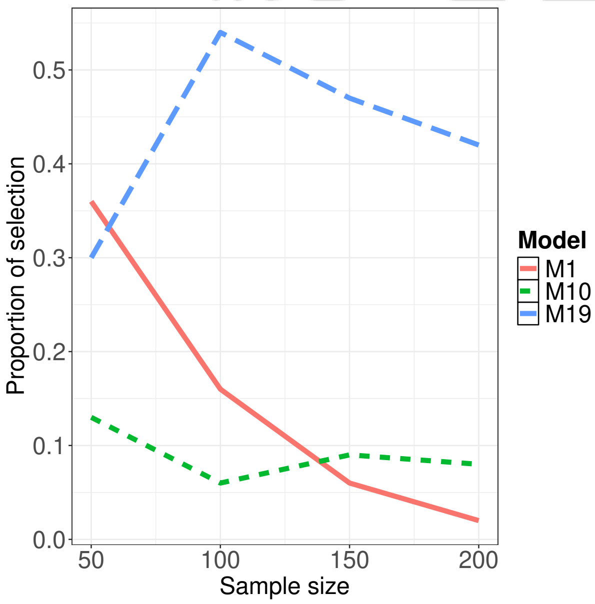
These two findings indicate the principal advantage of using the proposed model. When it is fitted to a given two-component load-sharing data, one can choose the best



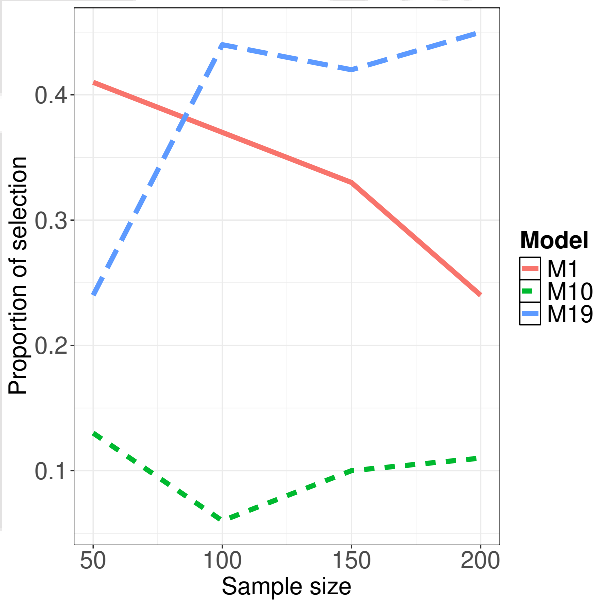
(a) Selection by AIC



(b) Selection by BIC



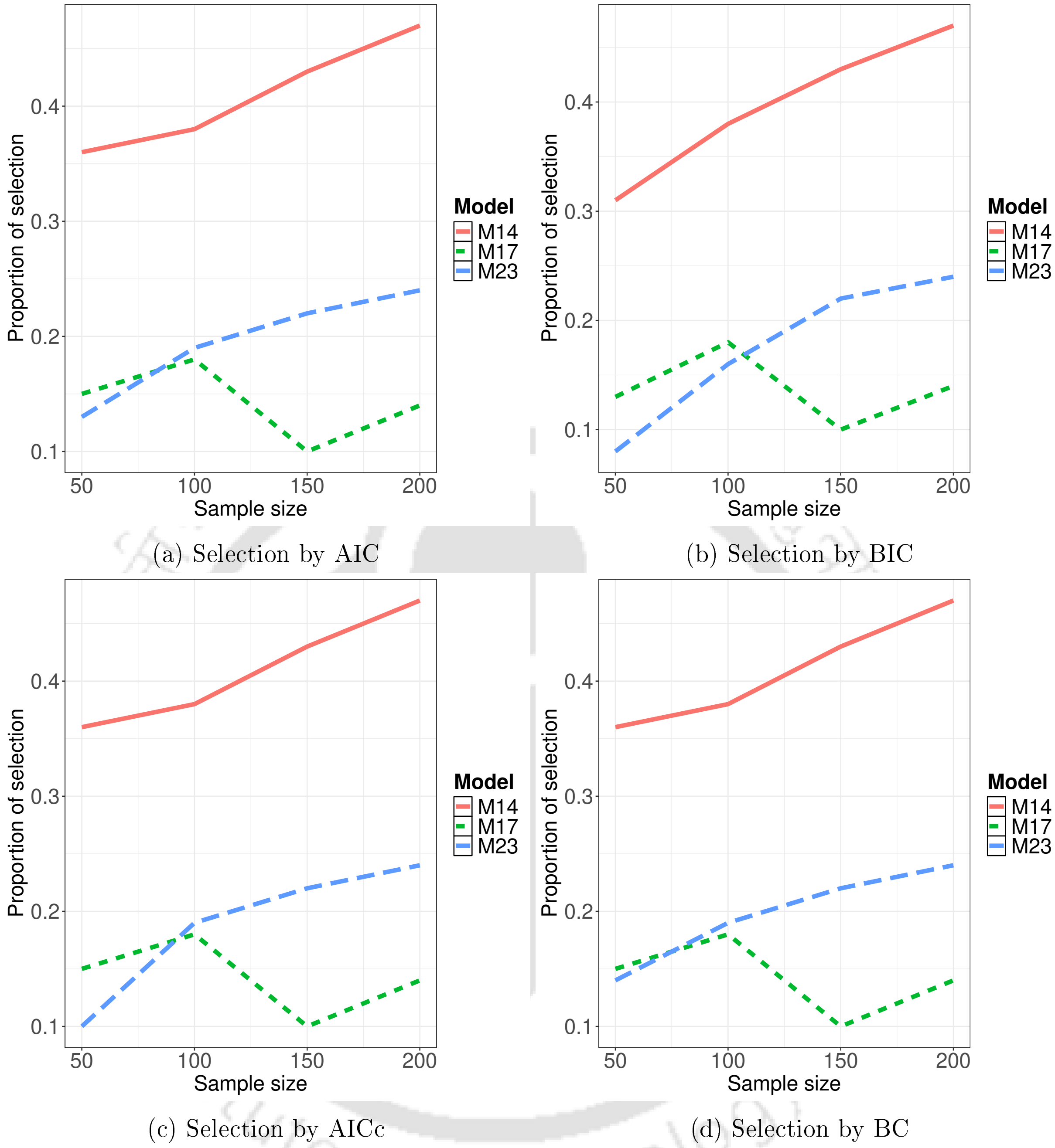
(c) Selection by AICc



(d) Selection by BC

Figure 3.1: Proportion of times the top three models get selected when the parent model is M1

model involving a specific combination of the baseline and frailty distributions within the GFB-GG model structure. Also, since the GFB and GG families are very flexible, the proposed modeling structure actually provides a broad variety of models to choose from, both for the baseline as well as the frailty distribution. Thus, this structure very effectively captures the dependence between the components of the load-sharing system, which in turn results in suitable inferences for various characteristics related to the system lifetime.



(a) Selection by AIC (b) Selection by BIC

(c) Selection by AICc (d) Selection by BC

Figure 3.2: Proportion of times the top three models get selected when the parent model is M14

3.6 Numerical Examples

3.6.1 A Simulation Case

A simulation case demonstrates the effectiveness of using the proposed GFB-GG model structure for two-component load-sharing systems. The simulated two-component load-sharing dataset is provided in Table 3.6.

The dataset is of size 200 and is generated using general choices for the baseline and frailty distributions within the GFB-GG model structure. The baseline distributions of the component lifetimes, before the first failure, are taken to be Weibull. After the first failure occurs, the lifetime of the remaining component is generated from a Gamma distribution. The shared frailty that connects the two components is generated from the GG distribution with general parameter choices, i.e., these parameter values for the GG distribution do not lead to any special member within the GG family of distributions. Specifically, the true parameter values are taken to be $\theta_1 = 0.25$, $\theta_2 = 0.50$, $\theta_1^* = 0.75$, $\theta_2^* = 1$, $\theta_B = 2$, $\theta_B^* = 1.5$, $k = 1.5$, $\beta = 1.5$, $\sigma = 1.5$.

To this simulated data, a total of 45 models with various baseline-frailty distribution combinations are fitted. In addition to the 27 model combinations presented in Table 3.2, another 18 baseline-frailty model combinations (9 each for lognormal and generalized gamma frailty combined with different baselines from the GFB model) are fitted to the simulated data by using the proposed fitting method. Table 3.7 presents the top three models among them, according to their AIC values. The table also provides the standard errors and lower and upper limits of 95% bootstrap CIs.

We observe that the model with Weibull-Gamma baselines within the GFB distribution, and the GG frailty is selected as the best model for this data, as this model has the lowest AIC value among the 45 candidate models. Recall that this combination was the true parent model, too. Note that, instead of fitting the general GFB-GG model structure with 45 candidate models to this two-component load-sharing data, if any special combination of baseline-frailty (such as the model of Asha et al. [4]) was fitted, it would have resulted in a sub-optimal model fit. Only when the GFB-GG model structure is fitted to this data, and a model selection is performed within the families, one can identify the best model for this data. This clearly demonstrates the utility of the proposed GFB-GG model structure for two-component load-sharing data in obtaining the optimal model fit.

3.6.2 Real Data: The Nuclear Reactor Data

In this section, we present an analysis of the dataset on nuclear reactor data provided in Subsection 1.2.3, by using the proposed GFB-GG model. All the 45 models with various baseline-frailty distribution combinations are fitted to the data, and model selection within the family has been performed in a similar fashion as explained in Section 3.4. The failure times of the nuclear plants are scaled before analysis by dividing them by 365. This will not affect the inference in any way. The top three models for the nuclear reactor data, in order of lowest AIC values, are presented in Table 3.8, along with the point estimates, standard errors of estimates, and bootstrap CIs of their parameters.

The best model is M12 according to AIC with an AIC value of −175.469, which is

Table 3.6: Simulated two-component load-sharing data: Baseline component lifetimes as Weibull($\theta_B = 2$) before the first failure, and Gamma($\theta_B^* = 1.5$) after the first failure, with PFR parameters ($\theta_1 = 0.25,\ \theta_2 = 0.5,\ \theta_1^* = 0.75,\ \theta_2^* = 1$); frailty from GG distribution ($k = 1.5,\ \beta = 1.5$)

n	Y_{1i}	Y_{2i}	n	Y_{1i}	Y_{2i}	n	Y_{1i}	Y_{2i}	n	Y_{1i}	Y_{2i}	n	Y_{1i}	Y_{2i}
1	0.55	0.44	2	5.77	1.30	3	1.86	0.77	4	3.66	0.31	5	4.51	1.64
6	1.15	4.30	7	0.79	2.96	8	1.89	1.87	9	0.83	1.37	10	1.36	1.27
11	3.90	1.35	12	2.52	1.37	13	1.83	4.37	14	1.95	0.88	15	1.10	1.89
16	1.07	1.19	17	0.67	0.22	18	2.69	3.04	19	1.74	1.21	20	0.61	3.57
21	1.22	1.14	22	1.77	1.28	23	1.00	1.19	24	3.86	1.02	25	1.14	0.86
26	1.30	3.42	27	1.70	1.04	28	3.05	0.57	29	4.70	1.04	30	3.79	0.62
31	1.06	0.57	32	3.30	4.94	33	0.44	1.18	34	0.70	2.70	35	4.96	2.62
36	1.73	1.68	37	11.32	0.94	38	2.68	0.53	39	1.67	1.06	40	1.18	5.93
41	0.89	1.43	42	1.67	0.65	43	5.70	2.70	44	10.17	2.17	45	3.61	1.04
46	1.29	7.00	47	0.22	2.08	48	2.11	6.11	49	10.95	0.95	50	1.11	1.77
51	1.11	0.95	52	0.58	1.30	53	1.77	2.88	54	2.13	1.47	55	5.31	1.60
56	2.74	1.47	57	1.77	1.48	58	2.57	1.29	59	1.55	1.28	60	0.37	1.34
61	9.32	1.83	62	0.19	1.25	63	9.55	4.45	64	1.09	0.63	65	1.86	1.45
66	1.16	0.61	67	1.43	0.95	68	4.38	0.24	69	0.56	1.80	70	0.79	0.37
71	8.57	4.46	72	1.39	5.69	73	2.23	0.52	74	2.88	0.83	75	0.98	1.56
76	1.15	3.02	77	36.40	4.03	78	2.71	3.36	79	0.82	1.24	80	12.17	0.66
81	0.72	0.79	82	0.93	1.01	83	2.51	0.54	84	0.66	0.69	85	2.19	0.73
86	1.73	8.69	87	3.25	0.19	88	2.62	1.13	89	1.79	1.42	90	1.29	3.70
91	0.96	1.06	92	0.70	0.31	93	1.13	1.21	94	0.95	3.19	95	8.02	0.54
96	1.11	2.12	97	2.23	2.23	98	1.50	0.56	99	1.39	0.22	100	2.66	1.30
101	0.52	1.45	102	4.96	1.58	103	0.94	0.72	104	0.34	1.84	105	7.91	1.77
106	2.61	0.33	107	1.82	0.90	108	3.35	1.31	109	5.00	0.28	110	1.85	8.25
111	1.73	4.73	112	2.39	1.12	113	3.18	0.88	114	0.93	1.57	115	1.45	3.94
116	0.65	0.24	117	0.99	1.55	118	4.16	1.29	119	1.15	0.87	120	0.50	0.53
121	0.45	0.76	122	0.83	2.19	123	1.63	0.62	124	0.54	0.22	125	1.80	0.65
126	6.89	0.57	127	5.72	1.50	128	4.04	1.38	129	0.94	4.43	130	3.08	0.88
131	3.12	2.63	132	0.95	0.85	133	1.44	1.69	134	0.52	5.25	135	0.55	1.04
136	0.90	2.12	137	2.47	1.53	138	1.24	2.07	139	2.43	1.49	140	1.44	0.53
141	0.63	1.06	142	0.92	2.88	143	1.35	0.87	144	1.10	0.68	145	0.54	3.58
146	2.39	0.21	147	1.76	0.08	148	1.94	1.24	149	3.51	0.99	150	2.13	0.75
151	2.19	1.40	152	0.91	1.64	153	1.62	1.78	154	0.96	0.86	155	5.73	0.98
156	3.82	1.72	157	4.51	1.17	158	2.12	3.65	159	1.43	2.61	160	1.76	1.44
161	3.48	1.69	162	0.91	0.73	163	0.64	6.09	164	0.04	0.60	165	0.77	1.95
166	8.23	2.24	167	1.31	2.63	168	7.52	1.04	169	4.62	0.65	170	0.37	1.09
171	1.19	0.77	172	0.10	1.01	173	2.30	2.39	174	1.25	0.59	175	1.08	1.17
176	6.64	1.47	177	3.35	0.52	178	1.09	1.04	179	1.53	2.04	180	5.59	0.98
181	0.17	0.65	182	1.66	0.62	183	2.52	0.95	184	1.12	1.27	185	3.76	0.75
186	1.39	0.68	187	1.75	0.47	188	1.00	1.81	189	3.18	3.06	190	1.94	0.89
191	2.87	1.08	192	2.10	0.98	193	1.20	0.88	194	1.83	1.36	195	1.74	1.43
196	1.06	0.66	197	5.53	2.09	198	1.83	0.72	199	3.84	0.94	200	1.10	0.48

the lowest among the 45 models considered. It may also be mentioned here that for the best model for this data, reported in Franco et al. [22]'s work has an AIC value -172.342. Recall that Franco et al. [22] also considered the GFB distribution for the baseline component lifetimes, but they did not consider a frailty that induced dependence due to unobserved random factors. Comparing the AIC values, we can conclude that the best model for the nuclear reactor data within the GFB-GG model structure provides a better fit compared to the best model of Franco et al. [22]. This once again shows the usefulness of the proposed model.

We also develop estimates of some reliability characteristics such as the MTTF, RMT, and MRT of load-sharing systems under the GFB-GG model. The MTTF of a load-sharing system is the expected time the system operates till its failure. Let T denote the system failure time; then, $T = \max(Y_1, Y_2)$. The MTTF of a load-sharing system is $E(T)$.

Table 3.7: Estimates of the parameters of the top three models based on simulated data given in Table 3.6

Model	Parameters	θ_1	θ_2	θ_1^*	θ_2^*	θ_B	θ_B^*	k	β	σ	AIC
	MLE	0.317	0.539	0.871	1.251	2.147	1.820	1.415	1.627	-	1278.079
R_B: Weibull	Boot SE	0.015	0.022	0.084	0.150	0.051	0.156	0.342	0.672	-	
R_B^*: Gamma	Boot LL	0.296	0.503	0.735	0.935	2.064	1.585	0.990	0.412	-	
Frailty: GG	Boot UL	0.354	0.588	1.064	1.524	2.264	2.196	2.332	3.047	-	
	MLE	0.350	0.596	1.132	1.674	2.205	2.127	-	-	0.746	1278.326
R_B: Weibull	Boot SE	0.017	0.031	0.111	0.186	0.056	0.130	-	-	0.047	
R_B^*: Gamma	Boot LL	0.314	0.531	0.921	1.291	2.102	1.845	-	-	0.649	
Frailty: Log-normal	Boot UL	0.382	0.653	1.357	2.020	2.321	2.355	-	-	0.834	
	MLE	0.314	0.535	0.399	0.572	2.153	1.258	1.862	1.087	-	1278.590
R_B: Weibull	Boot SE	0.017	0.020	0.028	0.036	0.051	0.052	0.336	0.296	-	
R_B^*: Weibull	Boot LL	0.292	0.492	0.342	0.502	2.095	1.158	1.378	0.399	-	
Frailty: GG	Boot UL	0.357	0.569	0.450	0.641	2.296	1.363	2.697	1.561	-	

Table 3.8: Estimates of the parameters of the top three models based on the nuclear reactor data given in Table 1.3

Model	Parameters	θ_1	θ_2	θ_1^*	θ_2^*	θ_B	θ_B^*	k	β	AIC
	MLE	3.751	108.749	0.663	12.674	-	0.379	110.001	-	-175.469
M12	Boot SE	0.590	4.978	0.061	0.109	-	0.057	0.003	-	
	Boot LL	2.573	99.488	0.516	12.482	-	0.242	109.990	-	
	Boot UL	4.887	119.002	0.754	12.910	-	0.465	110.002	-	
	MLE	3.765	109.212	0.666	12.700	-	0.382	-	273.502	-175.464
M21	Boot SE	0.548	5.495	0.056	0.135	-	0.063	-	0.017	
	Boot LL	2.635	98.154	0.525	12.462	-	0.264	-	273.452	
	Boot UL	4.785	119.696	0.743	12.989	-	0.510	-	273.518	
	MLE	8.000	226.000	0.761	15.000	1.148	0.440	-	12.000	-174.270
M24	Boot SE	1.553	1.587	0.079	0.054	0.019	0.065	-	4.838	
	Boot LL	4.591	222.549	0.619	14.900	1.120	0.326	-	5.866	
	Boot UL	10.679	228.769	0.928	15.111	1.195	0.581	-	24.829	

The RMT of a system is the probability that the system will operate till a desired time t_0. Thus, RMT at t_0 is $S(t_0) = P(T > t_0)$. It is clear that it is difficult to find the MTTF and RMT analytically under this model. The MRT of a system is the expected additional time the system will survive if it has already survived a given time t_0. That is,

$$\text{MRT}(t_0) = E(T - t_0 | T > t_0) = \int_{t_0}^{\infty} s f_{T|T>t_0}(s) ds - t_0.$$

However, for this model, obtaining closed-form expressions of MTTF, RMT, and MRT are difficult. Therefore, we propose to estimate these quantities using Monte Carlo simulations. For a Monte Carlo estimates, one needs to generate R data points t_i, $i = 1, 2, \ldots, R$

as realisations of the system lifetime T. Then, MTTF can be estimated by the mean of the generated t_i's. To estimate the RMT at a pre-specified time t_0, compute $\frac{R(t_0)}{R}$, where $R(t_0)$ is the number of realisations of the system lifetime that exceeds t_0.

For Monte Carlo estimation of MRT at t_0, we generate R data points t_i^*, $i = 1, 2, \ldots, R$ as realisations of the truncated lifetime $T|T > t_0$. Then, the estimated MRT at t_0 is

$$\widehat{\text{MRT}}(t_0) = \frac{\sum_{i=1}^{R} t_i^*}{R} - t_0.$$

For a reasonable estimates of these quantities, a large value of R should be used.

Here, we calculate these reliability characteristics for the top three models, viz., M12, M21, and M24, for the scaled nuclear reactor data based on 10000 Monte Carlo replications. The MTTFs under the models M12, M21, and M24 are estimated to be 0.740 years, 0.735 years, and 0.811 years, respectively. Monte Carlo estimates of the MRT and RMT are calculated at the 25th, 50th, and 75th sample percentile points of the system failure times and are presented in Table 3.9.

Table 3.9: Mean residual time and reliability in mission time for the nuclear reactor data

Model	t_0	0.111	0.574	1.066
M12	RMT(t_0)	0.698	0.375	0.227
	MRT(t_0)	0.949	1.142	1.228
M24	RMT(t_0)	0.704	0.373	0.225
	MRT(t_0)	0.903	1.084	1.163
M24	RMT(t_0)	0.723	0.392	0.239
	MRT(t_0)	0.996	1.214	1.360

3.7 Conclusions

In this chapter, we present a very general class of parametric models for two-component load-sharing systems. The proposed model with the GFB and GG family of distributions for the baseline component lifetimes and frailty distributions, respectively, can appropriately model the change in the lifetime distribution of the surviving component after the first failure occurs in the system, and can also accurately capture the dependence between the component lifetimes. A model fitting method based on an EM-type algorithm is described in detail, and extensive simulations are carried out. The fitting method is observed to perform satisfactorily. It is shown that the GFB-GG model structure can provide a better fit compared to existing models as one can choose from a wider array of candidate models in this case. This demonstrates the practical utility of the proposed model. Estimation of some important reliability characteristics, such as MTTF, RMT, and MRT, are discussed under the GFB-GG model.

CHAPTER 4

RELIABILITY ANALYSIS OF LOAD-SHARING SYSTEMS WITH PLA*

4.1 Introduction

Our aim in this chapter is to develop a flexible model for analysing load-sharing data, without using restrictive assumptions on the distribution of component lifetimes. The aim is also to suitably estimate important quantities such as RMT, quantile function, MTTF, and MRT of the underlying lifetime distribution of load-sharing systems. These quantities are of practical importance for making various strategies and plans.

We develop a model involving PLAs of the CHF of the distribution of system lifetimes between two successive component failures, capturing the unknown load-share rule at successive stages of component failures. At each of the successive stages of component failures, as the lifetime distributions of surviving operating components change, a new PLA for CHF is used. The model can be suitably tuned by choosing the number of linear pieces for PLA at each stage of failure. Usually, in parametric analyses of load-sharing systems, a specific lifetime distribution is used to model component lifetimes for the entire range. In contrast, the proposed PLA-based modelling approach does not require any such strong parametric assumptions for component lifetime distributions. The PLA-based model can be interpreted as approximating the underlying distribution of system lifetime between two consecutive component failures by several exponential distributions over short ranges defined by cut-points. In survival analysis, PLA-based models are used extensively. Balakrishnan et al. [7] proposed a PLA-based model for the hazard rate of a population with a cured proportion; see also the references therein. Recently, a

*This chapter is based on our published work: S. Biswas, A. Ganguly and D. Mitra. Reliability Analysis of Load-sharing Systems using a Flexible Model with Piecewise Linear Functions. *Applied Stochastic Models in Business and Industry*. https://doi.org/10.1002/asmb.2934, 2025.

PLA-based model for the cumulative hazard function was used to analyse left-truncated right-censored data by Ganguly et al. [24].

The proposed PLA-based model provides a good fit for load-sharing data. In particular, we show that the PLA-based model surpasses several existing load-sharing models with respect to the quality of model fit for the two-motor load-sharing data. We also discuss estimation of RMT, quantile function, MTTF, and MRT under the proposed PLA-based model. Since the accuracy of estimates of RMT, quantile function, MTTF, and MRT depends on the suitability of the model fitted, quite naturally, these important practical quantities are also better estimated under this model.

The main contributions of this chapter are as follows:

- We develop a flexible, data-driven PLA-based model for load-sharing data. The model does not require restrictive parametric assumptions on the underlying component lifetimes.
- We develop inference for the proposed PLA-based model based on multi-component load-sharing data.
- Under the proposed PLA-based model, we develop estimates for important reliability characteristics such as RMT, quantile function, MTTF, and MRT of load-sharing systems.

The rest of this chapter is structured as follows. In Section 4.2, the proposed PLA-based model is presented. Section 4.3 presents likelihood inference for the model based on data from multi-component load-sharing systems. This section also includes the construction of CIs for relevant model parameters and general guidance for practitioners on the selection of cut-points for defining the PLA-based model. Estimation of RMT, quantile function, MTTF, and MRT of load-sharing systems is discussed in Section 4.4. In Section 4.5, results of a detailed Monte Carlo simulation experiment investigating the efficacy and robustness of the PLA-based model are presented. Illustrative numerical examples demonstrating the application of the PLA-based model on two load-sharing datasets and estimation of various important reliability characteristics are presented in Section 4.6. In Section 4.7, we discuss an extension of the PLA-based model that can incorporate covariate information into the modelling. Finally, the chapter is concluded with some remarks in Section 4.8.

4.2 Model Description

Consider a J-component load-sharing system where the components are arranged in parallel. Assume a constant total load on the system and that the failed components of the system are not replaced or repaired. At the beginning when all components are operating, let $U_1^{(0)}, U_2^{(0)}, \ldots, U_J^{(0)}$ denote the latent lifetimes of the components, and $Y^{(0)}$ denote the system lifetime till the first component failure. Obviously,

$$Y^{(0)} = \min\left\{U_1^{(0)}, U_2^{(0)}, \ldots, U_J^{(0)}\right\}.$$

Similarly, for $j = 1, 2, \ldots, J-1$, let $Y^{(j)}$ denote the system lifetime between j-th and $(j+1)$-st failed components. Then,

$$Y^{(j)} = \min\left\{U_1^{(j)}, U_2^{(j)}, \ldots, U_{J-j}^{(j)}\right\},$$

where $U_1^{(j)}, U_2^{(j)}, \ldots U_{J-j}^{(j)}$ denote the latent lifetimes of the operating components after failure of j components, $j = 1, 2, \ldots, J-1$. For all values of j, $U_1^{(j)}, \ldots, U_{J-j}^{(j)}$ are assumed to be independent and identically distributed random variables. It is further assumed that $\left\{U_\ell^{(j)}, \ell = 1, 2, \ldots, J-j;\, j = 0, 1, \ldots, J-1\right\}$ are independent random variables.

Let $h^{(j)}(\cdot)$ and $H^{(j)}(\cdot)$ denote the HF and CHF, respectively, of the distribution of $U_\ell^{(j)}$, $\ell = 1, 2, \ldots, J-j$; $j = 0, 1, 2, \ldots, J-1$. Here, we assume that the HF $h^{(j)}(\cdot)$ is a non-decreasing function for all j. For $y > 0$, the SF of $Y^{(j)}$ is given by

$$P\left(Y^{(j)} > y\right) = P\left[\min\left\{U_1^{(j)}, U_2^{(j)}, \ldots, U_{J-j}^{(j)}\right\} > y\right] = e^{-(J-j)H^{(j)}(y)}.$$

Hence, for $y > 0$, the CDF and PDF of $Y^{(j)}$ are given by

$$F^{(j)}(y) = 1 - e^{-(J-j)H^{(j)}(y)}$$

and

$$f^{(j)}(y) = (J-j)h^{(j)}(y)\, e^{-(J-j)H^{(j)}(y)},$$

respectively.

Now, suppose there are n J-component load-sharing systems, and let $Y_i^{(j)}$ denote the system lifetime between j-th and $(j+1)$-st failed components for the i-th system, $i = 1, 2, \ldots, n$, $j = 0, 1, \ldots, J-1$. Suppose the observed values of $Y_1^{(j)}, Y_2^{(j)}, \ldots, Y_n^{(j)}$ are $y_1^{(j)}, y_2^{(j)}, \ldots, y_n^{(j)}$, respectively. Let, for $j = 0, 1, \ldots, J-1$, $\xi^{(j)} = \left\{\tau_0^{(j)}, \tau_1^{(j)}, \ldots, \tau_N^{(j)}\right\}$ denote a set of $N+1$ cut-points over the time scale $y_1^{(j)}, \ldots, y_n^{(j)}$, with the restrictions that

$$\tau_0^{(j)} < \tau_1^{(j)} < \tau_2^{(j)} < \ldots < \tau_N^{(j)},\ \tau_0^{(j)} \le \min\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\},\ \tau_N^{(j)} \ge \max\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}.$$

Initially, $\xi^{(j)}$ is taken to be fixed and known. In a later section, we discuss how to choose $\xi^{(j)}$.

The proposed model approximates the CHF $H^{(j)}(\cdot)$ by a piecewise linear function defined over intervals $[\tau_{k-1}^{(j)}, \tau_k^{(j)})$, $k = 1, 2, \ldots, N$, constructed by the consecutive cut points in $\xi^{(j)}$. Therefore, over the range $[\tau_0^{(0)}, \tau_N^{(0)})$, the CHF $H^{(0)}(\cdot)$ is approximated by $\Lambda^{(0)}(\cdot)$, where

$$\Lambda^{(0)}(t) = \sum_{k=1}^{N} (a_k + b_k t)\, \mathbf{1}_{[\tau_{k-1}^{(0)}, \tau_k^{(0)})}(t),$$

with a_k's and b_k's as real constants and

$$\mathbf{1}_A(t) = \begin{cases} 1 & \text{if } t \in A \\ 0 & \text{if } t \notin A. \end{cases}$$

One of the possible ways to extend the PLA beyond $\tau_N^{(0)}$ would be to extend the last line segment $a_N + b_N t$ over $[\tau_N^{(0)}, \infty)$. Therefore, the CHF corresponding to PLA over the range $[\tau_0^{(0)}, \infty)$ is

$$\Lambda^{(0)}(t) = \sum_{k=1}^{N} (a_k + b_k t)\, \mathbf{1}_{[\tau_{k-1}^{(0)}, \tau_k^{(0)})}(t) + (a_N + b_N t)\mathbf{1}_{[\tau_N^{(0)}, \infty)}(t),$$

with $\Lambda^{(0)}(\tau_0^{(0)}) = 0$. We also assume that $\Lambda^{(0)}(\cdot)$ is a continuous function. As $\Lambda^{(0)}(\tau_0^{(0)}) = 0$, using the assumption of continuity, a_i's can be expressed in terms of b_i's as follows:

$$a_1 = -b_1 \tau_0^{(0)} \qquad \text{and} \qquad a_k = \sum_{\ell=1}^{k-1} b_\ell \left(\tau_\ell^{(0)} - \tau_{\ell-1}^{(0)}\right) - b_k \tau_{k-1}^{(0)},$$

for $k = 2, 3, \ldots, N$.

Note that the above model can be equivalently described in terms of HFs. In this approach, $h^{(0)}(\cdot)$ over the range $[\tau_0^{(0)}, \tau_N^{(0)})$ is approximated by a piecewise constant function $\lambda^{(0)}(\cdot)$, where

$$\lambda^{(0)}(t) = \sum_{i=1}^{N} b_k \mathbf{1}_{[\tau_{k-1}^{(0)}, \tau_k^{(0)})}(t)\,. \tag{4.1}$$

When components fail within the system, the direct impact of the increased load will be an increased HF for the operating components. To incorporate this information, after the failure of j components, we approximate $h^{(j)}(\cdot)$ over $[\tau_0^{(j)}, \tau_N^{(j)})$, $j = 1, 2, \ldots, J-1$, using the piecewise constant function $\lambda^{(j)}(\cdot)$ as

$$\lambda^{(j)}(t) = \gamma_j \sum_{k=1}^{N} b_k \mathbf{1}_{[\tau_{k-1}^{(j)}, \tau_k^{(j)})}(t), \tag{4.2}$$

with γ_j's are real positive constants. The increase of HF at successive component failures is ensured by the conditions $1 < \gamma_1 < \gamma_2 < \ldots < \gamma_{J-1}$ and $0 < b_1 < b_2 < \ldots < b_N$. The PLAs to the CHFs, corresponding to the PLAs of the HFs given in Eq. (4.2) are given by

$$\Lambda^{(j)}(t) = \gamma_j \sum_{k=1}^{N} \left[\sum_{\ell=1}^{k-1} b_\ell \left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)} \right) + b_k \left(t - \tau_{k-1}^{(j)} \right) \right] \mathbf{1}_{[\tau_{k-1}^{(j)}, \tau_k^{(j)})}(t). \tag{4.3}$$

We treat $b_1, b_2, \ldots, b_N$ and $\gamma_1, \gamma_2, \ldots, \gamma_{J-1}$ as unknown parameters and estimate them from component failure data. It may be mentioned here that the PLA model can be interpreted as an approximation of the underlying lifetime distribution by several exponential models (with different rate parameters) over the ranges specified by the cut-points. Figure 4.1 depicts the approximation of the original CHF $H^{(j)}(\cdot)$ using $\Lambda^{(j)}(\cdot)$ consisting

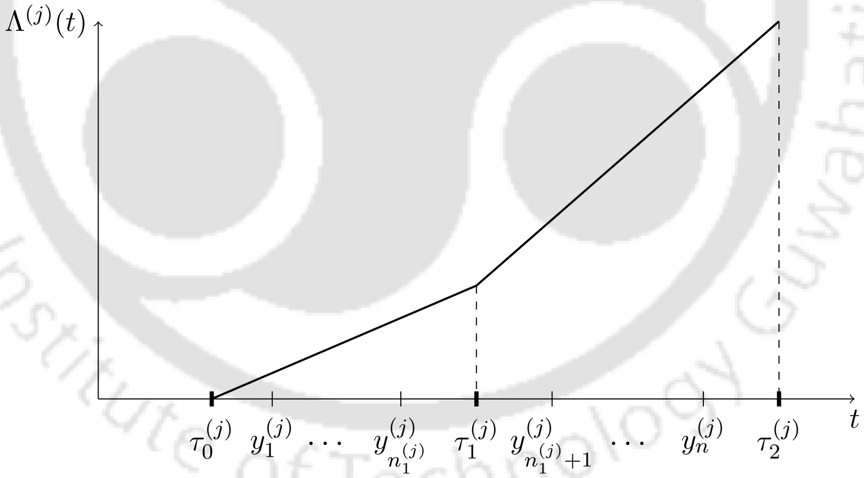


Figure 4.1: Pictorial representation of approximation of CHF for $N = 2$

two linear pieces. Thus, three cut-points $\tau_0^{(j)}$, $\tau_1^{(j)}$, and $\tau_2^{(j)}$ are placed such that all the observations on system lifetime between j-th and $(j+1)$-st failed components belong in the interval $[\tau_0^{(j)}, \tau_2^{(j)}]$. Then the original CHF is approximated using a piecewise linear function as shown in Figure 4.1, where $n_1^{(j)}$ denotes the number of observations between $\tau_0^{(j)}$ and $\tau_1^{(j)}$.

4.3 Likelihood Inference

The parameters involved in the PLA-based model are estimated from the data obtained from a set of load-sharing systems. The available data from n J-component load-sharing systems are of the form

$$\text{Data} = \left\{ y_i^{(j)} \,:\, i = 1,\, 2,\, \ldots,\, n;\; j = 0,\, 1,\, \ldots,\, J-1 \right\},$$

where $y_i^{(j)}$ is the observed system lifetime between j-th and $(j+1)$-st failed components for the i-th system. For $j = 0, 1, 2, \ldots, J-1$, and $k = 1, 2, \ldots, N$, define

$$I_k^{(j)} = \left\{ i : y_i^{(j)} \in \left[\tau_{k-1}^{(j)},\, \tau_k^{(j)} \right) \right\} \qquad \text{and} \qquad n_k^{(j)} = |I_k^{(j)}|.$$

Obviously, $\sum_{k=1}^{N} n_k^{(j)} = n$. The likelihood function for the PLA-based model is then given by

$$L\left(\Theta|\text{Data}\right) = \prod_{i=1}^{n} \prod_{j=0}^{J-1} \left[(J-j)\gamma_j \sum_{k=1}^{N} b_k \mathbf{1}_{[\tau_{k-1}^{(j)},\, \tau_k^{(j)})}\left(y_i^{(j)} \right) e^{-(J-j)\gamma_j \left[\sum_{\ell=1}^{k-1} b_\ell \left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)} \right) + b_k \left(y_i^{(j)} - \tau_{k-1}^{(j)} \right) \right]} \right],$$

where $\gamma_0 = 1$, and $\Theta = (\gamma_1,\, \gamma_2,\, \ldots,\, \gamma_{J-1}, b_1,\, b_2,\, \ldots,\, b_N)'$ is the vector of parameters. The corresponding log-likelihood function, ignoring the additive constant, can be expressed as

$$l\left(\Theta|\text{Data}\right) = \sum_{k=1}^{N} \left[\left(\sum_{j=0}^{J-1} n_k^{(j)} \right) \ln b_k - \left(\sum_{j=0}^{J-1} (J-j)\gamma_j T_k^{(j)} \right) b_k \right] + n \sum_{j=0}^{J-1} \ln \gamma_j, \tag{4.4}$$

where

$$T_k^{(j)} = \sum_{i \in I_k^{(j)}} \left(y_i^{(j)} - \tau_{k-1}^{(j)} \right) + \left(n - \sum_{\ell=1}^{k} n_\ell^{(j)} \right) \left(\tau_k^{(j)} - \tau_{k-1}^{(j)} \right),$$

for $k = 1, 2, \ldots, N$; $j = 0, 1, \ldots, J-1$. Equating the partial derivative of the log-likelihood function in Eq. (4.4) with respect to b_k to zero, we can express b_k in terms of the load-share parameters $\boldsymbol{\gamma} = (\gamma_1,\, \gamma_2,\, \ldots,\, \gamma_{J-1})$ as

$$b_k = b_k\left(\boldsymbol{\gamma}\right) = \frac{\sum_{j=0}^{J-1} n_k^{(j)}}{\sum_{j=0}^{J-1} (J-j)\gamma_j T_k^{(j)}}, \quad k = 1, \ldots, N. \tag{4.5}$$

Substituting $b_k(\boldsymbol{\gamma})$ from Eq. (4.5) in Eq. (4.4), the profile log-likelihood in $\boldsymbol{\gamma}$, ignoring the additive constant, is obtained as

$$\tilde{l}\,(\boldsymbol{\gamma}) = \sum_{k=1}^{N}\left[\left(\sum_{j=0}^{J-1} n_k^{(j)}\right)\left\{\ln\left(\sum_{j=0}^{J-1} n_k^{(j)}\right) - \ln\left(\sum_{j=0}^{J-1}(J-j)\gamma_j T_k^{(j)}\right)\right\}\right] + n\sum_{j=0}^{J-1}\ln\gamma_j.$$

For optimizing the profile log-likelihood $\tilde{l}\,(\boldsymbol{\gamma})$ with respect to $\boldsymbol{\gamma}$, any routine maximizer of a standard statistical software may be used. Once the MLEs $\widehat{\gamma}_1, \widehat{\gamma}_2, \ldots, \widehat{\gamma}_{J-1}$ of $\gamma_1, \gamma_2, \ldots, \gamma_{J-1}$ are obtained by numerical optimization of $\tilde{l}\,(\boldsymbol{\gamma})$, they can be plugged into Eq. (4.5) to get MLEs of b_k as

$$\widehat{b}_k = b_k\left(\widehat{\gamma}_1, \ldots, \widehat{\gamma}_{J-1}\right), \quad k = 1, 2, \ldots, N.$$

4.3.1 A Special Case: Two-component Load-sharing Systems

For analysing data from two-component load-sharing systems, if two linear pieces are used in the PLA-based model, MLEs can be derived analytically and explicitly. Consider the case when $J = 2$ and $N = 2$. In this case, the log-likelihood function simplifies to

$$l\,(\Theta|\text{Data}) = \sum_{k=1}^{2}\left[\left(\sum_{j=0}^{1} n_k^{(j)}\right)\ln b_k - \left(\sum_{j=0}^{1}(2-j)\gamma_j T_k^{(j)}\right) b_k\right] + n\sum_{j=0}^{1}\ln\gamma_j,$$

with

$$T_k^{(j)} = \sum_{i\in I_k^{(j)}}\left(y_i^{(j)} - \tau_{k-1}^{(j)}\right) + \left(n - \sum_{\ell=1}^{k} n_\ell^{(j)}\right)\left(\tau_k^{(j)} - \tau_{k-1}^{(j)}\right),$$

for $k = 1, 2$, $j = 0, 1$ and $\gamma_0 = 1$. Here, $\Theta = (\gamma_1, b_1, b_2)$.

Equating $\frac{\partial l(\Theta|\text{Data})}{\partial b_1}$ and $\frac{\partial l(\Theta|\text{Data})}{\partial b_2}$ to zero, we get

$$b_1 = \frac{n_1^{(0)} + n_1^{(1)}}{2T_1^{(0)} + \gamma_1 T_1^{(1)}} \tag{4.6}$$

$$b_2 = \frac{n_2^{(0)} + n_2^{(1)}}{2T_2^{(0)} + \gamma_1 T_2^{(1)}}. \tag{4.7}$$

Equating $\frac{\partial l(\Theta|\text{Data})}{\partial \gamma_1}$ to zero gives

$$\gamma_1 = \frac{n}{T_1^{(1)} b_1 + T_2^{(1)} b_2},$$

in which, substituting b_1 and b_2 from Eqs. (4.6) and (4.7), a quadratic equation in γ_1 is obtained as follows

$$Q(\gamma_1) = n\gamma_1^2 B_{0,12} + 2\gamma_1 \left\{ \left(n_1^{(0)} + n_1^{(1)} - n \right) B_{2,1} + \left(n_2^{(0)} + n_2^{(1)} - n \right) B_{1,2} \right\} - 4nB_{12,0} = 0,$$

with $B_{0,12} = T_1^{(1)}T_2^{(1)}$, $B_{1,2} = T_1^{(0)}T_2^{(1)}$, $B_{2,1} = T_2^{(0)}T_1^{(1)}$ and $B_{12,0} = T_1^{(0)}T_2^{(0)}$. Solving $Q(\gamma_1) = 0$, we have two values of γ_1 from which we choose the suitable one, and then from Eqs. (4.6) and (4.7) we get the MLEs of b_1 and b_2, respectively.

4.3.2 Confidence Intervals

As discussed above, the MLEs for the parameters of the PLA-based model are not available in explicit form in general, except for the special case of two-component load-sharing systems considered in Section 4.3.1. As a result, exact CIs for the model parameters cannot be obtained. Here, we develop bootstrap CIs for the parameters, as described below. Using the MLE $\widehat{\Theta}$, B bootstrap samples can be obtained in the same sampling framework; let $\widehat{\Theta}_s^* = \left(\widehat{\gamma}_{1s}^*, \widehat{\gamma}_{2s}^*, \ldots, \widehat{\gamma}_{(J-1)s}^*, \widehat{b}_{1s}^*, \widehat{b}_{2s}^*, \ldots, \widehat{b}_{Ns}^*\right)$ denote the bootstrap estimates, $s = 1, ..., B$. For $j = 1, 2, \ldots, (J-1)$, and $k = 1, 2, \ldots, N$, the bootstrap bias and standard error can be obtained as

$$bias_b(\widehat{\gamma}_j) = \overline{\widehat{\gamma}_j^*} - \widehat{\gamma}_j, \; bias_b(\widehat{b}_k) = \overline{\widehat{b}_k^*} - \widehat{b}_k,$$

$$SE_b(\widehat{\gamma}_j) = \sqrt{\frac{1}{B-1}\sum_{s=1}^{B}\left(\widehat{\gamma}_{js}^* - \overline{\widehat{\gamma}_j^*}\right)^2}, \; SE_b(\widehat{b}_k) = \sqrt{\frac{1}{B-1}\sum_{s=1}^{B}\left(\widehat{b}_{ks}^* - \overline{\widehat{b}_k^*}\right)^2},$$

where

$$\overline{\widehat{\gamma}_j^*} = \frac{1}{B}\sum_{s=1}^{B}\widehat{\gamma}_{js}^* \quad \text{and} \quad \overline{\widehat{b}_k^*} = \frac{1}{B}\sum_{s=1}^{B}\widehat{b}_{ks}^*.$$

Finally, a $100(1-\alpha)\%$ bootstrap CIs for γ_j, $j = 1, 2, \ldots, (J-1)$, can be calculated as $(LL_{\gamma_j}, UL_{\gamma_j})$, where

$$LL_{\gamma_j} = \widehat{\gamma}_j - bias_b(\widehat{\gamma}_j) - z_{\alpha/2}SE_b(\widehat{\gamma}_j) \quad \text{and} \quad UL_{\gamma_j} = \widehat{\gamma}_j - bias_b(\widehat{\gamma}_j) + z_{\alpha/2}SE_b(\widehat{\gamma}_j).$$

Bootstrap CIs for b_k, where $k = 1, 2, \ldots, N$, can be calculated similarly.

For percentile bootstrap CIs for, say $\gamma_j, j = 1, 2, \ldots, (J-1)$, the bootstrap estimates of $\widehat{\gamma}_j$ are first ordered in terms of magnitude:

$$\widehat{\gamma}_{j(1)}^* < \widehat{\gamma}_{j(2)}^* < \ldots < \widehat{\gamma}_{j(B)}^*.$$

Then, a $100(1-\alpha)\%$ percentile bootstrap CI for γ_j is $\left(\widehat{\gamma}_{j([\frac{\alpha B}{2}])}^*, \widehat{\gamma}_{j([(1-\frac{\alpha}{2})B])}^*\right)$. Similarly, percentile bootstrap CIs can be calculated for $b_k, k = 1, 2, \ldots, N$.

4.3.3 Choice of Cut Points

The number and position of the cut-points for constructing the PLA-based model need to be suitably chosen so that the model can closely approximate the underlying CHF but avoid overfitting. A large number of cut points would provide a close local approximation to the underlying CHF. However, apart from being computationally expensive, a close local approximation may also lead to overfitting, in which case it would be difficult to use the PLA-based model to predict future failures of components or systems.

One of the possible ways to choose the number and position of the cut-points is by looking at the plot of the nonparametric estimator of CHF. From such a plot, observing the areas where the nonparametric estimate changes significantly, one can determine the positions and number of cut-points.

More objectively, one can choose the positions of a given number of cut-points by maximizing the log-likelihood function. The procedure to find position of cut-points for $N = 2$ can be expressed as an algorithm as follows. Note that this algorithm can be extended for $N \geq 3$.

Algorithm:

- **Step 1:** The natural choices for $\tau_0^{(j)}$ and $\tau_2^{(j)}$ are $\min\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$ and $\max\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$, respectively.
- **Step 2:** Fix $0 < p_1 < p_2 < 1$.
- **Step 3:** Find the number of $y_1^{(j)}, \ldots, y_n^{(j)}$ that are between p_1-th and p_2-th sample quantiles of $\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$. Denote this number by l. Note that l does not depend on $j = 0, 1, \ldots, J-1$.
- **Step 4:** Set $a_{j1} = p_1$-th quantile of $\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$, $j = 0, 1, \ldots, J-1$.
- **Step 5:** Set LL_1= the value of log-likelihood function evaluated at MLE taking $\tau_1^{(j)} = a_{j1},\ j = 0, 1, \ldots, J-1$.
- **Step 6:** Set $a_{j2} = \min\left\{y_i^{(j)} > a_{j1}; i = 1, 2, \ldots, n\right\},\ j = 0, 1, \ldots, J-1$.
- **Step 7:** Set LL_2= the value of log-likelihood function evaluated at MLE taking $\tau_1^{(j)} = a_{j2},\ j = 0, 1, \ldots, J-1$.
- **Step 8:** Repeat the steps 5 and 6 to obtain $LL_1, LL_2, \ldots, LL_l$.
- **Step 9:** Set $k^* = \underset{1 \leq k \leq l}{arg\,max}\ LL_k$.

- **Step 10:** The final cut points are $\tau_1^{(j)} = a_{jk^*}$, $j = 0, 1, \ldots, J-1$.

Note that the value of p_1 should not be very close to 0 to avoid unreliable estimates of model parameters for the initial choice of $\tau_1^{(j)}$ due to inadequate number of data points between $\tau_0^{(j)}$ and $\tau_1^{(j)}$, $j = 1, 2, \ldots, J-1$. Similarly, the value of p_2 should not be very close to 1. For implementation, we notice that reasonable choices of p_1 and p_2 are 0.3 and 0.7, respectively.

As mentioned, the above algorithm can be extended for $N \geq 3$ with some modification as indicated below. For example, assume that $N = 3$. As before natural choices of $\tau_0^{(j)}$ and $\tau_3^{(j)}$ are $\min\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$ and $\max\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$, respectively, $j = 0, 1, \ldots, J-1$. To place the cut-points, $\tau_1^{(j)}$ and $\tau_2^{(j)}$, one can start with fixing $0 < p_1 < p_2 < 1$ and $p_1 = q_1 < q_2 < p_2$. Initially, place $\tau_i^{(j)}$ at q_i-th quantile of $\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$, $i = 1, 2$, $j = 0, 1, \ldots, J-1$. It is recommended to place cut-points such that there are at least m data points between any two consecutive cut-points. The value of m should be chosen so that the estimates of model parameters are reliable and the computation time is reasonable. For this set of cut-points, the value of the log-likelihood function at the MLEs of the unknown parameters is calculated. Next, for $j = 0, 1, \ldots, J-1$, keeping the positions of the cut-points $\tau_1^{(j)}$ fixed, the position of each of the cut-points $\tau_2^{(j)}$ is changed to the next ordered data point, and the value of log-likelihood at MLEs is calculated for the current set of cut-points. This procedure is continued till the position of $\tau_2^{(j)}$ reaches p_2-th quantile of $\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$, $j = 0, 1, \ldots, J-1$. Then, for $j = 0, 1, \ldots, J-1$, the position of $\tau_1^{(j)}$ is changed to next ordered data point, $\tau_2^{(j)}$ is placed such a way that there are m data points between $\tau_1^{(j)}$ and $\tau_2^{(j)}$ and calculate the value of log-likelihood function at MLEs. Now, $\tau_2^{(j)}$ is moved similarly, and the process is continued as before. The procedure is stopped when there is no scope to move the cut-points any further, *i.e.*, $\tau_2^{(j)}$ is at p_2-th quantile of $\left\{y_1^{(j)}, \ldots, y_n^{(j)}\right\}$ and $\tau_1^{(j)}$ is m data points away from $\tau_2^{(j)}$. Finally, among all these possible choices of cut-points, the set that maximizes the log-likelihood function is chosen as the optimal set of cut-points.

4.4 Estimation of Important Reliability Characteristics

A goal of fitting a model to load-sharing data, naturally, is an accurate estimation of the reliability characteristics of load-sharing systems. As the PLA-based model provides a good fit for load-sharing data due to its flexible nature, it is natural that different important reliability characteristics can also be estimated quite accurately under this model. In this section, we develop estimates of reliability characteristics such as the quantile function, MTTF, RMT, and MRT of load-sharing systems under the PLA-based model. Details of these derivations are given in Appendix 4.A for interested readers.

Under the PLA-based model, the quantile function of $Y^{(j)}$ which is the system lifetime between the j-th and $(j+1)$-st failed components, $j = 0, ..., J-1$, is given by

$$\eta(p) = \inf\left\{y \in \mathbb{R} : G^{(j)}(y) \geq p\right\}, \quad 0 < p < 1,$$

where $G^{(j)}(y) = 1 - e^{-(J-j)\Lambda^{(j)}(y)}$. Using the expression of $\Lambda^{(j)}(y)$ given in Section 4.2, it is possible to work out an explicit formula for the quantile function $\eta(p)$, as follows:

$$\eta(p) = \begin{cases} \tau_{k-1}^{(j)} - \frac{\log(1-p)}{(J-j)\gamma_j b_k} - \frac{1}{b_k} \cdot \sum_{\ell=1}^{k-1} b_\ell(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}) & \text{if } p \in \left[G^{(j)}(\tau_{k-1}^{(j)}), G^{(j)}(\tau_k^{(j)})\right), \\ & k = 1, 2, \ldots, N \\ \tau_{N-1}^{(j)} - \frac{\log(1-p)}{(J-j)\gamma_j b_N} - \frac{1}{b_N} \cdot \sum_{\ell=1}^{N-1} b_\ell(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}) & \text{if } p \in \left[G^{(j)}(\tau_N^{(j)}), 1\right). \end{cases}$$

The MTTF of a load-sharing system is the expected time the system operates till its failure. Let T denote the system failure time; then, $T = \sum_{j=0}^{J-1} Y^{(j)}$. The MTTF of a load-sharing system under the PLA-based model is given by

$$E(T) = \sum_{j=0}^{J-1} \sum_{s=1}^{N} \left\{ \frac{e^{-\kappa_{j,s-1}} - e^{-\kappa_{j,s}}}{(J-j)\gamma_j b_\ell} \right\},$$

where

$$\kappa_{j,s} = (J-j)\gamma_j \sum_{\ell=1}^{s} b_\ell \left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right).$$

The RMT of a system is the probability that the system will operate till a desired time t_0; it is calculated as the survival probability of the system at time t_0, i.e., $S(t_0) = P(T > t_0) = P\left(\sum_{j=0}^{J-1} Y^{(j)} > t_0\right)$. An explicit expression for RMT may be derived by using the distribution of the system lifetime T.

However, as $Y^{(j)}$s, $j = 0, ..., J-1$ are independent but not identically distributed, it isn't easy to obtain an explicit expression for the distribution of the system lifetime T. It is evident from the moment generating function $\phi_T(t)$ of T, which, under the PLA-based model, is given by

$$\phi_T(t) = \prod_{j=0}^{J-1} \sum_{s=1}^{N} \frac{(J-j)b_s\gamma_j}{(J-j)b_s\gamma_j - t} \left(e^{t\tau_{s-1}^{(j)} - \kappa_{j,s-1}} - e^{t\tau_s^{(j)} - \kappa_{j,s}}\right)$$

if $t < \gamma_1 b_N$. From here, it is clear that it is difficult to find the RMT analytically under this model. However, for this model, RMT can be estimated using Monte Carlo simulations.

For a Monte Carlo estimate of the RMT at a pre-specified time t_0, one needs to generate R data points t_i, $i = 1, 2, \ldots, R$ as realisations of the system lifetime T, and find $\frac{R(t_0)}{R}$, where $R(t_0)$ is the number of realisations of the system lifetime that exceeds t_0. For a reasonably good estimate of RMT, a large value of R should be used.

The MRT of a system is the expected additional time the system will survive if it has already survived a given time t. That is,

$$\text{MRT}(t) = E(T - t|T > t) = \int_t^{\infty} s f_{T|T>t}(s) ds - t.$$

Therefore, the analytical derivation of MRT requires the truncated distribution of the system lifetime T, and it is difficult to obtain the truncated distribution of T in this case. Instead, an estimate of the MRT can be given using Monte Carlo simulations. We generate R data points t_i^*, $i = 1, 2, \ldots, R$ as realisations of the truncated lifetime $T|T > t$, and a Monte Carlo estimate of the MRT for load-sharing systems under the PLA-based model is then given by

$$\widehat{\text{MRT}}(t) = \frac{\sum_{i=1}^{R} t_i^*}{R} - t.$$

4.5 Simulation Study

In this section, we present the results of a Monte Carlo simulation study that examines the performance of the proposed PLA-based model in two directions. First, based on samples generated from a parent process with piecewise linear CHF, we assess the performance of the proposed estimation method that is presented in Section 4.3. Then, the efficacy of the PLA-based model in fitting data generated from a parent process represented by some parametric models is also assessed. For the simulations, we consider two-component load-sharing systems.

4.5.1 Assessing Performance of the Estimation Method

To assess the performance of the estimation methods, we consider an underlying cumulative hazard that is made up of two linear pieces. To this effect, we generate samples from the model specified by Eqs. (4.1) and (4.2) with $J = 2$ and $N = 2$. The true parameter values are taken to be $b_1 = 0.01, 0.05$; $b_2 = 0.1, 0.5$; $\gamma_1 = 5$; $\tau_1^{(0)} = \frac{\ln 2}{2b_1}$; $\tau_1^{(1)} = \frac{\ln 2}{\gamma_1 b_1}$. The estimation is performed based on samples of size $n = 25,\ 50,\ 100$ and 200. The AE and MSE of the MLEs based on 5000 Monte Carlo replications are reported in Tables 4.1, 4.2, and 4.3. The CP and AL of 95% CIs are also reported in the same tables. In these tables, ~ 0 indicates that the corresponding quantity is less than 10^{-3}. From the Tables 4.1, 4.2

and 4.3, we observe that the average estimates of γ_1, b_1 and b_2 are very close to the true values, and the MSEs are quite small as desired. It is also noticed that the performance of all the constructed CIs is satisfactory, although the results improve when the sample size increases. These results demonstrate that the proposed inferential techniques can accurately estimate the parameters of the PLA-based model.

Table 4.1: Performance measures for estimates of γ_1

n	b_1	b_2	AE	MSE	P. bootstrap		Bootstrap	
					CP	AL	CP	AL
25	0.01	0.1	5.1564	1.6152	100.00	4.55	75.28	4.58
		0.5	5.4353	7.7556	100.00	9.36	79.52	10.25
	0.05	0.1	5.1561	1.6190	98.96	5.12	89.86	5.18
		0.5	5.1564	1.6153	100.00	4.55	75.28	4.58
50	0.01	0.1	5.0607	0.7212	99.98	3.24	81.50	3.26
		0.5	5.0579	0.8539	100.00	4.45	93.02	4.89
	0.05	0.1	5.0534	0.7252	98.98	3.35	88.64	3.37
		0.5	5.0607	0.7212	99.98	3.24	81.50	3.26
100	0.01	0.1	5.0231	0.3389	99.94	2.20	83.58	2.22
		0.5	5.0178	0.2881	99.94	2.13	86.68	2.20
	0.05	0.1	5.0210	0.3557	98.52	2.31	88.60	2.32
		0.5	5.0231	0.3389	99.94	2.20	83.58	2.22
200	0.01	0.1	5.0144	0.1472	99.90	1.50	85.06	1.51
		0.5	5.0127	0.1374	99.86	1.44	84.42	1.45
	0.05	0.1	5.0145	0.1698	98.56	1.62	89.56	1.62
		0.5	5.0144	0.1472	99.90	1.50	85.06	1.51

4.5.2 Assessing Efficacy of the PLA-based Model in Fitting Data from Other Models

Now, we examine the robustness of the PLA-based model in the following manner. We generate load-sharing data from some parametric models, and then fit the PLA-based model to the data. The model fit is then assessed with respect to an integrated measure that is suitably defined to reflect the quality of approximation provided by the PLA-based model. The measure, which we call the AIE, is as follows. For $j = 0,\, 1$, let $S^{(j)}_{TGP}(\cdot)$ and $H^{(j)}_{TGP}(\cdot)$ denote the SF and CHF of the lifetimes between j-th and $(j+1)$-st failures. Also, assume that the estimated SF and CHF based on PLA-based model are denoted by $\widehat{S}^{(j)}_{PLA}(\cdot)$ and $\widehat{H}^{(j)}_{PLA}(\cdot)$, respectively. Then the AIE, based on the SF and CHF, respectively,

Table 4.2: Performance measures for estimates of b_1

n	b_1	b_2	AE	MSE	P. bootstrap		Bootstrap	
					CP	AL	CP	AL
25	0.01	0.1	0.0120	~ 0	82.50	0.01	92.36	0.01
		0.5	0.0123	~ 0	92.80	0.01	84.90	0.01
	0.05	0.1	0.0569	0.0003	91.10	0.08	95.90	0.08
		0.5	0.0598	0.0003	82.50	0.06	92.36	0.06
50	0.01	0.1	0.0113	~ 0	80.98	0.01	91.10	0.01
		0.5	0.0116	~ 0	80.40	0.01	92.40	0.01
	0.05	0.1	0.0526	~ 0	95.74	0.04	94.36	0.04
		0.5	0.0563	0.0001	80.98	0.04	91.10	0.04
100	0.01	0.1	0.0108	~ 0	81.66	0.01	92.88	0.01
		0.5	0.0110	~ 0	68.70	0.01	93.96	0.01
	0.05	0.1	0.0513	~ 0	95.96	0.02	93.70	0.02
		0.5	0.0539	~ 0	81.66	0.03	92.88	0.03
200	0.01	0.1	0.0105	~ 0	81.74	0.01	93.48	0.01
		0.5	0.0106	~ 0	74.08	0.01	94.56	0.01
	0.05	0.1	0.0508	~ 0	95.14	0.02	93.08	0.02
		0.5	0.0525	~ 0	81.74	0.02	93.48	0.02

are defined as

$$AIE_{SF}^{(j)} = \frac{1}{R}\sum_{k=1}^{R}\frac{1}{y_M^{(j)} - y_m^{(j)}}\int_{y_m^{(j)}}^{y_M^{(j)}}\left|S_{TGP}^{(j)}(t) - \widehat{S}_{PLA}^{(j)}(t)\right| dt,$$

$$AIE_{CHF}^{(j)} = \frac{1}{R}\sum_{k=1}^{R}\frac{1}{y_M^{(j)} - y_m^{(j)}}\int_{y_m^{(j)}}^{y_M^{(j)}}\left|H_{TGP}^{(j)}(t) - \widehat{H}_{PLA}^{(j)}(t)\right| dt,$$

where $y_m^{(j)} = \min\left\{y_1^{(j)}, y_2^{(j)}, \ldots, y_n^{(j)}\right\}$, $y_M^{(j)} = \max\left\{y_1^{(j)}, y_2^{(j)}, \ldots, y_n^{(j)}\right\}$, $j = 0, 1$ and R is the number of replications.

For generating load-sharing data from parametric models, two scenarios are considered:

(a) Case - 1: This case assumes that the lifetimes of each component of a two-component load-sharing system are independent and identically distributed Weibull random variables, with the shape parameter α and scale parameter β when both components are working. After the first failure, the lifetime of the surviving component is assumed to follow a Weibull distribution with the same shape parameter α, but a different scale parameter $k\beta$, where $k > 2$ is to ensure the increase of load on the surviving component. For $\beta = 1$, $k = 3$, we take $\alpha = 1$ and 1.5. Here, the value of R is taken to be 5000.

Table 4.3: Performance measures for estimates of b_2

n	b_1	b_2	AE	MSE	P. bootstrap CP	P. bootstrap AL	Bootstrap CP	Bootstrap AL
25	0.01	0.1	0.0978	0.0009	88.00	0.11	85.24	0.12
		0.5	0.4592	0.0352	77.32	0.62	77.94	0.69
	0.05	0.1	0.1083	0.0009	96.32	0.12	94.94	0.12
		0.5	0.4891	0.0222	88.00	0.56	85.24	0.58
50	0.01	0.1	0.1007	0.0004	96.26	0.09	93.48	0.08
		0.5	0.5091	0.0096	97.00	0.52	97.16	0.52
	0.05	0.1	0.1059	0.0004	94.60	0.08	95.76	0.08
		0.5	0.5037	0.0096	96.26	0.42	93.48	0.42
100	0.01	0.1	0.1006	0.0002	96.60	0.05	94.00	0.05
		0.5	0.5067	0.0038	97.12	0.27	95.12	0.27
	0.05	0.1	0.1030	0.0002	95.46	0.05	94.70	0.05
		0.5	0.5030	0.0042	96.60	0.27	94.00	0.26
200	0.01	0.1	0.1002	0.0001	96.72	0.03	93.76	0.03
		0.5	0.5030	0.0017	96.52	0.17	93.60	0.17
	0.05	0.1	0.1011	0.0001	96.10	0.03	94.08	0.04
		0.5	0.5010	0.0018	96.72	0.17	93.76	0.17

(b) Case - 2: This case assumes that the component lifetimes are independent and identically distributed with a quadratic CHF $\kappa_1 t + \kappa_2 t^2$ when both components are working. After the first failure, the lifetime of the surviving component is assumed to follow a quadratic CHF with different parameters $\tilde{\kappa}_1$ and $\tilde{\kappa}_2$. We take several values of the parameters κ_1, κ_2, $\tilde{\kappa}_1$, and $\tilde{\kappa}_2$, ensuring the fact that the CHF increases after one component fails in the system. Here, for all cases, we take the value of R as 5000.

The numerical results are reported in Tables 4.4, 4.5, and 4.6. For all cases, it is observed that the values of AIE based on SF and CHF are reasonably small, indicating that the PLA-based model provides quite a satisfactory approximation to the data generated from different parent populations.

4.6 Data Analysis

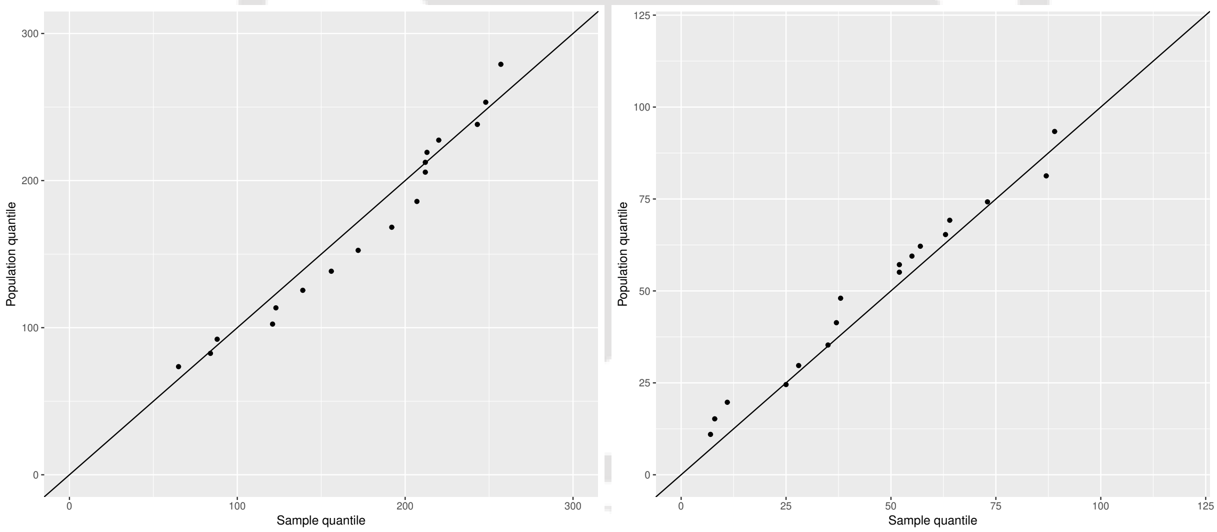
4.6.1 Analysis of Two-motor Data

In this section, we analyse the two-motor data introduced in Subsection 1.2.3 by applying the PLA-based model. We consider three cut points for the PLA-based model (i.e.,

Table 4.4: AIE based on SF and CHF for Weibull distribution with $k = 3$, $\beta = 1$

n	α	$AIE_{SF}^{(0)}$	$AIE_{SF}^{(1)}$	$AIE_{CHF}^{(0)}$	$AIE_{CHF}^{(1)}$
25	1.0	0.0545	0.0465	0.1755	0.3434
	1.5	0.0642	0.0681	0.1636	0.3283
50	1.0	0.0379	0.0291	0.1503	0.2981
	1.5	0.0436	0.0434	0.1329	0.2633
100	1.0	0.0266	0.0183	0.1282	0.2541
	1.5	0.0326	0.0301	0.1231	0.2440
200	1.0	0.0182	0.0117	0.1056	0.2100
	1.5	0.0276	0.0227	0.1388	0.2742

$N = 2$) for this data. The estimates of the model parameters are reported in Table 4.7. The Q-Q plots for $Y^{(0)}$ and $Y^{(1)}$ are given in Figures 4.2a and 4.2b, respectively. These plots indicate that the PLA-based model fits the data quite adequately. The plots of the estimated SF and CHF under the PLA-based model are also given in Figures 4.3 and 4.4, respectively. A Kolmogorov-Smirnov type test has been performed to test the following



(a) Q-Q plot for $Y^{(0)}$ (b) Q-Q plot for $Y^{(1)}$

Figure 4.2: Q-Q plots for the two-motor data

hypotheses:

H_0 : True model is specified by Eqs. (4.1) and (4.2)

against

H_1 : True model is not specified by Eqs. (4.1) and (4.2)

Table 4.5: AIE of the SF and the CHF of quadratic distribution for $\kappa_1 = 0.5,\ \tilde{\kappa}_1 = 2\kappa_1 = 1, \tilde{\kappa}_2 > 2\kappa_2$

n	κ_2	$\tilde{\kappa}_2$	$AIE^{(0)}_{SF}$	$AIE^{(1)}_{SF}$	$AIE^{(0)}_{CHF}$	$AIE^{(1)}_{CHF}$
25	0.50	1.50	0.0535	0.0562	0.1468	0.2954
		2.00	0.0534	0.0591	0.1470	0.2965
	0.70	1.50	0.0548	0.0556	0.1464	0.2953
		2.00	0.0546	0.0584	0.1465	0.2958
50	0.50	1.50	0.0380	0.0368	0.1262	0.2536
		2.00	0.0380	0.0389	0.1261	0.2555
	0.70	1.50	0.0389	0.0363	0.1261	0.2524
		2.00	0.0388	0.0383	0.1258	0.2539
100	0.50	1.50	0.0289	0.0262	0.1185	0.2506
		2.00	0.0289	0.0281	0.1178	0.2575
	0.70	1.50	0.0301	0.0257	0.1217	0.2465
		2.00	0.0299	0.0274	0.1206	0.2528
200	0.50	1.50	0.0245	0.0207	0.1320	0.2882
		2.00	0.0243	0.0226	0.1304	0.3020
	0.70	1.50	0.0263	0.0201	0.1400	0.2815
		2.00	0.0259	0.0219	0.1377	0.2945

based on the test statistics

$$T_n = \max_{1\le i\le n}\left|\widehat{G}^{(0)}\left(Y^{(0)}_{i:n}\right) - \frac{i}{n}\right| + \max_{1\le i\le n}\left|\widehat{G}^{(1)}\left(Y^{(1)}_{i:n}\right) - \frac{i}{n}\right|,$$

where $\widehat{G}^{(j)}(\cdot)$ is the estimated CDF corresponding to PLA-based model, and $Y^{(j)}_{i:n}$ is the i-th order statistics corresponding to $Y^{(j)}_i$, $j = 0,\ 1,\ i = 1,\ 2,\ \ldots,\ n$. The observed value of the test statistic T_n is found to be 0.414 based on this data. The Monte Carlo estimate of the corresponding p-value is 0.71. Therefore, the null hypothesis cannot be rejected at a significance level of 0.05, and we conclude that it is quite reasonable to use the PLA-based model for this data.

We compare the PLA-based model with some other models that have been used recently for load-sharing data, such as the extended sequential order statistics model by Pesch et al. [45], frailty-based model by Asha et al. [5], the model based on the GFB distribution by Franco et al. [22], and the exponential and Weibull models by Park [40]. The comparison is done with respect to AIC. For the PLA-based model, the penalty term of AIC is computed by considering the number of cut-points (in this case, it is 2) along with the number of model parameters (in this case, it is 3). Table 4.8 presents the AIC values of different models for the two-motor load-sharing data. It may be clearly observed

Table 4.6: AIE of the SF and CHF of quadratic distribution for $\tilde{\kappa}_1 > 2\kappa_1, \kappa_2 = 0.5,\ \tilde{\kappa}_2 = 2\kappa_2 = 1$

n	κ_1	$\tilde{\kappa}_1$	$AIE^{(0)}_{SF}$	$AIE^{(1)}_{SF}$	$AIE^{(0)}_{CHF}$	$AIE^{(1)}_{CHF}$
25	0.50	1.50	0.0546	0.0487	0.1475	0.3051
		2.00	0.0556	0.0476	0.1489	0.3170
	0.70	1.50	0.0529	0.0496	0.1501	0.3048
		2.00	0.0536	0.0480	0.1511	0.3145
50	0.50	1.50	0.0388	0.0313	0.1283	0.2570
		2.00	0.0397	0.0307	0.1309	0.2672
	0.70	1.50	0.0372	0.0314	0.1284	0.2567
		2.00	0.0377	0.0304	0.1301	0.2644
100	0.50	1.50	0.0306	0.0210	0.1265	0.2290
		2.00	0.0319	0.0206	0.1325	0.2285
	0.70	1.50	0.0278	0.0210	0.1184	0.2331
		2.00	0.0285	0.0198	0.1227	0.2271
200	0.50	1.50	0.0271	0.0152	0.1471	0.2305
		2.00	0.0292	0.0146	0.1578	0.2116
	0.70	1.50	0.0229	0.0153	0.1277	0.2422
		2.00	0.0242	0.0138	0.1369	0.2162

Table 4.7: Parameter estimates of the PLA-based model for the two-motor data

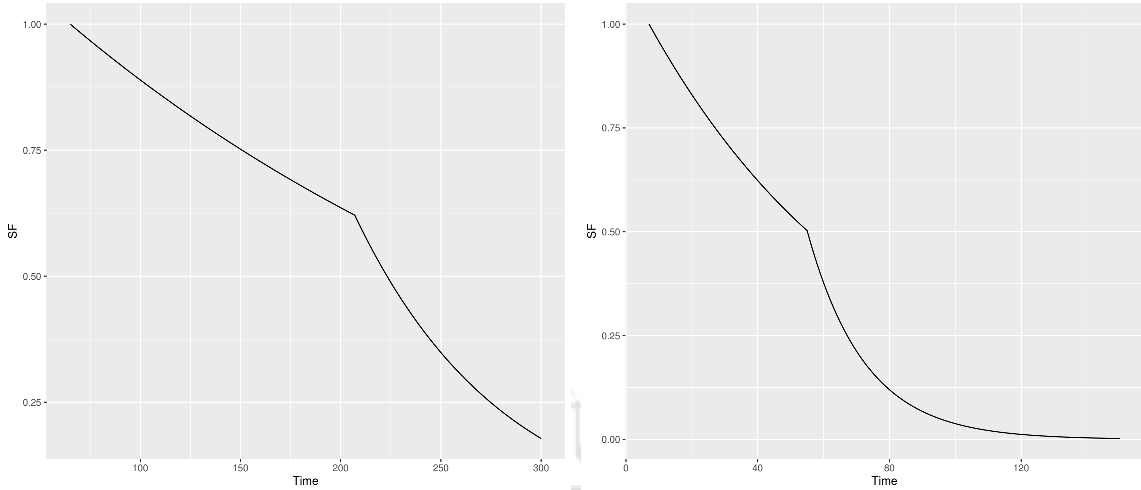
Parameter	MLE	Std. Error	P. Bootstrap	Bootstrap
γ_1	4.271	1.1901	(3.075, 8.028)	(0.846, 5.817)
b_1	0.003	0.0008	(0.002, 0.006)	(0.001, 0.005)
b_2	0.013	0.0039	(0.006, 0.021)	(0.008, 0.023)

that the AIC value for the PLA-based model turns out to be 373.34, which is the lowest among all. This implies that among these recent models considered here, the PLA-based model is the most suitable for the two-motor data.

Table 4.8: Comparison between various load-sharing models for the two-motor data

Model	No. of Parameters	AIC
Frailty-based model [Asha et al. [5]]	6	480.50
Generalised Freund bivaraite model [Franco et al. [22]]	5	409.65
Exponential model [Park [40]]	3	405.30
Weibull model [Park [40]]	5	378.42
Sequential order statistics model [Pesch et al. [45]]	4	431.72
Proposed PLA-based model	5	373.34

For the PLA-based model, the estimated value of γ_1 is 4.2712, which empirically



(a) Plot of SF for $Y^{(0)}$ (b) Plot of SF for $Y^{(1)}$

Figure 4.3: Plots of SFs for the two-motor data

implies that the load-sharing model is quite appropriate in this case. The same comment can also be made from the plots, by noting that the plot of the SF of the distribution of time between the first and second failed components diminishes to zero more quickly, compared to that of the lifetime of the component that failed first in Figure 4.3. The reliability characteristics of the two-motor load-sharing systems are also estimated by using the expressions and techniques described in Section 4.4. The MTTF is calculated to be 221.36 days. Monte Carlo estimates using 10000 replications of the MRT and RMT are calculated at different sample percentile points of the system failure times and are presented in Table 4.9.

Table 4.9: Mean residual time and reliability in mission time for the two-motor data

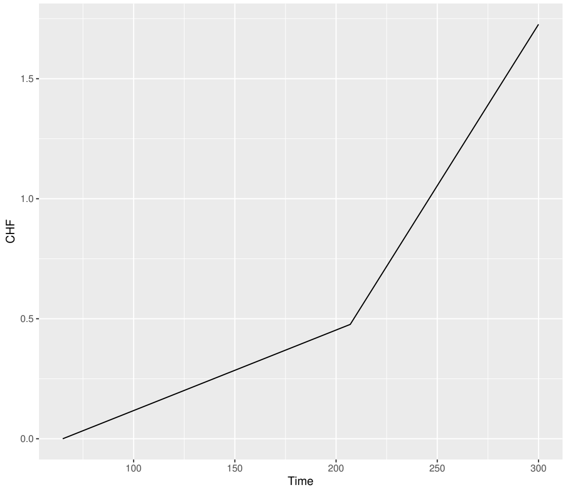
t_0	MRT(t_0)	RMT(t_0)
102.00	124.223	0.963
167.50	88.678	0.706
227.50	60.646	0.466
272.50	42.794	0.271
350.00	36.919	0.044

We perform Wilcoxon signed-rank test to check for stochastic ordering between two distributions due to load-sharing effect. We take the null hypothesis as:

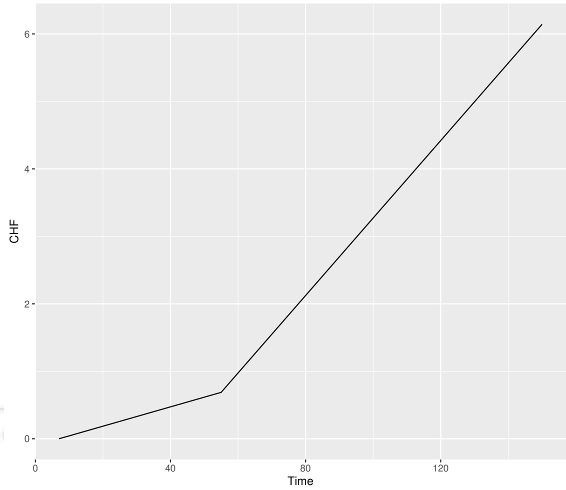
$$H_0 : F_{Y^{(0)}}(t) = F_{Y^{(1)}}(t) \quad \text{for all } t,$$

against the one-sided alternative

$$H_1 : F_{Y^{(1)}}(t) > F_{Y^{(0)}}(t) \text{ for some } t,$$



(a) Plot of CHF for $Y^{(0)}$



(b) Plot of CHF for $Y^{(1)}$

Figure 4.4: Plots of CHFs for the two-motor data

implying that $Y^{(1)}$ is stochastically smaller than $Y^{(0)}$. Using the Wilcoxon signed-rank test (one-sided), we obtain a test statistic $V = 2$ and p value $= 0.00015$. Since $p < 0.05$, we reject H_0 and conclude that the lifetime distribution of the lifetimes after failure is significantly smaller than that before failure. This provides strong evidence of a load-sharing effect.

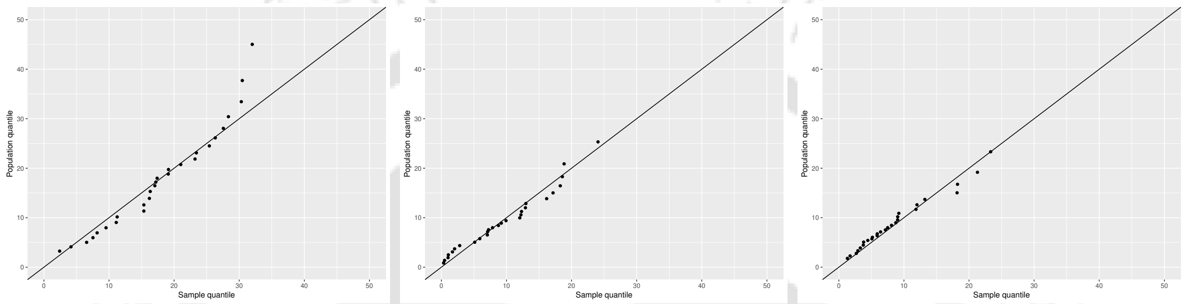
4.6.2 Analysis of Three-player Basketball Data

Analysis of the three-player basketball data, which is introduced in Subsection 1.2.2, is presented here by applying the PLA-based model. Several researchers have analysed this dataset as three sets of two-component load-sharing datasets due to the structural restriction of those load-sharing models; see Deshpande et al. [19], Asha et al. [4]. However, we employ the proposed PLA-based model for the original three-component load-sharing data as presented in Kvam and Peña [32].

We consider three cut points for the PLA-based model (i.e., $N = 2$) for this data. The point and interval estimates of the model parameters are reported in Table 4.10. Note that the estimated value of γ_1 in this case is 2.471 and γ_2 is 5.291, which empirically implies the presence of a load-sharing phenomenon in this data. The Q-Q plots for $Y^{(0)}$, $Y^{(1)}$ and $Y^{(2)}$ are given in Figures 4.5a, 4.5b and 4.5c, respectively. These Q-Q plots show that the PLA-based model fits the data quite well. The plots of the estimated SF and CHF under the PLA-based model are also given in Figures 4.6 and 4.7, respectively; which also demonstrate the presence of the load-sharing phenomenon in the data as we can see that the plot of the SF of the distribution of time between the first and second failed components diminishes to zero more quickly, compared to that of the lifetime of the component that failed first and the SF of the distribution of time between second and third failure diminishes to zero more quickly than that of the first and second failed components in Figure 4.6. For assessing the goodness-of-fit of the PLA-based model for

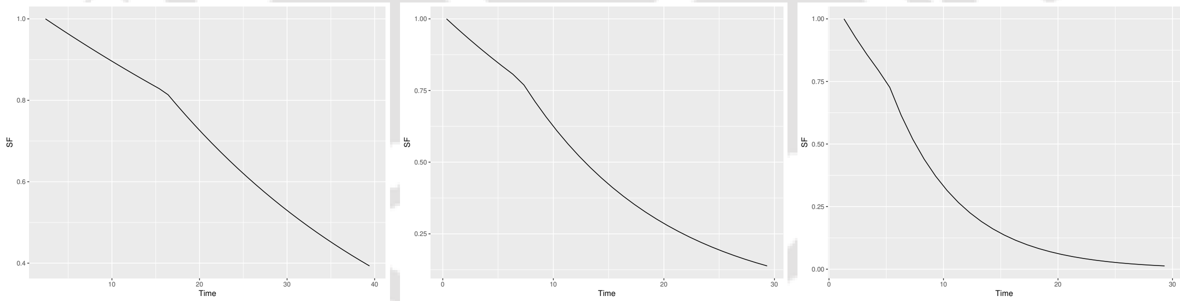
Table 4.10: Parameter estimates of the PLA-based model for the basketball data

Parameter	MLE	Std. Error	P. Bootstrap	Bootstrap
γ_1	2.471	0.614	(1.672, 4.026)	(1.089, 3.495)
γ_2	5.291	1.526	(3.893, 9.587)	(1.321, 7.303)
b_1	0.015	0.004	(0.011, 0.025)	(0.005, 0.019)
b_2	0.032	0.007	(0.021, 0.047)	(0.018, 0.045)



(a) Q-Q plot for $Y^{(0)}$ (b) Q-Q plot for $Y^{(1)}$ (c) Q-Q plot for $Y^{(2)}$

Figure 4.5: Q-Q plots for three-player Basketball data

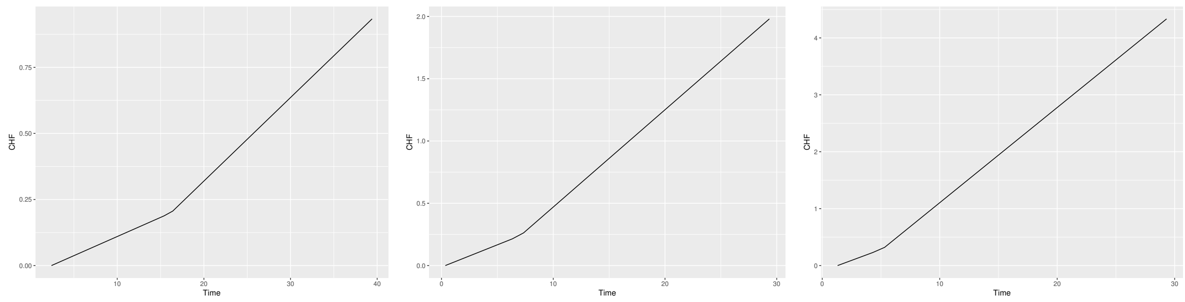


(a) Plot of SF for $Y^{(0)}$ (b) Plot of SF for $Y^{(1)}$ (c) Plot of SF for $Y^{(2)}$

Figure 4.6: Plots of SFs for three-player Basketball data

the three-player basketball data, a Kolmogorov-Smirnov type test has been performed. The observed value of the test statistic T_n is found to be 0.836, with a Monte Carlo estimate of the corresponding p-value as 0.67. Therefore, the null hypothesis for the suitability of the PLA-based model cannot be rejected at a significance level of 0.05, and we conclude that it is reasonable to use the PLA-based model for this data.

We compared the PLA-based model with some other recently used load-sharing models that can be applied to a three-component load-sharing system, such as the extended sequential order statistics model by Pesch et al. [45], and the exponential and Weibull



(a) Plot of CHF for $Y^{(0)}$ (b) Plot of CHF for $Y^{(1)}$ (c) Plot of CHF for $Y^{(2)}$

Figure 4.7: Plots of CHFs for three-player Basketball data

models by Park [40], on the basis of the AIC values of the models. For the PLA-based model, the penalty term of AIC is computed by considering the number of cut-points (in this case, it is 3) along with the number of model parameters (in this case, it is 4). Table 4.11 presents the AIC values of different models for the three-player basketball load-sharing data. It may be clearly observed from Table 4.11 shows that the AIC value for the PLA-based model turns out to be 566.46, which is the lowest among all. This implies that among these recent models considered here, the PLA-based model is the most suitable one for the three-player basketball data. Finally, the reliability characteristics for the three-

Table 4.11: Comparison of various load-sharing models for the basketball data

Model	No. of Parameters	AIC
Exponential model [Park [40]]	6	595.08
Weibull model [Park [40]]	9	573.05
Sequential order statistics model [Pesch et al. [45]]	9	712.52
Proposed PLA-based model	7	566.47

player basketball data are calculated, using techniques described in Section 4.4. The MTTF is calculated to be 37.31 minutes. Monte Carlo estimates using 10000 replications of the MRT and RMT are also calculated at different sample percentile points of the system failure times and are presented in Table 4.12. Practically, these have implications in predicting fouls in the games. For the datasets we will check for $Y^{(0)}$ and $Y^{(1)}$ and

Table 4.12: Mean residual time and reliability in mission time for the basketball data

t_0	MRT(t_0)	RMT(t_0)
19.09	20.404	0.904
32.98	14.396	0.565
40.18	12.889	0.364
42.46	12.496	0.323
46.44	12.051	0.250

again perform the test for $Y^{(1)}$ and $Y^{(2)}$ seperately. We first take the null hypothesis as:

$$H_0 : F_{Y^{(0)}}(t) = F_{Y^{(1)}}(t) \quad \text{for all } t,$$

against the one-sided alternative

$$H_1 : F_{Y^{(1)}}(t) > F_{Y^{(0)}}(t) \text{ for some } t,$$

implying that $Y^{(1)}$ is stochastically smaller than $Y^{(0)}$. Using the paired Wilcoxon signed-rank test (one-sided), we obtain a test statistic $V = 78$ and p-value $p = 0.00229$. Since $p < 0.05$, we reject H_0 and conclude that the lifetime distribution of the lifetime after first failure is significantly smaller than that before failure. This provides strong evidence of a load-sharing effect between the two components. We then take the null hypothesis as:

$$H_0 : F_{Y^{(1)}}(t) = F_{Y^{(2)}}(t) \quad \text{for all } t,$$

against the one-sided alternative

$$H_1 : F_{Y^{(2)}}(t) > F_{Y^{(1)}}(t) \text{ for some } t,$$

implying that $Y^{(2)}$ is stochastically smaller than $Y^{(1)}$. Using the Wilcoxon signed-rank test (one-sided), we obtain a test statistic $V = 174$ and p-value $= 0.2582$. Since $p > 0.05$, we can not reject H_0 at significance level 0.05. However, we can see that $Y^{(1)}$ is stochastically smaller than $Y^{(0)}$, and hence the basketball data exhibit a load-sharing phenomenon.

4.7 Further Extension Incorporating Covariates

In many reliability applications, information on concomitant variables (i.e., covariates) that include physical characteristics of the units being studied is available along with lifetimes. These covariates may provide valuable information about the distribution of component and system lifetimes. The proposed PLA-based model can be easily modified to incorporate covariates by using the Cox proportional hazards model, proposed by Cox [14]. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_p)$ denote the vector of covariates. Then the CHF given in Eq. (4.3) is modified as

$$\Lambda^{(j)}(t|\boldsymbol{x}) = e^{\boldsymbol{x}'\boldsymbol{\beta}} \left(\gamma_j \sum_{k=1}^{N} \left[\sum_{\ell=1}^{k-1} b_\ell \left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)} \right) + b_k \left(t - \tau_{k-1}^{(j)} \right) \right] \mathbf{1}_{[\tau_{k-1}^{(j)}, \tau_k^{(j)})}(t) \right),$$

for $j = 0, 1, 2, \ldots, J-1$. Here, $\boldsymbol{\beta}$ is the regression parameter vector, and $0 < b_1 < b_2 < \ldots < b_N$, $1 = \gamma_0 < \gamma_1 < \gamma_2 < \ldots < \gamma_{J-1}$.

Suppose the available data from n J-component load-sharing system, along with corresponding covariates, is of the form

$$Data = \left\{y_i^{(0)},\, y_i^{(1)},\, \ldots,\, y_i^{(J-1)},\, \boldsymbol{x}_i : i = 1,\, 2,\, \ldots,\, n\right\},$$

where $y_i^{(j)}$ is the observed system lifetime between j-th and $(j+1)$-st failed components for the i-th system, and $\boldsymbol{x}_i$ represents covariates. The log-likelihood function, ignoring the additive constant, is given by

$$l\left(\boldsymbol{\zeta}\right) = \sum_{k=1}^{N}\left[\left(\sum_{j=0}^{J-1} n_k^{(j)}\right)\ln b_k - \left(\sum_{j=0}^{J-1}(J-j)\gamma_j \widetilde{T}_k^{(j)}\right) b_k\right] + n\sum_{j=0}^{J-1}\ln\gamma_j + J\left(\sum_{i=1}^{n}\boldsymbol{x}_i\right)'\boldsymbol{\beta},$$

where $\boldsymbol{\zeta} = (\boldsymbol{\theta},\, \boldsymbol{\beta})$ and

$$\widetilde{T}_k^{(j)} = \sum_{i\in I_k^{(j)}} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\left(y_i^{(j)} - \tau_{k-1}^{(j)}\right) + \left(\sum_{l=k+1}^{N}\sum_{i\in I_l^{(j)}} e^{\boldsymbol{x}_i'\boldsymbol{\beta}}\right)\left(\tau_k^{(j)} - \tau_{k-1}^{(j)}\right),$$

for $k = 1,\, 2,\, \ldots,\, N$; $j = 0,\, 1,\, \ldots,\, J-1$. Here, $I_k^{(j)}$'s are as defined in Section 4.3.

4.8 Conclusions

In this chapter, a flexible PLA-based model is developed for analysing load-sharing data. Also, some important reliability characteristics, such as quantile function, RMT, MTTF, and MRT, are estimated under the proposed PLA-based model. The main advantages of the PLA-based model are that it is data-driven and does not depend on strong parametric assumptions for the underlying lifetime variables. Likelihood inference for the proposed model is discussed in detail. It is observed that for two-component load-sharing systems, explicit expressions for the MLEs of the model parameters can be obtained. Construction of CIs using bootstrap approaches is also discussed.

A Monte Carlo simulation study is carried out to examine (a) the performance of the methods of inference and (b) the efficacy of the PLA-based model to fit load-sharing data in general. It is shown that the PLA-based model performs quite satisfactorily in both cases. Analyses of the two-motor and three-player basketball data are presented as numerical illustrations. It is shown that the PLA-based model provides a better fit to these load-sharing data, compared to several load-sharing models that have been recently used for these datasets. In summary, in this chapter, an efficient modelling framework for load-sharing systems is discussed using a PLA-based model, and estimates of important reliability characteristics for load-sharing systems are developed in this setting.

Appendix

4.A Calculations of Some Important Reliability Characteristics

4.A.1 Derivation of the Quantile Function

Denote $p = G^{(j)}(y)$ for $y \in \left[\tau_{k-1}^{(j)}, \tau_k^{(j)}\right)$; then, $y = \eta(p)$ for $p \in \left[G^{(j)}(\tau_{k-1}^{(j)}), G^{(j)}(\tau_k^{(j)})\right)$, $k = 1, 2, \ldots, N$. Now,

$$p = 1 - e^{-(J-j)\gamma_j\left[\sum_{\ell=1}^{k-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right) + b_k\left(y - \tau_{k-1}^{(j)}\right)\right]}$$
$$\Longrightarrow y = \tau_{k-1}^{(j)} - \frac{\log(1-p)}{(J-j)\gamma_j b_k} - \frac{1}{b_k}\sum_{\ell=1}^{k-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right),$$

if $p \in \left[G^{(j)}(\tau_{k-1}^{(j)}), G^{(j)}(\tau_k^{(j)})\right)$ for $k = 1, 2, \ldots, N$. If $y \in \left[\tau_N^{(j)}, \infty\right)$, then $y = \eta(p)$ for $p \in \left[G^{(j)}(\tau_N^{(j)}), 1\right)$. Therefore,

$$p = 1 - e^{-(J-j)\gamma_j\left[\sum_{\ell=1}^{N-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right) + b_N\left(y - \tau_{N-1}^{(j)}\right)\right]}$$
$$\Longrightarrow y = \tau_{N-1}^{(j)} - \frac{\log(1-p)}{(J-j)\gamma_j b_N} - \frac{1}{b_N}\sum_{\ell=1}^{N-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right),$$

if $p \in \left[G^{(j)}(\tau_N^{(j)}), 1\right)$.

4.A.2 Derivation of MTTF

The MTTF of the system lifetime T is given by

$$E(T) = E\left(\sum_{j=0}^{J-1} Y^{(j)}\right)$$
$$= \sum_{j=0}^{J-1} E(Y^{(j)}), \tag{4.8}$$

where

$$E(Y^{(j)}) = \int_0^\infty P(Y^{(j)} > y)dy$$
$$= \int_0^{\tau_{N-1}^{(j)}} e^{-(J-j)\Lambda^{(j)}(y)}dy + \int_{\tau_{N-1}^{(j)}}^\infty e^{-(J-j)\Lambda^{(j)}(y)}dy$$

$$= I_1 + I_2 \text{ (say)}.$$

Here,

$$\begin{aligned}
I_1 &= \int_0^{\tau_{N-1}^{(j)}} \exp\left\{-(J-j)\gamma_j \sum_{k=1}^{N}\left[\sum_{\ell=1}^{k-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right) + b_k\left(y - \tau_{k-1}^{(j)}\right)\right]\mathbf{1}_{[\tau_{k-1}^{(0)},\tau_k^{(0)})}(y)\right\}dy \\
&= \sum_{s=1}^{N-1}\int_{\tau_{s-1}^{(j)}}^{\tau_s^{(j)}} \exp\left\{-(J-j)\gamma_j\left[\sum_{\ell=1}^{s-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right) + b_s\left(y - \tau_{s-1}^{(j)}\right)\right]\right\}dy \\
&= \sum_{s=1}^{N-1}\frac{1}{(J-j)\gamma_j b_s}\left[e^{-\kappa_{j,s-1}} - e^{-\kappa_{j,s}}\right],
\end{aligned}$$

and

$$\begin{aligned}
I_2 &= \int_{\tau_{N-1}^{(j)}}^{\infty} \exp\left\{-(J-j)\gamma_j \sum_{k=1}^{N}\left[\sum_{\ell=1}^{k-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right) + b_k\left(y - \tau_{k-1}^{(j)}\right)\right]\mathbf{1}_{[\tau_{k-1}^{(0)},\tau_k^{(0)})}(y)\right\}dy \\
&= \int_{\tau_{N-1}^{(j)}}^{\infty} \exp\left\{-(J-j)\gamma_j\left[\sum_{\ell=1}^{N-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right) + b_N\left(y - \tau_{N-1}^{(j)}\right)\right]\right\}dy \\
&= e^{-(J-j)\gamma_j\sum_{\ell=1}^{N-1} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right)}\int_{\tau_{N-1}^{(j)}}^{\infty} e^{-(J-j)\gamma_j b_N\left(y-\tau_{N-1}^{(j)}\right)}dy \\
&= \frac{e^{-\kappa_{j,N-1}}}{(J-j)\gamma_j b_N}.
\end{aligned}$$

Therefore,

$$E(Y^{(j)}) = \sum_{s=1}^{N}\left\{\frac{e^{-\kappa_{j,s-1}} - e^{-\kappa_{j,s}}}{(J-j)\gamma_j b_s}\right\}.$$

Now, from Eq. (4.8), we get the MTTF of a load-sharing system under the PLA-based model as

$$E(T) = \sum_{j=0}^{J-1}\sum_{s=1}^{N}\left\{\frac{e^{-\kappa_{j,s-1}} - e^{-\kappa_{j,s}}}{(J-j)\gamma_j b_\ell}\right\},$$

where

$$\kappa_{j,s} = (J-j)\gamma_j\sum_{\ell=1}^{s} b_\ell\left(\tau_\ell^{(j)} - \tau_{\ell-1}^{(j)}\right).$$

4.A.3 Derivation of the Moment Generating Function of System Lifetime

Note that the system lifetime MGF of T is $T = \sum_{j=0}^{J-1} Y^{(j)}$, where $Y^{(j)}$'s are independent for $j = 0, 1, \ldots, (J-1)$. Therefore, the MGF of T is

$$\phi_T(t) = \prod_{j=0}^{J-1} \phi_{Y^{(j)}}(t). \tag{4.9}$$

Now,

$$\begin{aligned}
\phi_{Y^{(j)}}(t) &= \int_0^{\tau_{N-1}^{(j)}} e^{ty}(J-j)\lambda^{(j)}(y)e^{-(J-j)\Lambda^{(j)}(y)}dy \\
&\quad + \int_{\tau_{N-1}^{(j)}}^{\infty} e^{ty}(J-j)\lambda^{(j)}(y)e^{-(J-j)\Lambda^{(j)}(y)}dy \\
&= I_3 + I_4 \text{ (say)}.
\end{aligned}$$

For $t \in \mathbb{R}$,

$$\begin{aligned}
I_3 &= \int_0^{\tau_{N-1}^{(j)}} e^{ty}(J-j)\gamma_j \sum_{k=1}^{N} b_k \mathbf{1}_{[\tau_{k-1}^{(j)}, \tau_k^{(j)})}(y) \\
&\qquad \times \exp\left\{-\sum_{k=1}^{N}\left[\kappa_{j,k-1} + (J-j)\gamma_j b_k\left(y - \tau_{k-1}^{(j)}\right)\right]\right\} dy \\
&= \sum_{s=1}^{N-1}(J-j)b_s\gamma_j \int_{\tau_{s-1}^{(j)}}^{\tau_s^{(j)}} e^{-\left\{\kappa_{j,s-1}+(J-j)\gamma_j b_s\left(y-\tau_{s-1}^{(j)}\right)-ty\right\}}dy \\
&= \sum_{s=1}^{N-1}\left\{(J-j)b_s\gamma_j e^{-(J-j)\gamma_j\sum_{\ell=1}^{s-1}b_\ell\left(\tau_\ell^{(j)}-\tau_{\ell-1}^{(j)}\right)}\left[\frac{e^{t\tau_{s-1}^{(j)}} - e^{-\left\{(J-j)\gamma_j b_s\left(\tau_s^{(j)}-\tau_{s-1}^{(j)}\right)-t\tau_s^{(j)}\right\}}}{(J-j)\gamma_j b_s - t}\right]\right\} \\
&= \sum_{s=1}^{N-1}\frac{(J-j)b_s\gamma_j}{(J-j)b_s\gamma_j - t}\left\{e^{-\left(\kappa_{j,s-1}-t\tau_{s-1}^{(j)}\right)} - e^{-\left(\kappa_{j,s}-t\tau_s^{(j)}\right)}\right\}.
\end{aligned}$$

For $t < (J-j)\gamma_j b_N$,

$$\begin{aligned}
I_4 = \int_{\tau_{N-1}^{(j)}}^{\infty} & e^{ty}(J-j)\gamma_j \sum_{k=1}^{N} b_k \mathbf{1}_{[\tau_{k-1}^{(j)}, \tau_k^{(j)})}(y) \\
&\times \exp\left\{-\sum_{k=1}^{N}\left[\kappa_{j,k-1} + (J-j)\gamma_j b_k\left(y - \tau_{k-1}^{(j)}\right)\right]\right\} dy
\end{aligned}$$

$$
\begin{aligned}
&= (J-j)b_N\gamma_j e^{-\kappa_{j,N-1}} \int_{\tau_{N-1}^{(j)}}^{\infty} e^{-\left\{(J-j)\gamma_j b_N\left(y-\tau_{N-1}^{(j)}\right)-ty\right\}} dy \\
&= (J-j)b_N\gamma_j e^{-\kappa_{j,N-1}} \left[\frac{e^{t\tau_{N-1}^{(j)}}}{(J-j)\gamma_j b_N - t}\right] \\
&= (J-j)b_N\gamma_j \frac{e^{t\tau_{N-1}^{(j)}-\kappa_{j,N-1}}}{(J-j)b_N\gamma_j - t}.
\end{aligned}
$$

Therefore, for $t < (J-j)\gamma_j b_N$,

$$
\phi_{Y^{(j)}}(t) = \sum_{s=1}^{N} \frac{(J-j)b_s\gamma_j}{(J-j)b_s\gamma_j - t} \left\{e^{-\left\{\kappa_{j,s-1}-t\tau_{s-1}^{(j)}\right\}} - e^{-\left\{\kappa_{j,s}-t\tau_{s}^{(j)}\right\}}\right\}.
$$

From Eq. (4.9), we get the moment generating function as

$$
\phi_T(t) = \prod_{j=0}^{J-1}\sum_{s=1}^{N} \frac{(J-j)b_s\gamma_j}{(J-j)b_s\gamma_j - t} \left(e^{t\tau_{s-1}^{(j)}-\kappa_{j,s-1}} - e^{t\tau_{s}^{(j)}-\kappa_{j,s}}\right)
$$

if $t < \gamma_1 b_N$.

CHAPTER 5

CONCLUSIONS AND FUTURE SCOPE

5.1 Conclusions

This thesis presents a detailed statistical analysis of load-sharing systems under various settings. Specifically, we solve three problems, each focusing on different aspects and challenges of analysing reliability in systems where component failures affect the system. First, we provide the preliminaries and some basic definitions regarding load-sharing, which are necessary to discuss the work presented throughout the thesis in Chapter 1. In the same chapter, we also discuss various inferential techniques used in the thesis. Three datasets - two pertaining to load-sharing systems and one about correlated bivariate data- are introduced in this chapter. These datasets are analysed in the latter chapters. A detailed literature review is performed to give a perspective of the load-sharing literature and the relevance of the work done in this thesis.

In Chapter 2, we develop a Bayesian estimation approach for the GFB model in a two-component load-sharing system. Assuming independent gamma priors for the parameters, we implement the estimation through the MCMC algorithm using the `rstan` package to `R` software. Simulation results show that the BEs successfully recover the true parameters, even in small-sample settings. A model selection study is performed to check the fit of the proposed model with respect to WAIC. As GFB model is a family of distributions, the choice of the best fitted model within GFB family of distributions for a given dataset is of interest. We discuss the model selection using WAIC. The simulation study indicates that the performance of WAIC in selecting the parent model is quite satisfactory. Two real datasets are analysed using this method.

In Chapter 3, we propose a generalized class of load-sharing models for two-component systems. Here, we combine the GFB distribution for component lifetimes with the GG distribution for the frailty component, which allows us to model both changes in lifetime behavior after the first failure and dependence between component lifetimes. This leads to a wider class of models that can accommodate a variety of practical situations. An EM algorithm is developed for parameter estimation, and simulation studies confirm that the proposed method performs well. We advocate use of the standard model selection criteria like AIC, AICc, BIC, BC to select the best fitted model within GFB-GG family of distributions. The simulation study shows that the performance of these criteria is quite satisfactory. Finally, we analyse a simulated load-sharing data and a real bivariate data. We also discuss an estimation procedure of some important reliability characteristics such as MTTF, MRT, and RMT under the GFB-GG model.

In Chapter 4, we introduce a semi-parametric approach using a PLA to model the CDF in multi-component load-sharing systems. This model is especially useful because it is data-driven and avoids strong parametric assumptions, making it suitable for a wide variety of real-world applications. Important reliability measures such as the quantile function, MTTF, MRT, and RMT are derived under the proposed model. Likelihood-based inference procedures are discussed in detail, including MLE and bootstrap CIs. A Monte Carlo simulation study is carried out to assess both the quality of the estimation procedure and the overall fit of the PLA-based model. It is found that the performance of the estimation method discussed is satisfactory. Through simulation, we show that the PLA-based model is quite flexible in fitting to datasets arising from different underlying processes. The model is applied to two-motor and three-basketball players' load-sharing datasets, and the analysis demonstrates that the PLA-based model is better in comparison to several existing models in terms of fit and flexibility for both datasets.

The main contributions of the thesis are listed below.

- We develop a Bayesian estimation framework for the GFB model in load-sharing systems using MCMC, and applied it successfully to real data.
- We provide a generalized frailty-based model (GFB-GG) that captures lifetime dependence and provides better flexibility in applications.
- We give a PLA-based semi-parametric model for analysing multi-component systems without strong distributional assumptions.
- We validate all methods through simulation studies.
- We compare our model with related existing models, and it is shown that our models perform better for the same data.

Overall, this thesis offers a flexible set of statistical models for modeling, analysing, and interpreting load-sharing systems. By considering Bayesian and classical (parametric and semi-parametric) approaches, the work contributes to both the theoretical development and practical applications of reliability modeling in complex systems.

5.2 Future Scope

Here, we discuss some open problems that are related to the work done in this thesis. We have a plan to work on these problems in the near future.

- In Chapter 2, the prior means were shifted from the true parameter values to assess the robustness of the posterior estimates under mild prior misspecification. We have a plan to examine the case where the prior means are centered at the true parameter values for Models B1-B9 to evaluate the impact of a well-specified prior and to include sensitivity analysis over multiple prior families and hyperparameter settings in future work.
- In Chapter 3, we use likelihood inference, and due to the small sample size, we face some problems in the optimization of the complicated GFB-GG models. To overcome this problem, we have a plan to implement a Bayesian framework in the GFB-GG model.
- In Chapter 4, we approximate CHF by a piecewise linear function in time. A natural generalization would be to consider a quadratic function or B-splines to approximate CHF.
- In all of our works, we do not consider censoring for load-sharing systems. We have a plan to study a load-sharing set-up when the observations are censored. It may be noted that incorporating censoring in this setting is a challenging problem; in such cases, the amount of information available for the components that fail later in the study may be substantially less. This incorporates a severe loss of information, or even no information at all, for these components. A suitable mechanism needs to be adopted to resolve this issue.

BIBLIOGRAPHY

[1] S. V. Amari and R. Bergman. Reliability analysis of k-out-of-n load-sharing systems. In *2008 Annual reliability and maintainability symposium*, pages 440–445. IEEE, 2008.

[2] S. V. Amari, K. B. Misra, and H. Pham. Tampered failure rate load-sharing systems: Status and perspectives. In *Handbook of Performability Engineering*, pages 291–308. Springer, 2008.

[3] S. V. Amari and H. Pham. Reliability analysis of dynamic fiber bundle models. In *2009 Annual Reliability and Maintainability Symposium*, pages 1–6. IEEE, 2009.

[4] G. Asha, J. K. M. Krishna, and D. Kundu. An extension of the Freund's bivariate distribution to model load-sharing systems. *American Journal of Mathematical and Management Sciences*, 35(3):207–226, 2016.

[5] G. Asha, A. V. Raja, and N. Ravishanker. Reliability modelling incorporating load share and frailty. *Applied Stochastic Models in Business and Industry*, 34:206–223, 2018.

[6] N. Balakrishnan, E. Beutner, and U. Kamps. Modeling parameters of a load-sharing system through link functions in sequential order statistics models and associated inference. *IEEE Transactions on Reliability*, 60(3):605–611, 2011.

[7] N. Balakrishnan, M. V. Koutras, F. S. Milienos, and S. Pal. Piecewise linear approximations for cure rate models and associated inferential issues. *Methodology and Computing in Applied Probability*, 18:937–966, 2006.

[8] N. Balakrishnan and Y. Peng. Generalized gamma frailty model. *Statistics in Medicine*, 25:2797–2816, 2006.

[9] Z. Birnbaum and S. C. Saunders. A statistical model for life-length of materials. *Journal of the American Statistical Association*, 53(281):151–160, 1958.

[10] M. Burkschat, U. Kamps, and M. Kateri. Sequential order statistics with an order statistics prior. *Journal of Multivariate Analysis*, 101(8):1826–1836, 2010.

[11] G. Claeskens. Statistical model choice. *Annual Review of Statistics and Its Application*, 3:233–256, 2016.

[12] B. D. Coleman. Time dependence of mechanical breakdown in bundles of fibers. I. constant total load. *Journal of Applied Physics*, 28(9):1058–1064, 1957.

[13] B. D. Coleman. Statistics and time dependence of mechanical breakdown in fibers. *Journal of Applied Physics*, 29(6):968–983, 1958.

[14] D. Cox. Regress models and life-tables. *Journal of the Royal Statistical Society Series B*, 34:187–220, 1972.

[15] E. Cramer and U. Kamps. Sequential order statistics and k-out-of-n systems with sequentially adjusted failure rates. *Annals of the Institute of Statistical Mathematics*, 48(3):535–549, 1996.

[16] E. Cramer and U. Kamps. Marginal distributions of sequential and generalized order statistics. *Metrika*, 58:293–310, 2003.

[17] H. E. Daniels. The statistical theory of the strength of bundles of threads I. *Proceedings of the Royal Society of London*, 83:405–435, 1945.

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

[19] J. V. Deshpande, I. Dewan, and U. V. Naik-Nimbalkar. Two components load sharing systems with applications to biology, 2007. `https://www.isid.ac.in/~statmath/2007/isid200706.pdf`.

[20] L. Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, New York, 1986.

[21] S. D. Durham, J. D. Lynch, and W. J. Padgett. A theoretical justification for an increasing average failure rate strength distribution in fibrous composites. *Naval Research Logistics*, 36(5):655–661, 1989.

[22] M. Franco, J.-M. Vivo, and D. Kundu. A generalized Freund bivariate model for a two-component load sharing system. *Reliability Engineering & System Safety*, 203:107096, 2020.

[23] J. Freund. A bivariate extension of the exponential distribution. *Journal of American Statistical Association*, 56:971–976, 1961.

[24] A. Ganguly, D. Mitra, N. Balakrishnan, and D. Kundu. A flexible model based on piecewise linear approximation for the analysis of left truncated right censored data with covariates, and applications to worcester heart attack study data and channing house data. *Statistics in Medicine*, 43(2):233–255, 2024.

[25] A. Gelman, J. Hwang, and A. Vehtari. Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6):997–1016, 2014.

[26] D. D. Hanagal. *Modeling survival data using frailty models, Second Edition.* Boca Raton: Chapman and Hall/CRC Boca Raton, 2011.

[27] D. G. Harlow and S. L. Phoenix. The chain-of-bundles probability model for the strength of fibrous materials I: analysis and conjectures. *Journal of Composite Materials*, 12(2):195–214, 1978.

[28] D. G. Harlow and S. L. Phoenix. Probability distributions for the strength of fibrous materials under local load sharing I: two-level failure and edge effects. *Advances in Applied Probability*, 14:68–94, 1982.

[29] P. Hougaard. *Analysis of Multivariate Survival Data.* New York: Springer, 2000.

[30] U. Kamps. A concept of generalized order statistics. *Journal of Statistical Planning and Inference*, 48(1):1–23, 1995.

[31] H. Kim and P. H. Kvam. Reliability estimation based on system data with an unknown load share rule. *Lifetime Data Analysis*, 10:83–94, 2004.

[32] P. H. Kvam and E. A. Peña. Estimating load-sharing properties in a dynamic reliability system, 2003. https://people.stat.sc.edu/pena/TechReports/KvamPena2003.pdf.

[33] P. H. Kvam and E. A. Peña. Estimating load-sharing properties in a dynamic reliability system. *Journal of the American Statistical Association*, 100(469):262–272, 2005.

[34] T. A. Louis. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):226–233, 1982.

[35] A. Maurya, J. Vaghasiya, and R. K. Maurya. Analyzing k-out-of-n load sharing systems under progressive censoring. *American Journal of Mathematical and Management Sciences*, 44(1):69–91, 2025.

[36] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. Wiley, New York, 2nd edition, 2008.

[37] A. Mettas and P. Vassiliou. Application of quantitative accelerated life models on load sharing redundancy. In *Annual Symposium Reliability and Maintainability, 2004-RAMS*, pages 293–296. IEEE, 2004.

[38] F. Mies and S. Bedbur. Exact semiparametric inference and model selection for load-sharing systems. *IEEE Transactions on Reliability*, 69(3):863–872, 2019.

[39] M. Miyakawa. Analysis of incomplete data in competing risks model. *IEEE Transactions on Reliability*, 33:293–296, 1984.

[40] C. Park. Parameter estimation for the reliability of load-sharing systems. *IIE Transactions*, 42:753–765, 2010.

[41] C. Park. Parameter estimation from load-sharing system data using the expectation-maximization algorithm. *IIE Transactions*, 45:147–163, 2013.

[42] C. Park and W. J. Padgett. New cumulative damage models for failure using stochastic processes as initial damage. *IEEE Transactions on Reliability*, 54(3):530–540, 2005.

[43] C. Park, M. Wang, R. M. Alotaibi, and H. Rezk. Load-sharing model under Lindley distribution and its parameter estimation using the expectation-maximization algorithm. *Entropy*, 22(11):1329, 2020.

[44] M. Park and J. Y. Kim. Cost analysis on warranty policies using Freund's bivariate exponential distribution. *Journal of Korean Society for Quality Management*, 42:1–14, 2014.

[45] T. Pesch, E. Cramer, A. Polpo, and E. Cripps. Estimation with extended sequential order statistics: A link function approach. *Applied Stochastic Models in Business and Industry*, 40(5):1313–1336, 2024.

[46] T. M. Pesch, E. Cramer, E. Cripps, and A. Polpo. Modeling failure risks in load-sharing systems with heterogeneous components. *IEEE Transactions on Reliability*, 73(3):1446–1458, 2024.

[47] S. Phoenix. A classical and Bayesian estimation of a k-components load-sharing parallel system. *SIAM Journal of Applied Mathematics*, 34:227–246, 1978.

[48] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2025.

[49] B. W. Rosen. Tensile failure of fibrous composites. *AIAA Journal*, 2(11):1985–1991, 1964.

[50] S. M. Ross. A model in which component failure rates depend on the working set. *Naval Research Logistics Quarterly*, 31(2):297–300, 1984.

[51] Z. Schechner. A load-sharing model: The linear breakdown rule. *Naval Research Logistics Quarterly*, 31(1):137–144, 1984.

[52] B. Singh and P. Gupta. Load-sharing system model and its application to the real data set. *Mathematics and Computers in Simulation*, 82:1615–1629, 2012.

[53] B. Singh and P. Gupta. Bayesian reliability estimation of a 1-out-of-k load-sharing system model. *International Journal of System Assurance Engineering and Management*, 5:562–576, 2014.

[54] B. Singh, K. K. Sharma, and A. Kumar. A classical and Bayesian estimation of a k-components load-sharing parallel system. *Computational Statistics & Data Analysis*, 52(12):5175–5185, 2008.

[55] Stan Development Team. *RStan: the R interface to Stan*. R Package Version 2.32.7., 2025.

[56] S. S. Sutar and U. V. Naik-Nimbalkar. Accelerated failure time models for load sharing systems. *IEEE Transactions on Reliability*, 63:706–714, 2014.

[57] S. S. Sutar, S. V. Rajguru, P. S. Patil, and S. D. Pawar. Reliability analysis of load sharing systems under unequal load-sharing rule with applications. *Communications in Statistics-Theory and Methods*, 54(13):3938–3960, 2025.

[58] D. Wang, C. Jiang, and C. Park. Reliability analysis of load-sharing systems with memory. *Lifetime Data Analysis*, 25:341–360, 2019.

[59] X. Zhao, B. Liu, and Y. Liu. Reliability modeling and analysis of load-sharing systems with continuously degrading components. *IEEE Transactions on Reliability*, 67:1096–1110, 2018.

BIODATA

Current Status

Research Scholar (July, 2019–Present), Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati 781039, India.

Education

- 2016–2018: **Master of Science** (Mathematics), CGPA: **8.04**
  Tezpur (Central) University,Tezpur, Assam India.

- 2013–2016: **Bachelor of Science** (Honours in Mathematics), percentage: **67.13**
  A.B.N. Seal College, University of North Bengal, West Bengal, India.

- 2011–2013: **Higher Secondary (12th)** (Science Stream), percentage: **85.60**
  Dinhata Soni Debi Jain High School, Dinhata, Coochbehar, West Bengal, India.

- 2011: **Secondary (10th)**, percentage: **85.38**
  Station Para Saranarthi High School, Dinhata, Coochbehar, West Bengal, India.

Teaching Experience

Teaching Assistant at IIT Guwahati

- MA 225 (Probabilty Theory and Random Processes) — July-Nov, 2020
- MA 102 (Linear Algebra & Ordinary Differential Equations) — Jan-May, 2021
- MA 323 (Monte Carlo Simulation) — July-Nov, 2021
- MA 102 (Linear Algebra & Ordinary Differential Equations) — Jan-May, 2022

- NPTEL (Discrete-time Markov Chains and Poisson Processes) Jan-April, 2022
- MA 323 (Monte Carlo Simulation) July-Nov, 2022
- MA 101 (Single Variable Calculus & Multivariable Calculus) Jan-May, 2023

Technical Skills

- Software: LaTeX.
- Languages: `R`, `rstan`, C++.
- Operating Systems: Linux and Windows.

Ph.D. Course Work

- Real Analysis
- Differential Equations
- Linear Algebra
- Statistical Foundation for Data Science
- Probability Theory
- Generalized Linear Models

Awards & Achievements

- Awarded West Bengal Merit-cum-means scholarship (2011-2013).
- Awarded DST-INSPIRE (SHE) scholarship during 2013-2018.
- Awarded DST-INSPIRE fellowship for pursuing PhD at Tezpur University (2018-2019).
- Qualified Graduate Aptitude Test in Engineering (GATE) 2019 in Mathematics.
- Qualified National Eligibility Test (NET) June 2017 in Mathematics (Conducted by Council of Scientific & Industrial Research, India).
- Qualified National Eligibility Test (NET) December 2017 in Mathematics (Conducted by Council of Scientific & Industrial Research, India).
- Awarded Ministry of Human Resource Development scholarship for pursuing PhD at IIT Guwahati (July 2019 to present).
- Awarded the 2022 CWS Prof. Suddhendu Biswas Student Travel Award.

Conference and Workshop Attended during PhD Tenure

- Annual Conference of the International Indian Statistical Association (IISA-2024), December 27 - 31, 2024, organized by Cochin University of Science and Technology, Kochi, Kerala.
- International Workshop Reliability Theory and Survival analysis (IWRTSA-2024), December 21 - 23, 2024, organized by Department of Statistics, I.Sc., Banaras Hindu University
- Research and Industrial Conclave (RIC-2024), August 9 - 11, 2024, organized by the Students' Academic Board, IIT Guwahati.
- National conference on Modeling, Analysis & Simulation (Comprising the realms of AI, ML and IoT) (MAS-2024), June 28 - 30, 2024, organized by Department of Mathematics & Computing, IIT (ISM), Dhanbad.
- 38-th Annual Conference of the Ramanujan Mathematical Society (RMS-2023), December 22 - 24, 2023, organized by Department of Mathematics, IIT Guwahati.
- Third International Conference on Applications of Mathematics to Nonlinear Sciences (AMNS-2023), May 25 - 28, 2023, organized by ANMA and NMS, Pokhara, Nepal.
- Annual Conference of the International Indian Statistical Association (IISA-2022), December 26 - 30, 2022, organized by Indian Institute of Science, Bengaluru.
- International Workshop Reliability Theory and Survival analysis (IWRTSA-2022), December 23 - 25, 2022, organized by Department of Mathematics and Statistics, Indian Institute of Science Education and Research Kolkata.
- Workshop on Applied Statistical Modeling & Data Analytics for Petroleum Engineering & Related Applications, November 22 - 24, 2022, organized by Department of Chemical Engineering IIT Guwahati.
- International Symposium on Probability and Statistics: New Frontiers (ISPS-2022) (online mode), August 12 - 14, 2022, organized by Department of Statistics, University of Calcutta & Calcutta Statistical Association.

List of Published and Communicated Papers

- **Shilpi Biswas**, Ayon Ganguly and Debanjan Mitra, **Reliability Analysis of Load-sharing Systems using a Flexible Model with Piecewise Linear Functions**, *Applied Stochastic Models in Business and Industry*. `https://doi.org/10.1002/asmb.2934`.
- **Shilpi Biswas**, Ayon Ganguly and Debanjan Mitra, **A Doubly-Flexible Model Based on Generalized Gamma Frailty for Two-component Load-sharing Systems** (Communicated to a journal).
- **Shilpi Biswas**, Ayon Ganguly and Debanjan Mitra, **Bayesian Analysis of Generalized Freund Bivariate Model for a Two-component Load-sharing System** (Communicated to a journal).