وزارة التعليم العالي والبحث العلمي

Ministry of Higher Education and Scientific Research

BADJI MOKHTAR-ANNABA UNIVERSITY

UNIVERSITE BADJI MOKHTAR ANNABA

جامعة باجي مختار ـ عنابة ـ

Faculty of Sciences
Department of Mathematics

Year: 2025/2026

THESIS

Presented with a view to obtaining the doctorate degree

About Frailty Models

Stream
Applied Mathematics

Speciality
Probability and Statistics

By
TEGHRI Samia

SUPERVISOR: GOUAL Hafida **M.C.A.** **U.B.M. Annaba**

In front of the jury

PRESIDENT: ZEGHDOUDI Halim	**Prof.**	**U.B.M. Annaba**
EXAMINER: LEULMI Sarra	**M.C.A.**	**U. Constantine 1**
EXAMINER: MEZHOUD Kenza Assia	**M.C.A.**	**U. Constantine 1**
EXAMINER: TALHI Hamida	**M.C.A.**	**U.B.M. Annaba**

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

BADJI MOKHTAR-ANNABA

UNIVERSITY

UNIVERSITE BADJI MOKHTAR
ANNABA

جامعة باجي مختار

ـعنابةـ

Faculté des Sciences

Département de Mathématiques

Année : 2025/2026

THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat

Sur Les Modèles De Fragilité

Filière
Mathématiques Appliquées

Spécialité
Probabilités et Statistique

Par
TEGHRI Samia

DIRECTEUR DE THÈSE: GOUAL Hafida **M.C.A. U.B.M. Annaba**

Devant le jury

PRESIDENT: ZEGHDOUDI Halim **Prof. U.B.M. Annaba**

EXAMINATEUR : LEULMI Sarra **M.C.A. U. Constantine 1**

EXAMINATEUR : MEZHOUD Kenza Assia **M.C.A. U.Constantine 1**

EXAMINATEUR: TALHI Hamida **M.C.A. U.B.M. Annaba**

حول نماذج الهشاشة

ملخّص

عادة ما تفترض نماذج البقاء أن الأفراد في المجتمع متجانسون فيما يتعلق بقابليتهم لحدوث الحدث، مثل الوفاة أو الفشل. ومع ذلك، في البيانات الحقيقية، يوجد عادةً تباين غير مرئي، حيث تؤثر العوامل الكامنة مثل العوامل الوراثية أو البيئية أو الاجتماعية على الأفراد. إذا تم تجاهل هذا التباين، فقد يؤدي ذلك إلى تقديرات منحازة لمعدلات البقاء ودوال الخطر. لمعالجة هذه المشكلة، نقترح نموذجًا جديدًا للضعف (Frailty) الذي يدمج التباين غير المرئي من خلال متغير الضعف الذي يتبع توزيع ليندلي ذو المعاملين (TPL). يتم تقدير النموذج باستخدام طريقة الاحتمالية القصوى، ونقوم بدراسة أدائه مع العديد من دوال الخطر الأساسية المستخدمة بشكل شائع، بما في ذلك توزيعات ويبول، الأسي، جومبيرتس، وبارتو. من خلال دراسات محاكاة موسعة، نقيم قدرة النموذج على التقاط التباين ونقارن أدائه مع نماذج الضعف الأخرى المستخدمة بشكل واسع. كما نستخدم اختبارات جودة الملاءمة من نوع نيكولين-راو-روبسون ونيكولين-باغدونايفيوس-نيكولين للتحقق من دقة النموذج. لإثبات قابليتها للتطبيق العملي، نقوم بتحليل مجموعة بيانات طبية حقيقية من مستشفى طوارئ في الجزائر، إلى جانب بيانات عن حالات النوبات القلبية، حيث يتفوق نموذج الضعف المعتمد على توزيع على النماذج التقليدية في التقاط التباين غير المرئي وتقديم تقديرات أكثر موثوقية للبقاء.

كلمات مفتاحية: نماذج الهشاشة، اختبارات مطابقة الجودة، دالة الخطر، تحويل لابلاس

About frailty models

Abstract

Survival analysis often assumes that individuals within a population are homogeneous with regard to their susceptibility to an event, such as death or failure. However, real-world data typically exhibit unobserved heterogeneity, where individuals are affected by latent factors such as genetic, environmental, or social influences. If this variability is ignored, it can lead to biased estimates of survival rates and hazard functions. To address this issue, we propose a novel frailty model that incorporates unobserved heterogeneity through a frailty variable following the Two-Parameter Lindley (TPL) distribution. The model is estimated using maximum likelihood estimation, and we examine its performance with several common baseline hazard functions, including Weibull, Exponential, Gompertz, and Pareto distributions. Through extensive simulation studies, we assess the model's ability to capture heterogeneity and compare it with other widely used frailty models. We also employ Nikulin-Rao-Robson and Bagdonavicius-Nikulin goodness-of-fit tests to validate the model's accuracy. To demonstrate its practical applicability, we analyze a real medical dataset from an emergency hospital in Algeria, along with heart attack data, where the proposed frailty model outperforms traditional models in capturing unobserved heterogeneity and providing more reliable survival predictions.

Keywords: Frailty models; Goodness-of-fit testing; Hazard function; Laplace transform

Sur les modèles de fragilité

Résumé

L'analyse de survie repose souvent sur l'hypothèse d'homogénéité au sein d'une population d'individus en ce qui concerne leur susceptibilité à un événement, tel que le décès ou la défaillance. Cependant, les données réelles présentent généralement une hétérogénéité non observée, où des facteurs latents tels que les influences génétiques, environnementales ou sociales affectent les individus. Si cette variabilité n'est pas prise en compte, cela peut entraîner des estimations biaisées des taux de survie et des fonctions de risque. Afin de remédier à cette problématique, nous proposons un nouveau modèle de frailty qui intègre l'hétérogénéité non observée par le biais d'une variable de frailty suivant la distribution de Lindley à deux paramètres (TPL). Le modèle est estimé à l'aide de la méthode du maximum de vraisemblance et nous examinons ses performances avec plusieurs fonctions de risque de base couramment utilisées, notamment les distributions Weibull, Exponentielle, Gompertz et Pareto. À travers des études de simulation approfondies, nous évaluons la capacité du modèle à capturer l'hétérogénéité et le comparons à d'autres modèles de frailty largement utilisés. Nous utilisons également les tests de bonté d'ajustement de Nikulin-Rao-Robson et de Bagdonavicius-Nikulin pour évaluer la précision du modèle. Pour démontrer son applicabilité pratique, nous analysons un ensemble de données médicales réelles provenant d'un hôpital d'urgence en Algérie, ainsi que des données relatives aux crises cardiaques, où le modèle de frailty proposé surpasse les modèles traditionnels en capturant l'hétérogénéité non observée et en fournissant des prédictions de survie plus fiables.

Mots-clés : Modèles de fragilité ; Tests d'adéquation ; Fonction de risque ; Transformée de Laplace

To Allah.

The Glorified and Exalted,
The only One worshipped,
The greatest Cherisher and Sustainer.

For his satisfaction, I have read and written.

He is the Best Disposer of affairs.
And, Him alone is sufficient for us.
May Him accept us as
his righteous slaves and submitters,
and pious worshippers.

All the praises to Allah.
My Almighty, Allah.
...

Aknowledgements

I would like to express my sincere gratitude to my supervisor **Dr. Goual Hafida** for their guidance and support during the preparation of this doctoral thesis.

I would like to thank my family, colleagues, and friends for their support throughout my doctoral studies.

I also thank the members of the jury for accepting to evaluate this work. My thanks go to **Pr. ZEGHDOUDI Halim**, **Dr. LEULMI Sarra**, **Dr. MEZHOUD Kenza Assia** and **Dr. TALHI Hamida** for their time and consideration.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Survival analysis is a core and essential area within mathematical statistics devoted to the study and analysis of time-to-event data, where the objective is to model and understand the duration between an initiating event, such as the diagnosis of a disease or the start of observation, and the occurrence of a defined endpoint, including death, disease relapse, equipment failure, or other significant events. This framework was formally developed by Cox (1972) [34] and further systematized by Kalbfleisch and Prentice (2002) [91]. Its applicability spans a wide range of disciplines, including medicine, epidemiology, public health, biology, reliability engineering, actuarial science, demography, economics, engineering, and the social sciences (Aalen and Tretli, 1999 [3]; Nelson, 1982 [130]; Lancaster, 1990 [103]; Kleinbaum and Klein, 2012 [97]; Collett, 2015) [31]. Owing to this broad relevance, survival analysis has become an indispensable methodological tool for analyzing lifetime and duration data across diverse applied contexts.

A distinctive and defining feature of survival data is the presence of censoring, a concept rigorously studied by Turnbull (1976) [157] and later summarized by Lawless (2003) [104]. Right-censoring, left-censoring, and interval-censoring frequently arise in practice due to incomplete follow-up, delayed entry, or study termination before the occurrence of the event of interest. In addition, truncation mechanisms and the inclusion of time-dependent covariates further complicate statistical inference, as discussed by Andersen and Keiding (2012) [13]. These inherent characteristics of survival data necessitate specialized statistical methodologies that extend beyond the scope of classical regression models.

The theoretical foundations of survival analysis rely on several key functions that describe lifetime behavior. The survival function and hazard function were formally de-

fined and popularized by Cox (1972) [34], while the cumulative hazard function was later developed within the counting process framework by Aalen (1978) [1]. Together, these functions provide complementary perspectives on the risk structure over time and form the basis of most parametric, semi-parametric, and non-parametric survival models.

A wide variety of classical parametric survival models have been extensively studied in the literature. The exponential distribution, introduced in reliability theory by Feller (1971) [47], assumes a constant hazard rate over time. The Weibull distribution, proposed by Weibull (1951) [164], allows for monotone increasing or decreasing hazard functions and remains one of the most widely used lifetime models. The Gompertz distribution was originally introduced by Gompertz (1825) [54] in demographic mortality studies, while the log-normal distribution was applied to survival data by Aitchison and Brown (1957) [8]. The log-logistic model was later examined by Bennett (1983) [23] for medical and biological applications. Despite their practical usefulness, these classical models typically assume homogeneous risk among individuals with identical observed covariates.

However, empirical evidence has repeatedly shown that the assumption of homogeneous risk is often unrealistic. A classical and often implicit assumption in standard survival models is that observed lifetimes are independently and identically distributed (IID). In real-world applications, this assumption is frequently violated due to the presence of unobserved heterogeneity among individuals. Vaupel, Manton, and Stallard (1979) [160] demonstrated that ignoring such heterogeneity may lead to biased inference and misleading conclusions. This phenomenon was further examined by Heckman and Singer (1984) [69], who showed that unobserved heterogeneity can induce spurious duration dependence and distort estimated covariate effects.

To address this fundamental limitation, frailty models were introduced into survival analysis. The concept of frailty was formally proposed by Vaupel et al. (1979) [160] as a latent, unobserved random variable that acts multiplicatively on the hazard function and captures individual-level susceptibility to the event of interest. Clayton (1978) [28] independently introduced frailty ideas in multivariate survival analysis, emphasizing the dependence induced among correlated survival times. Subsequent work by Struthers and

Kalbfleisch (1986) [154] and Henderson and Oman (1999) [71]further highlighted that ignoring frailty can result in biased regression parameter estimates and misinterpretation of covariate effects.

Frailty models were developed precisely to incorporate unobserved heterogeneity by introducing random effects at the individual or group level. The fundamental idea underlying frailty modeling is that observed survival times are governed by an unobserved latent variable that modifies the individual hazard function (Vaupel et al., 1979) [160]. As a result, frailty models provide a flexible and powerful framework for modeling heterogeneity in survival data and have found widespread applications in cancer research, epidemiology, clinical trials, reliability engineering, and social science studies.

Frailty models are commonly classified into shared frailty and univariate frailty formulations. In shared frailty models, individuals belonging to the same cluster such as members of the same family, patients treated within the same hospital unit, or components of a common mechanical system share a common frailty term due to shared genetic, environmental, or contextual factors (Aalen, 1988 [3]; Aalen and Aalen, 1992) [4]. In contrast, univariate frailty models assume that each individual possesses their own independent frailty term. When independence across individuals is plausible, univariate frailty structures lead to analytically tractable models that are widely used in biomedical and clinical research.

A variety of probability distributions have been proposed to model frailty. The gamma frailty model, introduced by Vaupel et al. (1979) [160] and later formalized by Hougaard (1986) [77], remains the most commonly used due to its mathematical tractability. The inverse Gaussian frailty model was proposed by Hougaard (1984) [76] to allow for heavier tails and increased heterogeneity. The log-normal frailty model was studied by McGilchrist (1993) [122], offering greater flexibility at the expense of computational complexity. The positive stable frailty model, also introduced by Hougaard (1986) [77], allows for strong dependence structures. Additional developments include the Power Variance Function (PVF) frailty model (Hougaard, 1986 [77]; Duchateau and Janssen, 2007) [42], the compound Poisson frailty model (Aalen, 1988; Aalen and Aalen,

1992) [3] [4], Lévy-based frailty models (Hougaard, 2000) [83], and log-gamma frailty models discussed by Kalbfleisch and Prentice (2002) [91].

Although these frailty distributions have proven useful, each exhibits limitations in terms of flexibility, tail behavior, or the ability to capture complex forms of heterogeneity observed in empirical data (Bretagnolle and Huber-Carol, 1988) [25]. Consequently, recent research has increasingly focused on the development of more flexible lifetime and frailty distributions. In this context, the Lindley distribution, originally introduced by Lindley (1958) [112] in a Bayesian framework, has attracted renewed attention and was later adapted for survival analysis by Ghitany et al. (2008) [52]. Several extensions of the Lindley distribution have since been proposed, including the two-parameter Lindley (TPL) distribution introduced by Shanker and Mishra (2013) [151], as well as further generalizations studied by Zakerzadeh and Dolati (2009) [173], Bakouch et al. (2012) [19], and Al-Zahrani (2015) [10].

Motivated by the desirable properties of the TPL distribution, including positive support, flexible skewness, and analytical tractability, a new two-parameter frailty (TPF) model was recently proposed as an alternative to classical gamma, compound Poisson, log-normal, and weighted Lindley frailty models (Mota et al., 2021) [126]. Building on these developments, the present thesis introduces and studies a novel extension, namely the Two-Parameter Lindley Frailty (TPLF) model. The TPLF model incorporates the two-parameter Lindley distribution as a frailty component, preserving its favourable properties while providing a richer and more flexible structure for modelling unobserved heterogeneity in survival data.

Assessing the adequacy of survival and frailty models is a crucial aspect of statistical modelling. Goodness-of-fit procedures based on cumulative hazard functions were proposed by Bagdonavičius and Nikulin (1997) [15], with further asymptotic developments in (2001) [16]. For complete data, adjusted chi-squared goodness-of-fit tests introduced by Nikulin (1973a, 1973b, 1973c) [132] [133] [134] and the Nikulin–Rao–Robson test (Rao and Robson, 1974) [140] are employed. For censored data, the Bagdonavičius–Nikulin (Bg-N) test has proved particularly effective in validating parametric survival and frailty

models (Bagdonavičius and Nikulin, 2011) [17]. These tests enable rigorous comparison between empirical observations and theoretical distributions.

Likelihood-based inference for frailty models has been extensively discussed by Therneau and Grambsch (2000) [156] and is adopted in this thesis for the proposed TPLF model. The statistical properties of the estimators are investigated through extensive simulation studies under varying levels of censoring, following the approaches of Ripatti and Palmgren (2000) [141] and Duchateau and Janssen (2007 [42]). These Monte Carlo experiments allow for the assessment of bias, efficiency, robustness, and overall performance of the proposed model.

The methodological developments are complemented by empirical applications based on medical data. The first application relies on newly collected data from an emergency hospital in Algeria. Using this dataset, we analyze time-to-event outcomes for patients with a specific clinical condition and compare several baseline hazard functions, including the Weibull, Gompertz, exponential, Pareto, and related families. The empirical results demonstrate that the proposed TPLF model provides an improved fit compared to conventional frailty models and offers enhanced interpretability of survival dynamics in the considered medical setting.

In addition, a second application is conducted using a heart attack dataset developed by the Hungarian Institute of Cardiology. This application follows the same modeling strategy and inferential procedure as the emergency care data analysis, including parameter estimation, baseline hazard comparison, and goodness-of-fit assessment. The purpose of this second application is to further validate the robustness, generalizability, and practical applicability of the proposed TPLF model across different medical contexts.

The primary objective of this thesis is therefore to develop, analyze, and apply a flexible frailty model capable of capturing complex unobserved heterogeneity in survival data. The thesis combines theoretical development, likelihood-based inference, simulation-based validation, goodness-of-fit assessment, and real-data applications to demonstrate the robustness and practical utility of the proposed Two-Parameter Lindley Frailty model.

The remainder of the thesis is organized as follows. Chapter 2 introduces the fundamentals of survival analysis. Chapter 3 reviews classical and modern frailty models. Chapter 4 presents Lindley and Two-Parameter Lindley distributions. Chapter 5 focuses on the estimation, simulation studies, goodness-of-fit testing, and applications of the new Two Parameter frailty model. Chapter 6 concludes the thesis and outlines directions for future research.

Chapter 2

Backround

Survival analysis is a statistical technique that addresses time-to-event data (Wienke (2010)) [165]. Survival data is identified as such because, historically, the occurrences primarily concerned were mortality (Hougaard, 1986) [77]. This term is now applicable to all types of time-to-event data. The occurrence can generally be perceived as a transition from one state to another (Kleinbaum and Klein, 1996).

The primary notion in survival analysis is modeling the time until an event occurs, sometimes referred to as the failure time. Survival time is defined as a period measured from a designated time origin until the occurrence of the event of interest (Hanagal, 2011) [65]. To determine the survival time, it is crucial to establish the time origin (beginning point), agree upon a scale for measuring the passage of time, and clearly define the occurrence (often referred to as failure) (Kleinbaum and Klein, 1996) [96].

Challenges within the field of survival analysis arise when specific subjects have undergone the event of interest, whereas others have not by the conclusion of the study, leading to uncertainty regarding the actual survival durations and introducing the concept of censoring.

Censoring happens when there is partial information regarding individual survival time, although the exact survival duration remains unknown (Selvin, 2004) [149]. Kleinbaum and Klein (1996) [96] delineate three classifications of censoring; the initial type is referred to as right censoring, which occurs when the event of interest transpires following the recorded survival duration. Consequently, right-censored survival time is inferior to the actual survival time. Secondly, left censoring occurs when we detect the existence of a condition but lack knowledge at the point of its initiation; in this particular instance, the true duration of survival is less than the recorded censoring interval. Interval censoring

is present in situations where it is known that an individual experienced an event during a specific time interval, but the precise survival time remains unknown.

A crucial assumption for the methods discussed below for analyzing censored survival data is that individuals who are censored possess the same probability of subsequent failure as those who remain alive and are not censored. This signifies that an individual whose survival duration is censored at the temporal point C must be indicative of all other individuals who have persisted until that moment. Should this be the case, the process of censoring can be characterized as non-informative (Wienke, 2010) [165].From a statistical perspective, in cases where the censoring mechanism demonstrates independence from the duration of survival, then

$$P(X \geq x, C \geq x) = P(X \geq x)P(C \geq x) \tag{2.1}$$

where X is the actual time until the event occurs. Consequently, independent censoring is a specific form of non-informative censoring (Wienke, 2010) [165].

2.1 Essential Theories in Survival Analysis

2.1.1 Survival and Hazard Functions

Let us examine a cohort of subjects characterized by survival durations denoted as $t_1, t_2, \ldots, t_N$, a subset of which may be influenced by censoring. The previous values can be considered as those of the continuous variable T, characterized by the probability density function $f(t)$ and the cumulative distribution function $F(t)$ (Wienke, 2010) [165], where $F(t)$ is defined as

$$F(t) = P(T < t) = \int_0^t f(u)du \tag{2.2}$$

This indicates the probability that the period of survival is less than a specific value t (Collett (1994)) [30]. The survivor function, denoting the probability that an individual

would survive beyond time t, is expressed as

$$S(t) = P(T \geq t) = 1 - F(t) \tag{2.3}$$

Due to the skewness present in survival distributions, along with the presence of censored data, the mean and variance are considered inadequate for representing the distribution of T; instead, the median and quantiles are more suitable for this application. These statistical measures can be extracted from the survival function (Kleinbaum and Klein, 1996) [96]. The median survival time is defined as the value t_m of T at which $S(t_m) = 0.5$.

For a continuous random variable, T (Wienke (2010)) [165], the probability density function (pdf) is defined as

$$f(t) = F'(t) = -S'(t), t \geq 0 \tag{2.4}$$

The hazard function expresses the immediate failure rate at time t, based on the assumption that the individual has successfully survived up to that particular temporal point, is defined as

$$h(t) = \lim_{\delta t \to 0} \frac{P(t \leq T < t + \delta t \mid T \geq t)}{\delta t}; t \geq 0 \tag{2.5}$$

Using the notions of conditional probability and the mathematical concept of derivatives, the previous equation can be reformulated as follows.

$$\begin{aligned} h(t) &= \lim_{\delta t \to 0} \frac{P(t \leq T \leq t + \delta t)}{\delta t P(T \geq t)} \\ &= \lim_{\delta t \to 0} \left[\frac{F(t + \delta t) - F(t)}{\delta t} \right] \frac{1}{P(T \geq t)} \\ &= \frac{f(t)}{S(t)} \end{aligned}$$

The association between $S(t)$ and $h(t)$ is shown below.

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d\log S(t)}{dt} \tag{2.6}$$

Similarly,

$$S(t) = \exp\left[-\int_0^t h(u)du\right] = \exp(-H(t)), t \geq 0 \tag{2.7}$$

$H(t) = \int_0^t h(u)\,du$ is referred to as the cumulative hazard function, obtained from the survival function, as $H(t) = -\log S(t)$, where the probability density function can be expressed as

$$f(t) = h(t)\exp\left[-\int_0^t h(u)du\right], t \geq 0 \tag{2.8}$$

All of these functions offer a mathematically analogous characterization of the distributions pertaining to the survival time variable T. If one of these functions were to be established, the remaining two could be deduced. The survival function is particularly advantageous for contrasting the survival progress of multiple groups. The hazard function provides a more effective characterization of the risk associated with failure at a specific temporal point. It signifies the immediate probability of experiencing failure at a specified moment, dependent on the individual having survived until that moment.

Estimation methods can be categorized into parametric and nonparametric approaches. Alternative methods, including the semi-parametric methodology introduced by Cox (1972) [34], known as the Cox proportional hazards model, has also been successfully validated. This project will provide a concise explanation of these strategies.

The survival function is consistently a decreasing function, whereas the risk function is usually an increasing function.

2.1.2 Parametric Distributions

Survival data are typically right-skewed, making symmetric distributions like the Normal distribution ineffective for modelling such data (Kleinbaum and Klein, 1996) [96].

Asymmetric distributions commonly include the exponential, Weibull, and log-logistic distributions (Wienke (2010)) [165]. This section will exclusively address the exponential and Weibull models. The objective is to outline essential associations based on the proposition of specific survival distributions.

Exponential Distribution

The exponential distribution is characterized by the subsequent probability density function.

$$f(t;\lambda) = \lambda e^{-\lambda t}, \quad t > 0 \tag{2.9}$$

The cumulative distribution function is expressed as

$$F(t) = 1 - e^{-\lambda t} \tag{2.10}$$

and the survival function is

$$S(t) = 1 - F(t) = e^{-\lambda t} \tag{2.11}$$

The hazard function is represented by

$$h(t) = \frac{f(t)}{S(t)} = \frac{\lambda e^{-\lambda t}}{e^{\lambda t}} = \lambda(\text{ fixed value }) \tag{2.12}$$

The exponential distribution indicates a constant hazard, implying that the mortality risk is time-independent. This assumption is quite unrealistic, as intuitively, the risk of death may fluctuate with age, for instance.

Based on Kosorok (2008) [98]. A significant characteristic of the exponential distribution is its memoryless structure. Let the random variable T represent survival time, which follows an exponential distribution with parameter λ. Evaluate the probability that an individual survives beyond time t_1, conditional on having survived up to time t_0.

Subsequently

$$\begin{aligned} P(T > t_1 \mid T > t_0) &= \frac{P(T > t_1 \text{ and } T > t_0)}{P(T > t_0)} \\ &= \frac{P(T > t_1)}{P(T > t_0)} \\ &= \frac{S(t_1)}{S(t_0)} \\ &= \frac{e^{-\lambda t_1}}{e^{-\lambda t_0}} \\ &= e^{\lambda(t_1 - t_0)} \end{aligned}$$

This may be viewed as follows. Conditional on surviving till time t_0, the excess lifetime beyond t_0 continues to follow an exponential distribution with parameter λ. This outcome demonstrates why the exponential distribution may not provide a realistic model for time-to-event data. Still, due to the simplicity of the equation and the relative ease of calculations, this model may be attractive in specific contexts and for illustrating fundamental characteristics of time-to-event observations.

Weibull Distribution

The probability density function characterized by two parameters for the Weibull distribution is represented as

$$f(t;\gamma,\lambda) = \lambda\gamma t^{\gamma-1} e^{-\lambda t^{\gamma}}, \quad t > 0 \tag{2.13}$$

In this context, γ is referred to as the shape parameter, whereas λ is designated as the scale parameter (Wienke, 2010) [165]. It is important to note that when $\gamma = 1$, the Weibull distribution is equivalent to the exponential distribution characterized by the parameter λ. The cumulative distribution function related to the Weibull distribution is formulated as

$$F(t) = 1 - e^{-\lambda t^{\gamma}}, \quad t > 0 \tag{2.14}$$

Consequently, the associated survival function is

$$S(t) = e^{-\lambda t^{\gamma}} \tag{2.15}$$

Thus, the hazard function is

$$h(t) = \frac{f(t)}{S(t)} = \lambda \gamma t^{\gamma-1} \tag{2.16}$$

It is evident that for $\gamma \neq 1$, the hazard function exhibits variability, in contrast to the constancy observed in the exponential distribution. The form of the hazard function depends on the shape parameter γ, as presented in table 2.1:

Table 2.1: Hazard function for various values of γ

Values of γ	Shape of $h(t)$
$0 < \gamma < 1$	Exponential decline
$\gamma = 1$	Fixed ($h(t) = \lambda$)
$\gamma = 2$	Linear path
$\gamma > 2$	Exponential growth

2.1.3 Non Parametric Methods

This part briefly outlines two frequent non-parametric methods used for the analysis and interpretation of time-to-event data. The Kaplan-Meier estimator of the survival function, in combination with the log-rank test (Mantel, 1963) [116], is applied to determine variations between two groups that present time-to-event results.

The Kaplan-Meier Approach to Survival Function Estimation

Let

$$t_{(1)} < t_{(2)} < \ldots < t_{(m)}$$

represents the distinct chronological actual times of mortality for m subjects are characterized from a participant group of N individuals, while omitting occurrences of censoring. Let d_i characterize the total fatalities at the designated time $t_{(i)}$, and let n_i illustrate the number of individuals who are alive just preceding $t_{(i)}$. This is the number susceptible to risk at time $t_{(i)}$. According to Kleinbaum and Klein (1996) [96], the Kaplan-Meier estimator, referred to as the product limit estimator of the survival function, denotes

$$\hat{S}(t) = \prod_{i:t_{(i)}<t} \left(1 - \frac{d_i}{n_i}\right) \tag{2.17}$$

Similarly, to survive until time t, an individual must first survive until $t_{(1)}$. The individual must then survive until $t_{(2)}$, having already survived until $t_{(1)}$, and the process continues. We suppose that there are no deaths between $t_{(i-1)}$ and $t_{(i)}$, hence the probability of death occurring during this interval is zero. The conditional probability of survival at time $t_{(i)}$ is the complement of $\frac{d_i}{n_i}$, expressed as $\left(1 - \frac{d_i}{n_i}\right)$. The conclusive unconditional probability of enduring until time t is achieved by multiplying the conditional probabilities throughout all relevant time intervals that precede t.

Non-Parametric Maximum Likelihood

Evaluate the likelihood effect of an incident that either endures an event or is subject to censoring at time t_i. Define c_i as the total count of individuals that have been censored within the interval from $t_{(i-1)}$ to $t_{(i)}$, and let d_i signify the number of individuals that die or experience the event at $t_{(i)}$ (Hanagal (2011)) [65]. The likelihood function supposes the following form, which is,

$$L = \prod_{i=1}^{m} \left[S\left(t_{(i-1)}\right) - S\left(t_{(i)}\right)\right]^{d_i} \left[S\left(t_{(i)}\right)\right]^{c_i} \tag{2.18}$$

as the product exceeds the m distinct values, given $t_{(0)} = 0$ with $S(t_{(0)}) = 1$ Based on Kleinbaum and Klein (1996) [96], for the intent of estimating m parameters that represent the values of the survival function at the identified death times $t_{(1)}, t_{(2)}, \ldots, t_{(m)}$, the

conditional probability of surviving from $S\left(t_{(i-1)}\right)$ to $S\left(t_{(i)}\right)$ is expressed by the notation $\pi_i = \frac{S(t_i)}{S(t_{i-1})}$

Thus, $S\left(t_{(i)}\right)$ can be formulated as

$$S\left(t_{(i)}\right) = \pi_1 \pi_2 \ldots \pi_i \tag{2.19}$$

and the likelihood is expressed as

$$L = \prod_{i=1}^{m} (1-\pi_i)^{d_i} \pi_i^{c_i} \left(pi_1 \pi_2 \ldots \pi_{i-1}\right)^{d_i + c_i} \tag{2.20}$$

Taking into account that every instance of mortality occurring at $t_{(i)}$ or those that are censored within the interval between $t_{(i)}$ and $t_{(i+1)}$ contribute a term π_j to each prior time of death from $t_{(1)}$ to $t_{(i-1)}$. Furthermore, individuals who die at $t_{(i)}$ contribute $1-\pi_i$, while censored cases make an additional π_i. Let $n_i = \sum_{j>i}(d_j + c_j)$ represent the overall number of individuals susceptible to risk at $t_{(i)}$. Consequently, combining the terms for each π_i, the likelihood is expressed as

$$L = \prod_{i=1}^{m} (1-\pi_i)^{d_i} \pi_i^{n_i - d_i} \tag{2.21}$$

a binomial likelihood. The maximum likelihood estimator for π_i is hence defined as

$$\hat{\pi}_i = \frac{n_i - d_i}{n_i} = 1 - \frac{d_i}{n_i} \tag{2.22}$$

The KP-M estimator is formulated through the multiplication of these conditional probabilities.

Greenwood's Formula

The previously established likelihood implies that the considerable sample variance of $\hat{p}_i$, dependent on the data n_i and d_i, is expressed through the traditional binomial formula, as

$$\mathrm{Var}\left(\hat{\pi}_i\right) = \frac{\pi_i\left(1-\pi_i\right)}{n_i} \tag{2.23}$$

Assuming that $\mathrm{cov}\left(\hat{\pi}_i, \hat{\pi}_j\right) = 0$ for $i \neq j$, it suggests that the covariances associated with contributions from various time points of mortality are entirely null. The aforementioned statement can be supported by conducting the calculations of logarithmic values, along with the determination of the first and second derivatives of the log-likelihood function. To determine the asymptotic variance of $\hat{S}(t)$, which denotes the Kaplan-Meier estimation of the survival function, we apply the delta method twice. (Kleinbaum and Klein, 1996) [96], initiating by taking logarithms in order to simplify the calculation of the variance of a sum rather than that of a product.

$$K_i = \log \hat{S}\left(t_{(i)}\right) = \sum_{j=1}^{i} \log \hat{\pi}_j \tag{2.24}$$

In order to evaluate the variance of the logarithm of $\hat{\pi}_i$, we use the delta method initially; therefore, the asymptotic variance of a function f related to a random variable X is

$$\mathrm{Vax}(f(x)) = \left(f'(X)\right)^2 \mathrm{var}(X) \tag{2.25}$$

Consequently, for the logarithmic function, the variance is expressed as

$$\mathrm{Var}\left(\log \hat{\pi}_i\right) = \left(\frac{1}{\pi_i}\right)^2 \mathrm{var}\left(\pi_i\right) = \frac{1-\pi_i}{n_i \pi_i} \tag{2.26}$$

Since K_i is a summation and the covariances of the $\hat{\pi}_i$ and therefore of the $\log \hat{\pi}_i$ are null, we conclude that Given that K_i represents a summation and the covariances of the $\hat{\pi}_i$ and consequently of the $\log \hat{\pi}_i$ are equal to zero, we reach the conclusion that

$$\mathrm{Var}\left(\log \hat{S}\left(t_{(i)}\right)\right) = \sum_{j=1}^{i} \frac{1-\pi_j}{n_j \pi_j} = \sum \frac{d_j}{n_j\left(n_j - d_j\right)} \tag{2.27}$$

Applying the delta approach once more, this instance aimed at deriving the variance

of the survival function from the variance of its logarithm, we achieve

$$\text{Var}\left(\hat{S}\left(t_{(i)}\right)\right) = \left[\hat{S}\left(t_{(i)}\right)\right]^2 \sum_{j=1}^{i} \frac{1-\hat{\pi}_j}{n_{j\hat{\pi}_j}} \tag{2.28}$$

The result is referred to as Greenwood's formula..

Mantel-Haenszel

Examine the issue of comparing many survival functions, such as urban versus rural in Lesotho. Let

$$t_{(1)} < t_{(2)} < \ldots < t_{(m)}$$

indicate the different times of death recorded in the overall sample, derived from the combination of all pertinent groups. Let

d_{ij} = the occurrence of mortality at the temporal point $t_{(i)}$ within the group j, and,

n_{ij} = subjects exposed to potential hazards at temporal point $t_{(i)}$ within the group j,

d_i = total number of mortality, and

n_i = child subjected to potential hazards at time $t_{(i)}$.

The d_i fatalities that occur at time $t_{(i)}$ should be distributed among the k cohorts in relation to the proportion of individuals at risk, assuming that the survival probabilities are consistent across all groups. Consequently, dependent on d_i and n_{ij},

$$E\left(d_{ij}\right) = d_i \frac{n_{ij}}{n_i} \tag{2.29}$$

An alternative interpretation of this calculation involves the application of a comprehensive failure rate $\frac{d_i}{n_i}$ to the n_{ij} subjects within group j, as demonstrated by the last term.

The Mantel-Haenszel statistic, commonly referred to as the log-rank test, assesses the null hypothesis, which claims that the risk of mortality is identical across two or more groups. For the sake of simplicity, we examine the scenario involving two groups (Klein-

baum and Klein, 1996) [96]. In fact, if the study focused on comparing child mortality rates within the rural and urban sectors of Lesotho, the null hypothesis would maintain that there is an absence of significant variation in the probability of infant mortality across the two cohorts. The examination of the statistical test is elaborated upon in detail in the subsequent section.

Let the two groups be denoted as 1 and 2, representing urban and rural locations, respectively, and let there be k distinct times, $t_1 < t_2 < \ldots < t_k$, across both groups. The test applies a conditional argument depending on the number at risk of failure before each recorded failure time. Let t_i be a certain time point, at which there are d_i total deaths and n_i individuals at risk. In this context, group 1 experiences d_{i1} deaths and has n_{i1} individuals at risk, whereas group 2 has d_{i2} deaths and n_{i2} subjects exposed to risk, satisfiying $d_{i1} + d_{i2} = d_i$ and $n_{i1} + n_{i2} = n_i$. For each recorded time of mortality t_i, this dataset may be presented in the table 2.2:

Table 2.2: Mortality number at time t_i

Group	Fatalities	Survivors	Total
1	d_{i1}	$n_{i1} - d_{i1}$	n_{i1}
2	d_{i2}	$n_{i2} - d_{i2}$	n_{i2}
Total	d_i	$n_i - d_i$	n_i

In addition to the correlated survival durations, $d_i = 1$, thus indicating that either d_{i1} or d_{i2} takes on the value of 0 or 1. If a child experiences censorship at a defined temporal point t_i, the child is still susceptible to potential risks at that moment and is incorporated into n_i. The assumption implies that censoring occurs after the event (Kleinbaum and Klein, 1996) [96]. If the null hypothesis holds, the mortality number at any given period is supposed to conform to the hypergeometric distribution; thus,

$$E(d_{i1}) = e_{i1} = \frac{n_{i1} d_i}{n_i} \tag{2.30}$$

and the variance is characterized by

$$\text{Var}(d_{i1}) = \frac{d_i(n_i - d_i)n_{i1}n_{i2}}{n_i^2(n_i - 1)} \tag{2.31}$$

The variation between d_{i1} and e_{i1} forms the basis for the statistical methodologies applied in the analysis of the null hypothesis. The log-rank test compiles these variations for each death time (Kleinbaum and Klein, 1996) [96]. The summation of the numerous measures across periods of mortality yields

$$O_1 = \sum_i d_{i1}$$
$$E_1 = \sum_i e_{i1}, \text{ and}$$
$$V_1 = \sum_i \text{Var}(d_{i1})$$

Here, E_1 illustrates the expected number of deaths in group 1 for the total duration. Conversely, O_1 indicates the overall count of fatalities that have been documented within the cohort. The variance associated with the deviation $O_1 - E_1$, under the presumption of independent event timings, is characterized as V_1. The test statistic is then defined by

$$\chi_1^2 = \frac{(O_1 - E_1)^2}{V_1} \tag{2.32}$$

which, under H_0, follows a chi-squared distribution with 1 degree of freedom (Mantel, 1963) [116]. If the value calculated exceeds the value associated with the chi-square distribution at an established significance level, the null hypothesis, which claims the absence of a significant difference, is thereafter invalidated, implying that the risk of death varies between the two groups (Kleinbaum and Klein, 1996) [96].

Alternatively, assuming the variations $d_{i1} - e_{i1}$, for $i = 1, 2, \ldots, k$, are considered to be independent,

$$Z = \frac{O_1 - E_1}{\sqrt{V_1}} \tag{2.33}$$

must indicate an approximately normal distribution, and the null hypothesis is considered invalidated for substantial values of Z. Specifically, at a significance level of 5%, the null hypothesis is dismissed if the observed Z surpasses 1.96. The ratios $\frac{O_1}{E_1}$ and $\frac{O_2}{E_2}$ are designated as the relative mortality rates, quantifying the proportion of the mortality rate within each distinct group regarding the collective mortality rate of the two groups. The proportional relation between these two comparative rates facilitates the estimation of the mortality rate for group 1 in relation to that of group 2.

The log-rank test can be modified for the analysis of mortality equality rates across $s > 2$ categories. The test statistic, possessing $(s-1)$ degrees of freedom, is then defined by

$$\chi^2_{s-1} = \frac{(O_1 - E_1)^2}{V_1} + \frac{(O_2 - E_2)^2}{V_2} + \frac{(O_3 - E_3)^2}{V_3} + \dots \tag{2.34}$$

If the evaluated statistic surpasses the established critical value at the α significance minimum level, we will refuse the null hypothesis that suggests an absence of variability in survival or hazard functions among the groups examined.

Several critical insights pertinent to the formulation of the log-rank statistic are delineated in the following discussion. At the outset, the vector representing the difference between observed and anticipated failures lacks independent constituents, and the central limit theorem typically employed to establish asymptotic normality is not applicable in this context. Furthermore, differences between observed and expected failures have an equal importance regardless of the risk set, particularly, the total number of cases remaining under investigation at the point of documented failures is of considerable importance, which will have an effect on the cumulative test statistic. The supplementary facets of significance testing are thoroughly analyzed in the comprehensive work concerning counting processes and survival analysis demonstrated by Fleming and Harrington (Fleming and Harrington, 2013) [49].

2.2 Regression Models in Survival Analysis

The non-parametric techniques discussed in the preceding chapter, specifically the log-rank test and the Kaplan-Meier estimator, are unable to account for covariates; consequently, extensions that incorporate covariates are required. These non-parametric techniques do not account for covariates and necessitate the use of categorical predictors. When multiple prognostic variables are present, multivariate methods are required; however, multiple linear regression or logistic regression cannot be employed due to their inability to handle censored observations (Cox, 1972) [34]. An alternative approach is required to model survival data in the presence of censoring. A widely used model in the field of survival analysis is the Cox proportional hazards (PH) model, which was first proposed by Cox (Cox, 1972) [34].

2.2.1 Cox Proportional Hazard (PH) Model

Cox's (1972) [34] introduction of the proportional hazard model provides a regression model where the temporal characteristics of events act as the dependent variable. On the other hand, the regression framework is developed on the assumption of the hazard function on various covariates. It enables the integration of pertinent information regarding observed covariates into models of time-to-event data in a simple way.

The Cox proportional hazards model is described as

$$h(t \mid X) = h_0(t) \exp\left(\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p\right) = h_0(t) \exp\left(\beta' X\right) \tag{2.35}$$

where $h_0(t)$ is characterized as the baseline hazard function, denoting the hazard function for an individual for whom all variables in the model are fixed at zero, $X = (x_1, x_2, \ldots, x_p)$ denotes the attributes of the vector of explanatory variables pertinent to a particular subject, and $\beta' = (\beta_1, \beta_2, \ldots, \beta_p)$ denotes a vector of regression coefficients.

The survival function associated with the covariate X is

$$S(t \mid X) = [S_0(t)]^{\exp(\beta' X)} \tag{2.36}$$

where $S_0(t) = \exp\left(-\int_0^t h_0(u)du\right)$ denotes the baseline survival function, while the components of the vector β consist of unidentified regression coefficients. As a consequence, the survival function that corresponds to an individual delineated by the covariate vector X is represented as a power function of the baseline survival function.

This framework is designated as the semi-parametric model, as it imposes no restrictions on the structure of $h_0(t)$ while assuming a parametric specification for the effect of predicting factors on the hazard (Hanagal, 2011) [65]. Despite the fact that the baseline hazard is unspecified, we can still obtain accurate estimates for the regression coefficients β, hazard ratios, and modified hazard curves (Cox, 1972) [34].

The measure used to quantify the effect is known as the hazard ratio. The hazard ratio associated with two individuals exhibiting differing baseline or covariates constant over time, X and X^*, is

$$\hat{HR} = \frac{h_0(t)\exp\left(\hat{\beta}' X\right)}{h_0(t)\exp\left(\hat{\beta}' X^*\right)} = \exp\left(\sum \hat{\beta}'\left(X - X^*\right)\right) \tag{2.37}$$

The hazard ratio is time-independent, so it is known as the proportional hazard model. The risk associated with an individual possessing covariates X shows a proportional relationship to the risk associated with an individual possessing covariates X^*, as demonstrated by equation (2.37).

Partial Likelihood Evaluation in the Context of the Cox PH Model

The primary objective in the application of the Cox proportional hazards model is to estimate the regression coefficients, $(\beta_1, \ldots, \beta_p)$, with each individual parameter β_j representing the log-hazard ratio associated with the covariate or predictor variable X_j (Cox (1975)) [35]. If X_j constitutes a binary classifier, then $\exp\left(\beta_j\right)$ signifies the hazard ratio

of one group in relation to the reference group. Conversely, in the case where X_j is a continuous variable, $\exp(\beta_j)$ indicates the hazard ratio associated with a rise of one unit in X_j. One approach involves optimizing the likelihood function for the observed data concurrently in relation to β. A prevalent technique was introduced by Cox (1975) [35], where a partial likelihood function independent of $h_0(t)$ is established for β. The partial likelihood represents a statistical technique formulated to estimate the regression parameters while accommodating a nuisance parameter $(h_0(t))$ within the Cox proportional hazards model.

Let $t_1, t_2, \ldots, t_n$ denote the recorded survival times for n people. Denote the ordered event times of r persons who undergo the event of interest as $t_{(1)} < t_{(2)} < \cdots < t_{(r)}$. Let $R(t_{(j)})$ represent the risk set immediately preceding $t_{(j)}$, with r_j indicating its size. Thus, $R(t_{(j)})$ indicates the cohort of individuals who are both alive and uncensored immediately before $t_{(j)}$. The conditional probability that the individual indexed as i^{th} fails at $t_{(j)}$, dependent on the occurrence of one death among the individuals within the risk set denoted by $R(t_{(j)})$ at $t_{(j)}$, is

$$\frac{\exp(\beta' X_i(t_j))}{\sum_{k \in R(t_{(j)})} \exp(\beta' X_k(t_{(j)}))} \tag{2.38}$$

Subsequently, the formulation of the partial likelihood function pertinent to the Cox proportional hazards model is defined as

$$L(\beta) = \prod_{j=1}^{r} \frac{\exp(\beta' X_i(t_j))}{\sum_{k \in R(t_{(j)})} \exp(\beta' X_k(t_{(j)}))} \tag{2.39}$$

$X_i(t_{(j)})$ denotes the vector of covariate values associated with individual i who experiences death outcome at time point $t_{(j)}$. The methodology of partial likelihood was initially proposed by Cox in 1975 [35]. This particular likelihood function is restricted to individuals whose data is not censored. Consider the survival times observed for n individuals denoted as $t_1, t_2, \ldots, t_n$ while δ_i shows the event occurrence indicator, which is null if the i^{th} survival time is subject to censoring and one in all other instances. The

likelihood function delineated in the preceding equation can be defined as

$$L(\beta) = \prod_{j=1}^{n} \left[\frac{\exp\left(\beta' X_i\left(t_j\right)\right)}{\sum_{k \in R\left(t_{(j)}\right)} \exp\left(\beta' X_k\left(t_{(j)}\right)\right)} \right]^{\delta_i} \tag{2.40}$$

where $R(t_i)$ signifies the group of individuals exposed to risk at the temporal point t_i.

The partial likelihood is still applicable in the absence of ties within the dataset; that is, no two subjects have equal event times.

Examination of the Cox Proportional Hazards Assumption

The core assumption that forms the foundation of the Cox proportional hazards model is the principle of proportional hazards (Kleinbaum and Klein, 1996) [96]. The concept of proportional hazards suggests that the hazard function associated with one individual is proportional to the hazard function of another individual, thereby implying that the hazard ratio remains constant over time. Several approaches exist to verify whether a model corresponds to the proportionality assumption, including the graphical method, the incorporation of time-dependent covariatess (Kleinbaum and Klein, 1996) [96].

Graphical Method

The extraction of the Cox proportional hazards survival function can be achieved by exploring the connection between the hazard function and the survival function, which is delineated as follows:

$$S(t,X) = [S_0(t)]^{\exp\left(\sum_{i=1}^{p} \beta_i x_i\right)} \tag{2.41}$$

where $X = (x_1, x_2, \ldots, x_p)$ indicates the values of the explanation variable vector related to a particular individual. By employing the logarithmic function twice, we can effectively illustrate that

$$\ln[-\ln S(t,X)] = \sum_{i=1}^{p} \beta_i x_i + \ln[-\ln S_0(t)] \tag{2.42}$$

The variation identified in the log-log graphs for two distinct subjects defined by variables $X_1 = \left(x_{11}, x_{12}, \ldots, x_{1p}\right)$ and $X_2 = \left(x_{21}, x_{22}, \ldots, x_{2p}\right)$ is formulated as

$$\ln[-\ln S(t,X_1)] - \ln[-\ln S(t,X_2)] = \sum_{i=1}^{n} \beta_i (x_{1i} - x_{2i}) \tag{2.43}$$

which is independent of t, considering that the two covariate vectors X_1 and X_2 show no temporal dependence. This correlation is essential for identifying scenarios where proportionate hazards may exist. Plotting estimated $\log(-\log(\text{ survival }))$ against survival time for two groups will result in parallel curves if hazards are proportional (Kleinbaum and Klein, 1996) [96]. This approach is ineffective for continuous predictors or categorical predictors with multiple levels due to the resulting clutter in the groups. Moreover, the curves are sparse when limited time points are present, making it challenging to determine how close to parallel is close enough (Kleinbaum and Klein, 1996) [96].

Nevertheless, examining the K-M curves and $\log(-\log(\text{ survival }))$ alone is insufficient to determine proportionality, as these represent univariate analyses that do not guarantee that hazards will remain proportional when a model incorporates multiple other predictors; however, they support our assertion of proportionality. Nonetheless, alternative statistical methods exist for assessing proportionality (Kleinbaum and Klein, 1996) [96].

Incorporating a Time-Dependent Term in Cox PH Model

In order to formulate a term that is dependent on time, we establish a relationship between covariates and the survival function, subsequently integrating it into the model. For instance, if the variable of interest is X_j, we construct a time-dependent term or covariate represented as $X_j(t)$, where $X_j(t) = X_j \times g(t)$, with $g(t)$ representing a temporal function, such as t, $\log t$, or the Heaviside step function of t. The methodological framework evaluating the validity of the proportional hazards assumption for X_j, whilst accounting for additional covariates, is

$$h(t,X(t)) = h_0(t) \exp\left[\beta_1 x_1 + \beta_2 x_2 + \ldots \beta_j x_j + \ldots \beta_p x_p + \delta x_j \times g(t)\right] \tag{2.44}$$

where $X(t) = \left(x_1, x_2, \ldots, x_p, x_j(t)\right)'$ denotes the vector of covariates corresponding to

a specific individual. The null hypothesis employed for the evaluation of the PH assumption pertaining to X_j is articulated as $\delta = 0$, while the alternative hypothesis is $\delta \neq 0 (H_0 : \delta = 0$ virsus $H_a : \delta \neq 0$). The evaluation process can be conducted using either a Wald test or a likelihood ratio test. In the context of the Wald test, the statistic is obtained based on estimation.

$$W = \left(\frac{\hat{\delta}}{\text{se}(\hat{\delta})} \right)^2 \tag{2.45}$$

which asymptotically adheres to a chi-square distribution characterized by one degree of freedom. The likelihood ratio (LR) test statistic assesses the likelihoods proposed under the null hypothesis H_0 in comparison to the alternative hypothesis H_a. The theoretical framework corresponding to the null hypothesis, H_0, must be integrated within the alternative hypothesis model, H_a. As a result, the likelihood ratio test statistic can be articulated as

$$LR = -2 \ln \left(\frac{L_0}{L_a} \right) = -2 (\ell_0 - \ell_a) \tag{2.46}$$

where ℓ_0 and ℓ_a denote the log likelihoods associated with the respective hypotheses. The statistical measure conforms to a chi-square distribution with one degree of freedom in the context of the null hypothesis. In instances where a time-dependent covariate demonstrates statistical significance, leading to the rejection of the null hypothesis, the predictor is considered to violate the proportional hazards assumption. Similarly, we can evaluate the PH assumption for several predictors sequentially.

2.2.2 Parametric PH Model

The Cox proportional hazards model delineated in the preceding section represents the most prevalent statistical methodology used for modeling survival data, especially in health research such as clinical trials. This can be associated with the premise that this model yields insights into the parameters without requiring any specific distribution concerning the survival duration. Nonetheless, in cases where the proportional hazards

assumption is violated, the suitability of these models is considered inadequate. On the other hand, the PH assumption may be valid, but an appropriate distribution for the time-to-event variable should be considered. In this situation, a fully parametric model may be utilized. In this segment, we delineate parametric models that presuppose particular probability distributions pertaining to survival durations. Initially, we shall introduce the parametric proportional hazards model. Secondly, we shall present the accelerated failure time (AFT) model and engage in a comprehensive discourse regarding the exponential, Weibull, log-logistic, log-normal, and gamma AFT models.

The parametric proportional hazards (PH) model constitutes a parametric adaptation of the Cox proportional hazards model. It is expressed in a way that is consistent with the framework of the Cox proportional hazards model. The hazard function at a specific temporal point t for an individual of interest is characterized by a set of p covariates $(x_1, x_2, \ldots, x_p)$ is expressed as follows:

$$h(t \mid X) = h_0(t) \exp\left(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p\right) = h_0(t) \exp\left(\beta' X\right) \tag{2.47}$$

The principal distinction between the two classes of models resides in the assumption that the baseline hazard function adheres to a specified distribution, leading to a fully parametric proportional hazards model, while the Cox PH model imposes no such restriction. In the Cox model, coefficients are estimated using partial likelihood, whereas in the parametric proportional hazards model, they are estimated by full maximum likelihood estimation. Apart from this, the two kinds of models are equivalent. Hazard ratios retain their interpretation, and the proportional hazards assumption remains valid in both cases. Various parametric PH models can be formulated by selecting various hazard functions. The models that are often employed in research involve the Exponential, Weibull, and Gompertz models.

Exponential PH Model

The exponential proportional hazards model constitutes a particular instance of the Weibull model when the parameter γ is equal to one. It is postulated that the hazard function associated with the exponential proportional hazards model maintains a constant value throughout the duration of the study. The baseline survival and hazard functions pertinent to this model are delineated as follows:

$$S(t) = \exp(-\lambda_0 t) \tag{2.48}$$

and the hazard function is defined as

$$h(t) = \lambda_0 \tag{2.49}$$

The hazard model for a specific individual characterized by covariates $x_{i1}, x_{i2}, \ldots, x_{ip}$ is expressed as follows

$$h(t \mid X) = \lambda_0 \exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p) = \lambda_0 \exp\left(\beta' X\right) \tag{2.50}$$

The exponential **PH** model can be extended to the piecewise constant exponential model (Breslow, 1974) [24]. In this model, the follow-up period, denoted as $[0, T]$, is segmented into K intervals $\left(t_j, t_{j+1}\right]$ for $j = 1, 2, \ldots, K$, with $t_1 = 0$ for simplification. The method assumes that the baseline hazard remains constant within each defined interval, while exhibiting variability across different intervals, such that $h_0(t) = \exp\left(\alpha_j\right) = \lambda_j$ for t_j, implying a step-function approximation of the baseline hazard. The segmented exponential model is characterized as

$$\lambda_{ij} = \lambda_j \exp\left(\beta' x_i\right) \tag{2.51}$$

where λ_{ij} denotes the hazard for individual i during interval j, and $\exp(\beta' x_i)$ characterizes the corresponding hazard for an individual with a covariate value x_i in relation to

the baseline at any specified temporal point.

The segmented exponential methodology employs a log-linear framework for the assessment of the effects of covariates as well as the primary hazard function. The derivation of estimates for the hazard function and regression coefficients can be obtained through the process of maximum likelihood estimation. The maximum likelihood approach for estimating the baseline hazard function in the interval j, given the regression coefficients β, is delineated as

$$\hat{\lambda}_j = \frac{d_j}{\sum_{i \in R_j} \exp\left(\hat{\beta}' x_i\right) t_{ij}} \tag{2.52}$$

Here, d_j represents the number of occurrences in interval j, R_j indicates the risk set entering interval j, and t_{ij} denotes the observed survival time for individual i within interval j (Holford and Sheiner (1982) [73]; Holford and Sheiner (1981) [72]). The principal challenge associated with the implementation of the piecewise exponential model lies in determining a suitable grid of temporal points requisite for its formulation. The merit of this methodology is its capacity to integrate covariates that vary with time. For any temporal covariates, the corresponding values at the initiation of each interval may be attributed to the records relevant to that particular time interval.

Weibull PH Model

The temporal duration of survival is expressed via the Weibull distribution, which is distinguished by the scale parameter λ and the shape parameter γ. As a result, the survival and hazard functions of a $W(\lambda, \gamma)$ distribution are specified as follows:

$$S(t) = \exp\left(-\lambda t^{\gamma}\right) \tag{2.53}$$

and the baseline hazard function is

$$h(t) = \lambda_0 \gamma t^{\gamma - 1} \tag{2.54}$$

where $\lambda, \gamma > 0$. The hazard rate demonstrates an increasing trend when $\gamma > 1$ and

exhibits a decreasing trend when $\gamma < 1$ as time progresses. Within the framework of the Weibull Proportional Hazards model, the hazard function for a specific individual characterized by covariates $x_{i1}, x_{i2}, \ldots, x_{ip}$, where i represents the subject, is formulated as

$$h(t \mid X) = \lambda_0 \gamma t^{\gamma-1} \exp\left(\beta' X\right) \tag{2.55}$$

The Weibull family, characterized by a constant γ, adheres to the proportional hazards (**PH**) property. This suggests that the covariates within the model influence the scale parameter of the distribution, while the shape parameter remains unchanged. The corresponding survival function is articulated as

$$S(t \mid X) = \exp\left\{-\exp\left(\beta' X\right) \lambda_0 t^{\lambda}\right\} \tag{2.56}$$

The baseline survival function (without covariates) of the Weibull distribution can be adjusted to yield the appropriate corresponding equation.

$$\log(-\log(S(t))) = \log \lambda_0 + \gamma \log t \tag{2.57}$$

A graphical representation of $\log(-\log(S(t)))$ against $\log(t)$ will exhibit a nearly linear relationship if the presumption of the Weibull distribution holds true. The y-intercept and the gradient of the linear equation will function as approximate estimations for $\log \lambda_0$ and γ, correspondingly. If the two lines depicting the two cohorts in this graph exhibit nearly parallel characteristics, it signifies the robustness of the proportional hazards model. Additionally, in situations where the linear trajectory demonstrates a gradient close to unity, the basic exponential distribution gains significance. In an exponential distribution, it follows that $\log S(t) = -\lambda_0 t$. Consequently, we can examine the graph of $-\log S(t)$ in relation to $\log t$. If the exponential distribution is suitable, this ought to represent a linear trajectory intersecting the origin.

Gompertz PH Model

According to Gompertz proportional hazards model, the foundational survival and hazard distributions are characterized by

$$S(t) = \exp\left(\frac{\lambda_0}{\theta}\left(1 - e^{\theta t}\right)\right) \tag{2.58}$$

and the hazard function is

$$h(t) = \lambda_0 \exp(\theta t) \tag{2.59}$$

where t resides within the interval $[0, \infty)$ and $\lambda_0 > 0$. The parameter θ illustrates the shape of the hazard function. At the point where $\theta = 0$, the hazard function aligns with that of an exponential distribution. The exponential distribution constitutes a specific instance of the Gompertz distribution. Analogous to the Weibull hazard function, the Gompertz hazard function reveals a continuous increase or decrease. For the Gompertz distribution, $\log h(t)$ shows a linear relation with respect to t.

According to the Gompertz PH model, the hazard function for an individual is expressed as

$$h(t \mid x) = \lambda_0 \exp(\theta t) \exp\left(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p\right) = \lambda_0 \exp\left(\beta' X\right) \exp(\theta t) \tag{2.60}$$

The Gompertz model demonstrably possesses the PH feature (Hougaard, 1986) [77]. Despite its limited application in reality, the Gompertz distribution possesses some attractive characteristics, including the PH property.

2.2.3 Accelerated Failure Time (AFT) Model

A different parametric structure for the assessment of time-to-event data is represented by the accelerated failure time model. Such models operate under the assumption of a particular probability distribution that determines survival time. Within the framework

of the AFT model, we assess the direct effect of the explanatory variables on survival time duration rather than on hazard, as previously stated in the **PH** model. This approach allows a simpler interpretation of the outcomes, since the parameters directly assess the influence of a specific covariate on the duration of survival time. At this time, the AFT model is not frequently applied to the evaluation of clinical trial data, despite its notable presence in the manufacturing industry. In proximity to the Proportional Hazards (PH) model, the Accelerated Failure Time (AFT) model delineates the association between survival probabilities and a variety of covariates.

The assumption of the AFT model can be expressed mathematically as $S_2(t) = S_1(\eta t)$ for $t \geq 0$, where $S_1(t)$ and $S_2(t)$ represent the survival functions corresponding to group one and group two, respectively, and η denotes a constant referred to as the acceleration factor that compares the two groups. Within the context of the regression framework, the acceleration factor η can be characterized as $\exp(\alpha)$, where α serves as a parameter that requires estimation from the empirical data. Employing this particular parameterization, the Accelerated Failure Time (AFT) assumption can be articulated as $S_2(t) = S_1(\exp(\alpha)t)$ or, conversely, $S_2(\exp(-\alpha)t) = S_1(t)$ for the domain $t \geq 0$. The AFT assumption may also be articulated in terms of random variables that denote the duration of survival as opposed to the survival function itself. If T_2 is construed as a random variable indicative of the survival time for the second cohort, whereas T_1 is a random variable that expresses the survival time for the first cohort, the Accelerated Failure Time (AFT) premise can be formulated as $T_1 = \eta T_2$.

The acceleration factor acts as an essential indicator for clarifying the connection established from an Accelerated Failure Time (AFT) model. It provides an evaluation of the impact of predictor variables on survival time, similar to how the hazard ratio enables the analysis of predictor variables concerning hazard rates.

The acceleration factor illustrates the elaboration of survival functions when performing a comparative analysis across different groups. The acceleration factor is defined as the ratio of survival times associated with a specific value of $S(t)$; supplementary illustrations will be presented in the upcoming subsections.

In this segment, we aim to perform an estimation of the Exponential Accelerated Failure Time model, the Weibull Accelerated Failure Time model, and the Log-logistic Accelerated Failure Time model.

Exponential AFT Model

For an exponential distribution, the survival function is expressed as $S(t) = \exp(-\lambda t)$, while the hazard function is denoted by $h(t) = \lambda$. In this subsection, we demonstrate the reparameterization of $S(t)$ as an Accelerated Failure Time (AFT) model. The AFT assumption for comparing two levels of covariates maintains that the ratio of durations for any designated value of $S(t) = q$ is invariant across any probability q. We develop the model utilizing the survival function and derive t as a function of $S(t)$. Subsequently, we normalize t concerning the predictor variable. For example, considering the exponential form of the survival function

$$S(t) = \exp(-\lambda t) \tag{2.61}$$

Determining the variable t entails the preliminary computation of the natural logarithm, followed by the multiplication of both sides by -1, and subsequently the multiplication by the reciprocal of λ, leading to the conclusion that

$$t = [-\ln S(t)] \times \frac{1}{\lambda} \tag{2.62}$$

Considering $\frac{1}{\lambda} = \exp(\alpha_0 + \alpha_1 \text{ GROUP })$ where GROUP is denoted as 1 for the second group and as 0 for the first group. Accordingly, t is adjusted into

$$t = [-\ln S(t)] \times \exp(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.63}$$

Time corresponding to a specified survival probability $S(t)$ is scaled through the pre-

dictor variable GROUP. Let $S(t) = q$, it follows that t will be

$$t = [-\ln q] \times \exp(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.64}$$

The acceleration factor η is derived from the calculation of the ratio of the time intervals necessary to reach $S(t) = q$ for $\mathbf{GROUP} = 1$ relative to $\mathbf{GROUP} = 0$, distinctly illustrated as

$$\eta = \frac{[-\ln(q)] \times \exp(\alpha_0 + \alpha_1)}{[-\ln(q)] \times \exp(\alpha_0)} = \exp(\alpha_1) \tag{2.65}$$

Upon the annulment of the preceding conditions, η is diminished to the expression $\exp(\alpha_1)$.

The estimations of parameters may be utilized to ascertain the temporal value $\hat{t}$ that corresponds to any specified value of q; as an illustration, one can estimate the duration in years associated with the first quartile ($q = 0.25$), the median ($q = 0.5$), and the third quartile ($q = 0.75$). The primary attribute of the exponential model is its associated acceleration factor and hazard ratio. $\mathbf{GROUP} = 1$ and $\mathbf{GROUP} = 0$ are reciprocals, as seen in the table 2.3:

Table 2.3: Acceleration factor and hazard ratio

AFT		HR	
$\eta > 1 \Rightarrow$	Exposure positive to survivability	$HR > 1 \Rightarrow$	Exposure negative to survivability
$\eta < 1 \Rightarrow$	Exposure negative to survivability	$HR < 1 \Rightarrow$	Exposure positive to survivability
$\eta = 1 \Rightarrow$	No impact from exposure	$HR = 1 \Rightarrow$	No impact from exposure

While the exponential (**PH**) and accelerated failure time (AFT) models are predicated on distinct basic assumptions, they are, in fact, equivalent. The unique differentiation is based on their parameterization. The obtained estimations for the survival function, hazard function, and median survival exhibit similarities across these models.

Weibull AFT Model

The representation of an Accelerated Failure Time (AFT) model can further be expressed via the Weibull distribution. The parameterization of the AFT is constructed in a way comparable to that used in association with the exponential model, by inverting the survival function to obtain t for a fixed $S(t)$. The Weibull survival function is expressed as

$$S(t) = \exp(-\lambda t^{\gamma}) \tag{2.66}$$

To solve for t, one first takes the natural logarithm, then the operation involves multiplying both sides by -1, subsequently elevating the resultant outcome to the exponent of $\frac{1}{\gamma}$, and thereafter multiplying by the inverse of $\lambda^{\frac{1}{\gamma}}$, culminating in the following expression for t.

$$t = [-\ln S(t)]^{\frac{1}{\gamma}} \times \frac{1}{\lambda^{\frac{1}{\gamma}}} \tag{2.67}$$

letting

$$\frac{1}{\lambda^{\frac{1}{\gamma}}} = \exp(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.68}$$

then t becomes

$$t = [-\ln S(t)]^{\frac{1}{\gamma}} \times \exp(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.69}$$

By reparameterizing, we have $\frac{1}{\lambda^{\frac{1}{\gamma}}} = \exp(\alpha_0 + \alpha_1 \text{ GROUP })$ By incorporating the predictor variable GROUP, time is scaled to correspond to a particular value of $S(t)$.

For any predetermined probability, $S(t) = q$. In order to derive an expression for the median survival time t_m, set $q = 0.5$, resulting in t becoming

$$t = [-\ln q]^{\frac{1}{\gamma}} \times \exp(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.70}$$

Regardless, the median survival duration, indicated by $q = 0.5$, denotes that

$$t = [-\ln(0.5)]^{\frac{1}{\gamma}} \times \exp(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.71}$$

In a comparable manner, the acceleration coefficient, η(**GROUP** $= 1$ versus GROUP $= 0$) is defined by

$$\eta = \frac{[-\ln(q)]^{\frac{1}{\gamma}} \exp(\alpha_0 + \alpha_1)}{[-\ln(q)]^{\frac{1}{\gamma}} \exp(\alpha_0)} = \exp(\alpha_1) \tag{2.72}$$

the acceleration factor η is formulated as the proportional relationship of the durations to $S(t) = q$ for (GROUP $= 1$ in comparison to GROUP $= 0$). Upon simplification, η reduces to $\exp(\alpha_1)$. Similar to the **PH** model, this outcome is dependent on γ, rather than the status of variation by groups; otherwise, η would be dependent on q.

Log-Logistic AFT Model

The log-logistic distribution includes an Accelerated Failure Time (AFT) model; however, it does not integrate a Proportional Hazards (PH) model. The hazard function is illustrated below.

$$h(t) = \frac{\lambda \gamma t^{\gamma-1}}{1 + \lambda t^{\gamma}}, p > 0, \lambda > 0 \tag{2.73}$$

In contrast to the Weibull model, the log-logistic accelerated failure time (AFT) model is not classified as a proportional hazards model. Instead, the log-logistic AFT model is correctly identified as a proportional odds (PO) model. This model is characterized by the assumption that the odds ratio remains invariant over time. This exhibits similarities to a proportional hazards model in which the hazard ratio is presumed to remain constant across time frames.

We establish the AFT parameterization by determining the value of t as a function of a fixed $S(t)$, where

$$S(t) = \frac{1}{1 + \lambda t^{\gamma}} = \frac{1}{1 + \left(\lambda^{\frac{1}{\gamma}}\right)^{\gamma}} \tag{2.74}$$

To solve t from the expression for $S(t)$, one must first take the reciprocals, subtract-

ing 1, elevate the resultant value to the exponent of $\frac{1}{\gamma}$, subsequently multiplying by the inverse of $\lambda^{\frac{1}{\gamma}}$, thereby yielding the expression for t as delineated below:

$$t = \left[\frac{1}{S(t)} - 1\right]^{\frac{1}{\gamma}} \times \frac{1}{\lambda^{\frac{1}{\gamma}}} \tag{2.75}$$

letting

$$\frac{1}{\lambda^{\frac{1}{\gamma}}} = \exp\left[\alpha_0 + \alpha_1 \text{GROUP}\right] \tag{2.76}$$

Similarly, t becomes into

$$t = \left[\frac{1}{S(t)} - 1\right]^{\frac{1}{\gamma}} \times \exp\left[\alpha_0 + \alpha_1 \text{GROUP}\right] \tag{2.77}$$

Through the process of reparameterization, we have $\frac{1}{\lambda^{\frac{1}{\gamma}}} = \exp\left[\alpha_0 + \alpha_1 \text{ GROUP }\right]$; We permit the predictor variable GROUP to enable the multiplicative transformation of time to any fixed value of $S(t)$. The formulation for t initiated with the substitution $S(t) = q$, which results in the following for t:

$$t = \left[\frac{1}{q} - 1\right]^{\frac{1}{\gamma}} \times \exp\left[\alpha_0 + \alpha_1 \text{GROUP}\right] \tag{2.78}$$

The acceleration factor η is determined through the calculation of the ratio of the time interval necessary to achieve $S(t) = q$ for GROUP $= 1$ relative to GROUP $= 0$. With the terms canceled, η is adjusted to $\exp(\alpha_1)$; thus, $\eta(\text{GROUP} = 1, \text{GROUP} = 0)$ is

$$\eta = \frac{\left[q^{-1} - 1\right]^{\frac{1}{\gamma}} \exp(\alpha_0 + \alpha_1)}{\left[q^{-1} - 1\right]^{\frac{1}{\gamma}} \exp(\alpha_0)} = \exp(\alpha_1) \tag{2.79}$$

Association Between Weibull AFT and PH Coefficients

The coefficients derived from the PH and AFT variants of the Weibull models are connected by the equation $\beta_j = -\alpha_j$ for the j^{th} covariate. This is most clearly demonstrated by reformulating the parameterization in terms of $\ln(\lambda)$ for both the PH and AFT forms

of the model, as illustrated below. For AFT: by redefining the parameterization with respect to $\ln(\lambda)$ for both the **PH** and AFT configurations of the model, as shown in the subsequent illustration. Regarding AFT:

$$\lambda^{\frac{1}{\gamma}} = \exp[-(\alpha_0 + \alpha_1 \text{ GROUP })] \tag{2.80}$$

Applying the natural logarithm to both sides results in

$$\frac{1}{\gamma}\ln\lambda = -(\alpha_0 + \alpha_1 \mathbf{GROUP}) \tag{2.81}$$

solving for $\ln\lambda$ that is,

$$\ln\lambda = -\gamma(\alpha_0 + \alpha_1 \text{ GROUP }) \tag{2.82}$$

and for proportional hazard

$$\lambda = (\beta_0 + \beta_1 \text{ GROUP }) \tag{2.83}$$

solving for $\ln\lambda$ yields

$$\ln\lambda = \beta_0 + \beta_1 \mathbf{GROUP} \tag{2.84}$$

Therefore, the relationship among the coefficients is

$$\beta_j = -\alpha_j\gamma \tag{2.85}$$

so that

$$\beta = -\alpha \tag{2.86}$$

for exponential ($\gamma = 1$).

Chapter 3

Frailty models

The framework, characteristics, and application of survival models to real-world issues are based on the characteristics of time-to-event data, additional information regarding significant causes and processes, and the objectives of the research. This chapter examines the study of univariate data, specifically single-spell data concerning unrelated individuals.

Traditional survival models address the fundamental context of data that is independent and identically distributed. This framework is predicated upon the presumption that the population under investigation is homogeneous. Conversely, observations in medical statistics show significant variability between individuals. The effects of a medicine or the impact of several explanatory variables. This heterogeneity is known as variability and is frequently considered a significant source of variability in medical and biological contexts. This chapter addresses heterogeneity in survival analysis. This heterogeneity may be challenging to evaluate; yet, it remains very important. In recent decades, numerous publications on 'frailty models' have come out. The basic concept of these models is that individuals possess different levels of frailty, with the most frail individuals experiencing earlier mortality than those with lesser frailty. As a result, the systematic selection of robust individuals takes place, hence impacting the observed outcomes. When estimating death rates, one may be interested in their variations over time or age. Frequently, they increase at the start of the observation period, reach a peak, and then decrease (unimodal intensity) or remain stable. This illustrates the usual death rates of cancer patients, indicating that the longer a patient survives beyond a specific duration, the greater their chance of survival becomes. Unimodal intensities are likely the result of selection rather than indicative of individual-level development. The total population begins to decrease

as a result of the mortality of high-risk individuals. The intensity of a specific individual might keep increasing.

If covariates are identified, they may be incorporated into the study, such as by the application of the proportional hazards model previously mentioned. However, it is frequently impossible to incorporate all significant risk factors, possibly due to a lack of individual-level information (this is particularly true for population research, where the only known risk factors are commonly sex and age). Moreover, we may be unaware of the significance of the risk factor or even its existence. In certain situations, measuring the risk factor may be difficult without significant resources or time investment. In such circumstances, two sources of variability in duration data need consideration: variability due to observable risk factors (which is theoretically predictable) and heterogeneity resulting from unknown covariates, making it theoretically unpredictable even when all pertinent information is available at that time. The latter is of particular interest in this context, while the topic of observable covariates is addressed mainly for the sake of comprehensively. Hougaard (1991) [80] mentioned that there are benefits to analyzing these two types of variability independently: heterogeneity explains some 'unexpected' outcomes or provides an alternate interpretation of specific results, such as non-proportional or diminishing risks. If certain individuals present a greater risk of failure, the other persons at risk usually constitute a relatively selected group with lower risk. An estimate of the individual hazard rate that neglects unobserved frailty will progressively underestimate the hazard function over time.

To identify such selection effects, mixture models may be used. The population is presumed to consist of individuals with different, at least partially unknown risks. The unobservable risks are characterized by the mixture variable known as frailty. It is a random variable that follows a specific distribution. Moreover, since the value is unknown, it necessitates the process of marginalization. The specific attribute of the relationship between individual and population aging is influenced by the distribution of frailty across individuals. A multitude of distributions pertaining to the unobserved covariates may be examined, including binary, gamma, and log-normal, which expose disparities of both

a qualitative and quantitative character. Specifically, the variance associated with the frailty distribution serves as a metric for assessing the degree of heterogeneity present within the study population.

To deal with the issue of heterogeneity in a population caused by unobserved factors, Vaupel et al. (1979) [160] proposed a random effects model for lifetimes. They presented the concept of frailty and used it in demographic data analysis.

3.1 Mathematical Foundations of Frailty Models

The traditional and widely used frailty model asserts a proportional hazards framework that is influenced by the random effect referred to as frailty. The risk faced by an individual is further impacted by an unobservable, age-independent random variable Z, which multiplicatively affects the baseline hazard function λ_0.

$$\lambda(t,Z) = Z\lambda_0(t) \tag{3.1}$$

In this context, Z appears as a random mixture variable that fluctuates within the population. It is important to recognize that a scale factor applicable to all individuals in the population may be incorporated into the baseline hazard function $\lambda_0(t)$, therefore normalizing frailty distributions to $\mathbf{E}Z = 1$. The variance parameter $\sigma^2 = \mathbf{V}(Z)$ serves as an indicator of variability within the population regarding baseline risk. When the variance, denoted as σ^2, reaches its minimum threshold, the corresponding values of Z exhibit a high degree of concentration around the value of one. Conversely, in scenarios where σ^2 is elevated, the resultant values of Z show higher dispersion, resulting in greater heterogeneity in the individual hazards $Z\lambda_0(t)$. Frailty raises the individual's risk and may also be referred to as liability or susceptibility in different contexts. All individuals, except for a specific constant individual Z, are presumed to conform to an identical mortality pattern. What can be observed within a population is not the hazard for an individual, but rather the overall result for a group of individuals having varying values of the random

variable Z.

The following analysis demonstrates the validity of model (3.1). Let $\lambda(t,Z)$ denote an individual hazard. Through Taylor series expansion

$$\lambda(t,Z) = \lambda(t,0) + Z\lambda'(t,0) + o(Z) \tag{3.2}$$

where $o(Z)$ indicates the terms involving Z of order greater than one. By excluding these terms, we obtain

$$\lambda(t,Z) \approx \lambda(t,0) + Z\lambda'(t,0) \tag{3.3}$$

Assuming that zero frailty (susceptibility) results in zero mortality (excluding background mortality), it follows that

$$\lambda(t,Z) \approx Z\lambda_0(t) \tag{3.4}$$

where $\lambda_0(t) = \lambda'(t,0) = \frac{\partial\lambda(t,z)}{\partial z}\Big|_{z=0}$. From this, it follows that the foundational risk function $\lambda_0(t)$ represents the partial derivative of the individual hazard function while taking into account frailty evaluated at the point $Z = 0$.

A multiplicative frailty model, as presented in (3.1), clearly offers a simplified perspective on the influence of heterogeneity. Nonetheless, basic mathematical models provide a method to analyze the consequences associated with variability. The assertions that frailty operates independently of chronological age and that it exerts a multiplicative effect on the baseline hazard function are essentially arbitrary; yet, they have served as the foundation for extensive following research on unobserved heterogeneity in survival analysis.

We will include additional frailty models that do not rely on the proportional hazards assumption only for the sake of completeness. For instance, the additive frailty model, in which the frailty has an additive effect on the baseline hazard function. For further information regarding this unusual case, refer to Rocha (1996) [144]. The investigation of proportional odds frailty models is conducted in Lam et al. (2002) [100] and Lam and Lee

(2004) [101]. Murphy et al. (1997) [129] elucidated the association that exists between proportional odds models and frailty models. Prior studies concerning Accelerated Failure Time (AFT) frailty models encompass research conducted by Anderson and Louis (1995) [14], Keiding et al. (1997) [92], Klein et al. (1999) [95], Schnier et al. (2004) [148], and Chang (2004) [27].

In a conventional manner, it is feasible to integrate known covariates within the model (3.1):

$$\lambda(t,Z,X) = Z\lambda_0(t)e^{\Sigma_{i=1}^{k}\beta_i X_i} = Z\lambda_0(t)e^{\beta^T X} \tag{3.5}$$

Let $X = (X_1,\ldots,X_k)$ denote the covariates and $\beta = (\beta_1,\ldots,\beta_k)$ represent the regression parameters. Thus, a frailty model serves as a generalization of the established proportional hazards model. Let $S(t \mid Z)$ denote the survival function of an individual reliant on the frailty variable Z, which signifies

$$S(t \mid Z) = e^{-\int_0^t \lambda(s,Z)ds} = e^{-Z\int_0^t \lambda_0(s)ds} = e^{-Z\Lambda_0(t)} \tag{3.6}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(s)\,ds$ represents the cumulative baseline hazard function.

3.2 Unconditional Survival and Hazard Functions

At this point, the model has been defined at the individual level. This specific model is not observable. Therefore, it is imperative to assess the population level comprehensively. The survival function pertinent to the overall population is determined as the mean of individual survival functions, contingent upon the frailty distribution. The concept can be interpreted as the survival function of a randomly chosen individual. The survival and density functions, in conjunction with the mean and variance of the frailty distribution, are characterized by the Laplace transform associated with the frailty distribution.

$$S(t) = \mathbf{E}S(t \mid Z) = \mathbf{E}e^{-Z\Lambda_0(t)} = \mathbf{L}(\Lambda_0(t)) \tag{3.7}$$

$$f(t) = -\lambda_0(t)\mathbf{L}'(\Lambda_0(t)) \tag{3.8}$$

$$\mathbf{E}Z = -\mathbf{L}'(0) \tag{3.9}$$

$$\mathbf{V}(Z) = \mathbf{L}''(0) - \big(\mathbf{L}'(0)\big)^2 \tag{3.10}$$

which uses (3.6) and accentuates the critical role of the Laplace transform in the context of frailty models. Hence, if the aforementioned reveal an explicit structure, executing this calculation is relatively easy. The association with the Laplace transform was initially identified and exploited by Hougaard (1984, 1986a, 1986b) [76] [77] [78]. Consequently, while determining distributions for the frailty variable Z, it is logical to use those with explicit Laplace transforms. This facilitates the application of standard maximum-likelihood techniques for parameter estimation.

Theorem 2 (Vaupel et al., 1979) [160]. Examine a frailty model articulated through equation (3.1). The hazard function of the population is delineated as $\lambda(t) = \frac{f(t)}{S(t)}$, which may generally be represented as $\lambda(t) = \mathbf{E}(\lambda(t,Z) \mid T > t)$, or more particuly,

$$\lambda(t) = \int_0^\infty \lambda(t,z) f(z \mid T > t)dz = \lambda_0(t) \int_0^\infty z f(z \mid T > t)dz \tag{3.11}$$

where $f(z \mid T > t)$ signifies the probability density function of frailty within a population of individuals who have survived until age t.

Proof: Beginning with relation (3.1), we obtain

$$\lambda(t,z) = \frac{f(t \mid z)}{S(t \mid z)} = z\lambda_0(x)$$

$$f(t \mid z) = z\lambda_0(t)S(t \mid z)$$

$$f(t,z) = z\lambda_0(t)S(t \mid z)f_Z(z)$$

$$f(t) = \lambda_0(t) \int_0^\infty zS(t \mid z)f_Z(z)dz$$

where f_Z denotes the probability density function of the frailty distribution. Therefore,

$$\lambda(t) = \frac{\lambda_0(t) \int_0^\infty zS(t \mid z) f_Z(z) dz}{S(t)}$$

Survival at age t implies a death age above t, hence it follows that

$$f(z, T > t) = f_Z(z) \int_t^\infty z\lambda_0(s) S(s \mid z) ds = f_Z(z) S(t \mid z)$$
$$f(z \mid T > t) = \frac{f_Z(z) S(t \mid z)}{S(t)}$$

This concludes the proof.

The population hazard is therefore defined as the average of individual hazards among the survivors, as seen in (3.11). Frail individuals presenting higher Z values are likely to fail first. Consequently, the average frailty of the surviving population $\int_0^\infty zf(z \mid t) dz$ will diminish with age. Thus, equation (3.11) indicates that the force of death for individuals increases faster than population; in this context, individuals'age more rapidly'than their population.

3.3 Different Types of Frailty Models

3.3.1 Univariate Frailty Models

Gamma Frailty Model

Numerous studies clarify the application of the gamma distribution as a composite distribution, illustrating notable contributions by Greenwood and Yule (1920) [59], Vaupel et al. (1979) [160], Congdon (1995 [32]), dos Santos et al. (1995) [38], and Hougaard (2000) [83]. From a numerical and methodical viewpoint, it exhibits a high degree of efficacy in aligning with failure data attributable to the ease of obtaining closed-form expressions for unconditional survival, cumulative density, and hazard functions. This arises from the inherent simplicity associated with the Laplace transform. This further elucidates the extensive application of this distribution in the majority of applications

published to date. It is a flexible distribution that assumes various forms as the parameter k varies: when $k = 1$, it corresponds with the well-known exponential distribution; as k increases, it adopts a bell-shaped curve close to a normal distribution. Although these advantages are significant, it is important to recognize that there is no biological justification that makes the gamma distribution more advantageous than alternative frailty distributions. Almost all arguments supporting the gamma distribution are founded on mathematical and computational considerations. The research presented by Abbring and van den Berg (2007) [7] highlights the rationale for employing gamma distributions in the context of frailty within the examination of time-to-event data. This was illustrated across an extensive range of both univariate and multivariate frailty models; the distribution of frailty within the population of survivors approximates a gamma distribution, provided that certain mild regularity conditions are satisfied.

Frailty cannot take negative values, and the gamma distribution, as well as the log-normal distribution, are regarded as some of the most frequently utilized distributions for the modelling of variables that exhibit inherent positivity. In addition, it seems that the assertion regarding the initial frailty at the start of follow-up, adhering to a gamma distribution, provides some interesting mathematical results.
including

- The frailty of survivors at any time t follows a gamma distribution with a shape parameter k equal to the value recorded either at birth or at the commencement of the study period, whereas the second parameter is now referred to as $\lambda + \Lambda_0(t)$, with $\Lambda_0(t)$ indicating the cumulative baseline hazard function.

- The frailty of individuals who die at any age t follows a gamma distribution, characterized by the same parameter $\lambda + \Lambda_0(t)$ as that of those who survive to t, but with a shape parameter of $k+1$.

The mean frailty among deaths at age t is $\frac{\lambda+1}{\lambda+\Lambda(t)}$, but among survivors at the same age, it is $\frac{\lambda}{\lambda+\Lambda(t)}$. This illustrates the selection through the mortality of high-risk people, such as those with high frailty values Z.

To ensure the model's identifiability, it is suitable to enforce the parameter constraint $k = \lambda$ for the gamma distribution, thereby yielding $\mathbf{E}Z = 1$. Let $\sigma^2 := \frac{1}{\lambda}$ denote the variance associated with the frailty variable. The unconditional hazard and survival functions are articulated as follows.

$$\lambda(t) = \frac{\lambda_0(t)}{1+\sigma^2\Lambda_0(t)} \text{ and } S(t) = \mathbf{L}(\Lambda_0(t)) = \left(1+\sigma^2\Lambda_0(t)\right)^{-\frac{1}{\sigma^2}} \tag{3.12}$$

The model was proposed by Vaupel et al. (1979) [160].

Positive Stable Frailty Model

A distribution is classified as positive stable if the appropriately normalized summation of n independent random variables originating from this distribution retains the identical distribution. The process of normalization is articulated as $n^{\frac{1}{\gamma}}$, with the index γ required to drop inside the interval $(0,1]$ to get a distribution of positive values. Despite the absence of explicit mathematical expressions for the probability density or survival function of a random variable exhibiting a positive stable distribution, the Laplace transform is characterized by a mathematical formulation that is expressed in closed form:

$$\mathbf{L}(s) = e^{-\frac{ks^\gamma}{\gamma}} \tag{3.13}$$

To ensure identifiability, we limit the two-parameter frailty distribution to the scenario where $k = \gamma$. As a result,

$$S(t) = \mathbf{L}(\Lambda_0(t)) = e^{-\Lambda_0(t)^\gamma} \tag{3.14}$$

$$f(t) = \gamma\lambda_0(t)\Lambda_0(t)^{\gamma-1}e^{-\Lambda_0(t)^\gamma} \tag{3.15}$$

$$\lambda(t) = \gamma\lambda_0(t)\Lambda_0(t)^{\gamma-1} \tag{3.16}$$

Hougaard (1986b) [78] proposed this distribution as a frailty distribution, which was

thereafter applied by Wang et al. (1995) [161] and Manatunga and Oakes (1999) [115]. Within the framework of the shared positive stable frailty model, Fine et al. (2003) [48], and Martinussen and Pipper (2005) [119] developed innovative estimation approaches. It was further enhanced through the utilization of Hougaard's power variance function distribution (Hougaard, 1986a) [77] as well as Aalen's compound Poisson distribution (Aalen, 1988, 1992) [3] [4]. Every moment associated with this distribution exhibits an infinite property. This conclusion is critically significant in the context of the identifiability challenges outlined by Elbers and Ridder (1982) [45]. It was established that the existence of a finite mean within the frailty distribution constitutes one requisite condition essential for the identifiability of univariate frailty models. This was the primary objective underlying the initiation of the positive stable distribution as a frailty distribution. Particularly in bivariate and multivariate applications, significant focus has been given on resolving confounding issues within shared (gamma) frailty models.

The positive stable model suggests several interesting characteristics, as illustrated by the associated instances. A principal characteristic of the positive stable distribution is its distinctive role as the only frailty distribution that preserves the proportional hazards property in the unconditional hazard functions subsequent to the integration of the frailty variable.

Inverse Gaussian Frailty Model

Initially introduced by Hougaard (1984) [76] as a variant for the gamma distribution, the inverse Gaussian distribution was thereafter employed by Manton et al. (1986) [117], Klein et al. (1992) [94], Keiding et al. (1997) [92], and Price and Manatunga (2001) [139]. The probability density function associated with a random variable that adheres to an Inverse Gaussian distribution, characterized by a mean of one and a variance σ^2, is

$$f(z) = \frac{1}{\sqrt{2\pi\sigma^2 z^3}} e^{-\frac{1}{2\sigma^2 z}(z-1)^2} \tag{3.17}$$

Thus, the Laplace transform of the inverse normal distribution is expressed as

$$\mathbf{L}(s) = e^{\frac{1}{\sigma^2}\left(1-\sqrt{1+2\sigma^2 s}\right)} \tag{3.18}$$

Consequently, the conditional survival and hazard functions are posited to exhibit the following formulations

$$S(t) = e^{\frac{1}{\sigma^2}\left(1-\sqrt{1+2\sigma^2\Lambda_0(t)}\right)} \tag{3.19}$$

and

$$\lambda(t) = \frac{\lambda_0(t)}{(1+2\sigma^2\Lambda_0(t))^{1/2}} \tag{3.20}$$

PVF Frailty Model

A generalized class of frailty models, which includes gamma, inverse Gaussian, and positive stable distributions, constitutes the family of power variance function distributions, proposed by Tweedy (1984) [158] and further developed separately by Hougaard (1986a) [77]. This represents a three-parameter family referred to as $PVF(\gamma,k,\lambda)$. The Laplace transform is

$$\mathbf{L}(s) = e^{-\frac{k}{\gamma}((\lambda+s)^\gamma-\lambda^\gamma)} \tag{3.21}$$

The expectation and variance of a random variable Z distributed according to a PVF are

$$\mathbf{E}Z = k\lambda^{\gamma-1} \text{ and } \mathbf{V}(Z) = k(1-\gamma)\lambda^{\gamma-2} \tag{3.22}$$

Thus, the survival function is written as

$$S(t) = e^{-\frac{k}{\gamma}\left((\lambda+\Lambda_0(t))^\gamma-\lambda^\gamma\right)} \tag{3.23}$$

and the non-conditional hazard function is articulated as

$$\lambda(t) = k\lambda_0(t)\,(\lambda+\Lambda_0(t))^{\gamma-1} \tag{3.24}$$

Considering the constraint $\mathbf{E}Z = 1$ and the relation $\sigma^2 = \frac{1-\gamma}{\lambda}$ (see (3.22)), it follows that

$$\lambda(t) = \frac{\lambda_0(t)}{\left(1 + \frac{\sigma^2}{1-\gamma}\Lambda_0(t)\right)^{1-\gamma}}. \tag{3.25}$$

For specific parameter values, the analytic expression referenced earlier lacks immediate clarity in its definition; however, it ought to be comprehensively established through the principle of continuity. If $\gamma = 0$, the gamma distributions denoted as $\Gamma(k,\lambda)$ are formulated utilising an identical parameterisation. In the case where $\gamma = 0.5$, the inverse Gaussian distributions are established, and when $\lambda = 0$, the positive stable distributions are recognised. The parameter set is characterized by the constraints $0 \leq \gamma \leq 1$, $k > 0$, with $\lambda \geq 0$ applicable when $\gamma > 0$, and $\lambda > 0$ when $\gamma \geq 0$. The distribution of individuals surviving at age t is denoted as $PVF(\gamma, k, \lambda + \Lambda_0(t))$. An examination of the lifetimes of Danish twins was conducted by Hougaard et al. (1992) [81] employing this particular model.

Compound Poisson Frailty Model

The conceptual framework of frailty modelling utilising the compound Poisson distribution was initially established by Aalen(1988,1992) [3] [4]. A significant feature of the model is its prediction of a subgroup showing zero frailty, which illustrates long-term survival. This model is useful in the fields of medicine and demography. Although the continuous part's density is expressed only as an infinite series requiring numerical computation, the distribution exhibits significant mathematical characteristics. This distribution arises as the sum of a random number of independent, identically distributed gamma random variables, where the number of terms follows a Poisson distribution. Under this interpretation, individuals are subject to a randomly determined number of hits, each having a random size, which together define a hit model.

$$\mathrm{Z} = \begin{cases} X_1 + X_2 + \ldots + X_N & \text{if } N > 0 \\ 0 & \text{if } N = 0 \end{cases} \tag{3.26}$$

Let N be a Poisson random variable with expectation ρ, and let the sequence of independent random variables $X_1, X_2, \ldots$ each conform to a gamma distribution, specifically, where $X_i \sim \Gamma(k, \lambda)$. The Laplace transforms corresponding to the gamma and Poisson distributions are articulated as $\mathbf{L}_X(s) = \left(1 + \frac{s}{\lambda}\right)^{-k}$ and $\mathbf{L}_N(s) = e^{-\rho + \rho e^{-s}}$, respectively. The subsequent standard deviation may now be applied:

$$\mathbf{L}(s) = \mathbf{E}e^{-sZ} = \mathbf{E}e^{-s(X_1 + \ldots + X_N)} = \mathbf{E}\mathbf{L}_X(s)^N = \mathbf{L}_N\left(-\ln\left(\mathbf{L}_X(s)\right)\right) \tag{3.27}$$

Using the expressions given above yields the corresponding Laplace transform of Z:

$$\mathbf{L}(s) = e^{-\rho + \rho\left(1 + \frac{s}{\lambda}\right)^{-k}} \tag{3.28}$$

We will apply an alternative parameterization as follows:

$$\rho = -\frac{k\lambda^{\gamma}}{\gamma}, \quad \lambda = \lambda, \quad k = -\gamma \tag{3.29}$$

It naturally results that the Laplace transform pertinent to the compound Poisson distribution is

$$\mathbf{L}(s) = e^{-\frac{k}{\gamma}\left((\lambda + s)^{\gamma} - \lambda^{\gamma}\right)} \tag{3.30}$$

The parameter γ delineates the category of distributions into two predominant subclasses: For $\gamma \geq 0$, the distribution is categorized as a power variance function distribution (PVF). Aalen (1988) [3] proposed the extension to $\gamma < 0$, resulting in the compound Poisson distribution. The two subclasses are distinct by the gamma distribution ($\gamma = 0$). Aalen used an alternative parameterization. The notation $cP(\gamma, k, \lambda)$ denotes a compound Poisson distribution.

The implementation of the aforementioned Laplace transform produces the marginal survival and hazard functions within the framework of a compound Poisson frailty model:

$$S(t) = e^{-\frac{k}{\gamma}\left((\lambda+\Lambda_0(t))^{\gamma}-\lambda^{\gamma}\right)} \quad \text{and} \quad \lambda(t) = k\lambda_0(t)\left(\lambda+\Lambda_0(t)\right)^{\gamma-1} \tag{3.31}$$

This leads to

$$S(t) = e^{-\frac{1-\gamma}{\gamma\sigma^2}\left(\left(1+\frac{\sigma^2}{1-\gamma}\Lambda_0(t)\right)^{\gamma}-1\right)} \quad \text{and} \quad \lambda(t) = \frac{\lambda_0(t)}{\left(1+\frac{\sigma^2}{1-\gamma}\Lambda_0(t)\right)^{1-\gamma}} \tag{3.32}$$

In alignment with the methodological framework established by the PVF frailty model, we examine that the integral of $\lambda(t)$ (cumulative hazard function) over $[0,\infty)$ is finite for $\gamma < 0$. As a result, the survival function is inadequate since a portion of individuals have zero frailty and will never experience the event in question. Aalen (1992) [4] employed the model to analyze the incidence of marriage among women born in Denmark. Marriage illustrates an occurrence that does not occur for everyone. A specific fraction of individuals remain unmarried, requiring that models of marriage incidence incorporate this factor. The observed peak in incidence around age 23, followed by a decline after age 30, is thus recognised as a selection effect resulting from heterogeneity, indicating that individuals who are most likely to marry tend to do so at an earlier age, while the remaining population shows a decreased tendency to marry. A second application of Aalen (1992) [4] corresponds to fertility data among Norwegian women. It is a widely recognised fact that approximately 5% to 10% of all couples are incapable of conceiving children. Consequently, $\mathbf{P}(Z=0)$ (where Z signifies the frailty associated with the capacity to achieve conception) represents the probability of infertility, whereas the variability inherent in the frailty variable Z illustrates the disparate fertility rates across fertile couples.

Hougaard et al. (1994) [82] used the model to analyse data relating to the diagnosis of diabetic nephropathy, a severe complication experienced by certain individuals with diabetes. In an earlier publication, Aalen and Tretli (1999) [5] implemented the compound Poisson distribution to analyze testicular cancer. Testicular cancer presents two notable

epidemiological characteristics. Firstly, its incidence has risen significantly over the recent decades. Secondly, the frequency is highest among younger males and then decreases with advancing age. The concept of the model claims that a specific subgroup of males shows a higher risk of testicular cancer, leading to a process of natural selection over time. The model has been adapted to incidence data obtained from the Norwegian Cancer Registry, covering the period from 1953 to 1993. Building upon this foundational research, Moger et al. (2004) [124] engaged in further investigations, while Haukka et al. (2003) [68] performed an analysis of schizophrenia-related data concerning the Finnish birth cohort from 1950 to 1968, utilising the specified model. It was ascertained that a restricted segment of the population is vulnerable to schizophrenia, and an increasing individual risk was observed with advancing age within the population characterised by heightened susceptibility.

Log-normal Frailty Models

Log-normal frailty models exhibit significant advantages for the elucidation of dependence structures within multivariate frailty models, as evidenced by empirical investigations conducted by McGilchrist and Aisbett (1991) [121], McGilchrist (1993) [122], Lillard (1993) [108], Lillard et al. (1995) [109], Xue and Brookmeyer (1996) [166], Sastry (1997) [147], Gustafson (1997) [63], Ripatti and Palmgren (2000) [141], Ripatti et al. (2002) [142], and Huang and Wolfe (2002) [85].

In spite of this, the log-normal distribution has been utilized within univariate frameworks, as evidenced by the work of Flinn and Heckman (1982) [50]. There exist two distinct variations of the log-normal model. It is posited that a normally distributed random variable W produces frailty in the form $Z = e^W$. The two configurations of the model are delineated by the conditions $\mathbf{E}W = 0$ and $\mathbf{E}Z = 1$, with the first one being significantly more common in the literature. Unfortunately, an explicit form of the unconditional likelihood is unobtainable. Therefore, estimation methods employing numerical integration within the maximum likelihood method are necessary.

Cure Model with Univariate Frailty

The Cox proportional hazards model is commonly employed in the examination of survival time datasets. This model, along with frailty models that exclude the compound Poisson distribution, implicitly assumes that all individuals will experience the event of interest. In designated circumstances, an identifiable group of individuals is not predicted to encounter the event of interest; these individuals are perceived as cured. For instance, researchers may be involved in examining the incidence of recurrence associated with a particular disease. A significant number of individuals may never encounter a recurrence of that disease; therefore, a segment of the population includes those who have recovered from the illness. Historically, cure models have been applied to estimate the proportion of the population that has attained recovery. These models facilitate a deeper comprehension of time-to-event data while supporting the derivation of more precise inferences than those previously available. These conclusions would not be attainable through an analysis that does not consider a cured proportion within the population. In the absence of a cured fraction component, the analysis reduces to typical survival analysis approaches. Cure models claim that individuals subject to the event of interest present homogeneous risks. This section addresses extensions of cure models to handle heterogeneity within the at-risk population by applying frailty models. Alternatively, from a certain perspective, it examines extensions of frailty models to incorporate a cured fraction within the study population. In this instance, the characterisation of frailty encompasses a combination of both discrete and continuous probability distributions. For instance, in instances where the variable of interest is the consumption of alcohol, a notable fraction of individuals refrain from engaging in the act of drinking.

Spilerman (1972) [153] viewed the 'spiked-gamma' as a relevant illustration of this type of distribution. In the framework of cured models, the population is divided into two distinct sub-groups: an individual is rendered cured with a probability of $1-\phi$, or possesses a legitimate survival function $S_0(t)$ with a probability of ϕ. For individuals classified as cured, the event of interest will not occur, and their survival time is treated as

infinite. Consequently, the hazard and survival functions of cured persons are established as zero and one, respectively, for all finite values of t. A survival time model that includes a cured proportion is shown as

$$S^*(t) = (1-\phi) + \phi S(t) \tag{3.33}$$

The idea of frailty cure, or cure-mixture, models was introduced by Longini and Halloran (1996) [113] as an extension of standard frailty models. The random variable pertaining to frailty in the prior context possesses a point mass at zero with a probability of $1-\phi$, while the heterogeneity across individuals experiencing the event of interest is assumed to follow a continuous distribution with probability ϕ. The survival function under the gamma frailty cure model is delineated by

$$S^*(t) = (1-\phi) + \phi \left(1+\sigma^2 \Lambda_0(t)\right)^{-1/\sigma^2} \tag{3.34}$$

The concept underlying this model is related but distinguished from Aalen's compound Poisson frailty model (1988, 1992) [3] [4]. Price and Manatunga (2001) [139] proposed a comprehensive overview of this field and employed various cure, frailty, and cure frailty models in the examination of data pertaining to leukaemia remission outcomes. They deduce that frailty models are accurate in the representation of data characterized by a cured fraction and note that the gamma frailty cure model provides a superior fit to their remission data when compared to the conventional cure model.

This illustration expands the model mentioned earlier to incorporate censored observations. Two manifestations of disease: incidence and age at which symptoms first present. Risk assessment models concerning overall vulnerability that exclusively consider the initial occurrence by categorizing the disease as a binary trait affected or unaffected may yield incorrect conclusions, as it is frequently indeterminate whether individuals without the disease will develop it due to censoring. In contrast, survival analysis models generally presume that all individuals possess identical susceptibility to the disease and will ulti-

mately be affected given an adequately extended observation period. It is possible that these models do not precisely characterize the risk factors associated with the disease. In models that incorporate both categories of expressions, the influence of a covariate may have implications for either the general susceptibility, the age of onset, or potentially both.

The application of mixture models for the integrative examination of aggregate disease risk in conjunction with the age-at-onset distribution within affected populations constitutes a well-established approach (Farewell 1977 [46]; Kuk and Chen 1992 [99]; Lam et al. 2005 [102]). Susceptibility is assigned to an individual if he or she will ultimately develop the disease after a long time of observation. Establish

$$\mathrm{Y} = \begin{cases} 1: & \text{if the individual is susceptible} \\ 0: & \text{if otherwise} \end{cases} \tag{3.35}$$

Let T be the age at which onset occurs when $Y = 1$. Let $\phi = \mathbf{P}(Y = 1)$ and $S(t) = \mathbf{P}(T > t \mid Y = 1)$ define the distribution of the variable Y and the associated failure time T. While Y is not subject to direct observation, it remains feasible to ascertain whether an individual has undergone the occurrence within the designated follow-up timeframe.

- For observations characterised by the occurrence of the event, it is established that $\Delta = 1$. Clearly, $Y = 1$ and the survival function for the uncensored data is expressed as $\mathbf{P}(Y = 1)\mathbf{P}(T \leq C \mid Y = 1) = \phi(1 - S(C))$, where C represents the censoring time.

- For the remaining observations, no failure is detected ($\Delta = 0$). This may happen either due to $Y = 0$ or because the observation is actually censored. Consequently, $\mathbf{P}(Y = 0) + \mathbf{P}(Y = 1)\mathbf{P}(T > C \mid Y = 1) = (1 - \phi) + \phi S(C)$.

Considering these results, the likelihood function assumes the following form.

$$L(t, \delta) = \delta\phi f(t) + (1 - \delta)(1 - \phi + \phi S(t)) \tag{3.36}$$

3.3.2 Multivariate Frailty Models

At this point, we have concentrated on the frailty model as a method for addressing potential heterogeneity resulting from unobserved factors. This represents the primary interpretation of frailty in relation to univariate time-to-event data. This leads to selection effects over time, illustrated by levelling-off or crossing-over effects in population risks.

A distinctive feature of that method is its application in modelling statistical dependence, as shown by Clayton (1978) [28]. Most statistical models and approaches for failure time data, the Cox proportional hazards model, in particular, were conceived under the premise that the observations obtained from individual subjects exhibit statistical independence. Despite the fact that this approach is appropriate for various applications, it has become clear that this assumption is invalid in alternative scenarios that are more common than originally expected. The subsequent three examples, as presented in Liang et al. (1995) [107], serve to illustrate this concept:

- Diabetic retinopathy is among the primary causes of loss of vision and blindness. Considering the relatively high incidence of this disease within the population, the development of novel interventions that prevent the onset of serious blindness is essential. In 1971, the Diabetic Retinopathy Study was initiated to evaluate the efficacy of laser photocoagulation. This randomised, controlled clinical trial included over 1,700 patients registered across 15 medical centres in the United States. Individuals diagnosed with diabetic retinopathy and showing a visual acuity of 20/100 or higher in both eyes were regarded as suitable for inclusion in the research. The experimental technique randomly selected one treated eye from each participant, while the opposite eye was observed without treatment. The incident was defined as visual acuity below 5/200 at two successive follow-ups executed four months apart. This strategy contrasts with traditional randomised trials by having each patient act as their own control. Thus, each patient provided two related observations for the study, one from each eye.

- Family studies are essential for evaluating the influence of genetics on the disease

process. Statistical techniques, such as variance component models and path analysis, have been formulated and utilised to examine familial data regarding quantitative features, such as cholesterol levels. Variance component models seek to quantify the proportion of overall variation in a quantitative trait attributable to familial correlations, including the degree of correlation among full siblings and other relatives. Typical approaches are inappropriate when the trait under consideration is the age of disease onset. This is primarily due to the censoring and truncation characteristics of this variable, as well as the necessity for a measure of within-family correlation that takes into consideration time, a trait lacking in standard measures like the correlation coefficient.

- The gamma interferon trial, as described by Fleming and Harrington (2013) [49], illustrates repeated events. The investigation analyzed the prevalence of significant infections in subjects diagnosed with chronic granulomatous disease (CGD) who were randomly assigned to receive either gamma interferon or a placebo. Infections are capable of recurring. Consequently, both the duration and frequency of infection occurrences may provide information about the treatment's efficacy.

These three cases possess a significant characteristic: the failure times of observations within the same cluster show correlation. In Examples 1 and 3, the cluster represents a person, whereas in Example 2, it denotes a family. Generally, the sizes of the clusters are small in comparison to the total number of clusters. These three examples, however, vary in their scientific purposes. The main goal of Examples 1 and 3 is to evaluate the efficacy of a novel treatment, which can presumably be described by regression modelling. In these instances, the within-cluster correlation is typically of minor importance; yet, neglecting it may result in incorrect results. The intra-family correlation for Example 2 is of major significance; nevertheless, regression correction for each related individual is essential to reduce the possibility that the observed correlation is principally related to shared environmental factors within the family. Additionally, the mechanisms underlying within-cluster associations may differ, necessitating the use of various statistical models

to precisely describe these associations. It is evident, for example, while a singular mechanism may elucidate the relationship observed between the two eyes of an individual, an alternative mechanism may govern the correlation of observations obtained from the same eye over an extended temporal framework. Multivariate survival analysis can serve as an effective method for extracting information from multiple or recurrent events in the scenarios discussed above. According to Wei and Glidden (1997) [163], statistical models are predominantly classified into two principal categories: marginal models and frailty models. Marginal analysis techniques delineate frameworks for assessing the impact of covariates on the risks associated with singular occurrences (the margins), while accounting for the correlation among recorded event durations without explicitly expressing this correlation (Wei et al. 1989 [162]; Lee et al. 1992 [105]; Cai and Prentice 1995 [26]). The dependence between events is treated as a nuisance parameter, while the marginal baseline hazards can be specified using different approaches (Wei et al. 1989) [162] or assumed to share a common functional form (Lee et al. 1992) [105]. In a manner analogous to the examination of longitudinal data, regression parameters are derived utilizing generalized estimating equations, and the corresponding variance-covariance estimators are appropriately adjusted to accommodate the dependence structure. An extensive examination of the comprehensive and thoroughly established marginal approach is presented in Lin (1994) [110].

The marginal method is optimal for estimating the population average effect of risk factors on failure time. Nonetheless, it offers little understanding of the multivariate correlation among failure times. These issues are addressed by frailty models, which explicitly account for the correlation among different events. Frailty models offer intuitive appeal and explain the relationship between failures.

Mahé and Chevret (1999) [114] offer a method that represents the intersection between the two previously discussed models, allowing for the estimation of regression coefficients using traditional interpretation along with correlations.

A prevalent and general method for modelling multivariate data is assuming conditional independence of observed items given unobserved factors. A multivariate model for

the observed data is obtained by integrating over an assumed distribution of the latent variables. The dependence framework within the multivariate model emerges when dependent latent variables are integrated into the conditional models for various observable data items, with the dependence parameters frequently interpretable as variance components. Frailty models for multivariate survival data are formulated based on a conditional independence assumption by including latent factors that have a multiplicative effect on the baseline hazard.

Let us now consider multivariate models characterised by dependent random hazards, as discussed above. This notion may be regarded as a multivariate extension of the classical univariate frailty model introduced by Vaupel et al. (1979) [160], permitting the assessment of the mutual dependence among the lifetimes of related individuals in the analysis of survival data. Survival models for correlated lifetimes are valuable because they permit the examination of more complex issues concerning mechanisms of ageing, disease advancement, and mortality.

The initial prominent methodology pertains to the notion of shared frailty. Within the framework of a shared frailty model, frailty is defined as an indicator of the comparative risk that individuals within a collective group mutually possess.

Consequently, the frailty variable corresponds to groups of individuals rather than to individuals themselves. The hazard model for each individual, however, matches the usual univariate frailty model:

$$\lambda(t,Z) = Z\lambda_0(t) \tag{3.37}$$

The theoretical framework posits that the occurrence of all failure times is independent, given the presence of the frailties. The lifetimes maintain a conditionally independent relationship. The frailty variable Z exhibits temporal invariance and is shared among the constituents of the group, thereby engendering dependence. This dependency is invariably positive. The conditional survival function within the context of the bivariate model is

$$S(t_1,t_2 \mid Z) = S(t_1 \mid Z)\,S(t_2 \mid Z) = e^{-Z\Lambda_0(t_1)}e^{-Z\Lambda_0(t_2)} \tag{3.38}$$

The correlated frailty model constitutes the second principal concept within the framework of multivariate frailty models. This serves as an expansion of the shared frailty model. In the correlated frailty model, members within a cohort share only specific components of the frailty Z. This framework facilitates the integration of an extra correlation parameter, thereby enabling an analysis of the genetic and environmental factors influencing personal frailty. The conditional survival function within the context of the bivariate framework is delineated as

$$S(t_1,t_2 \mid Z_1,Z_2) = S(t_1 \mid Z_1)S(t_2 \mid Z_2) = e^{-Z_1\Lambda_0(t_1)}e^{-Z_2\Lambda_0(t_2)} \tag{3.39}$$

where Z_1 and Z_2 represent two correlated random variables. The concept of conditional independence holds significant importance in the genetic analysis of survival. Within the boundaries of this assertion, the frailty variable is revealed as the only way to explain the genetic effect on longevity. The underlying hazard signifies only a non-genetic impact on lifetime. Univariate frailty models lack this characteristic.

In the scenario of a degenerated frailty Z, there is no dependence among the lifetimes inside a group. Various groups are regarded as independent. The number of members in a group is presumed to be known. We shall examine the bivariate scenario of individual pairs more deeply to clarify the fundamental concepts. Twin studies illustrate bivariate event data and will be examined in depth later. Another example is the time to failure for several similar human organs, such as the duration until blindness occurs in the right and left eyes, as observed in investigations on diabetic retinopathy. Higher-dimensional extensions are provided where needed.

In every instance, it is imperative to assign a distribution to the frailty variable. The presumption of a random frailty facilitates the incorporation of the frailty component within the formulations, thereby allowing the evaluation of multivariate survival times. Nearly all calculations can be conducted utilizing the Laplace transform of the associated frailty distribution.

Shared Frailty Model

A shared frailty model within the domain of survival analysis can be delineated in the following manner. Consider the existence of n clusters, wherein the i-th cluster comprises n_i individuals and is linked to an unobservable frailty denoted as Z_i, where $(1 \leq i \leq n)$. A vector $X_{ij}(1 \leq i \leq n, 1 \leq j \leq n_i)$ is representative of the ij-th comprehensive survival duration T_{ij} pertaining to the j-th subject positioned within the i-th cluster.Considering the inherent frailties Z_i, it is assumed that the survival durations are independent, with their hazard functions delineated by the subsequent formulation

$$\lambda(t) = Z_i \lambda_{0j}(t) e^{\beta^T X_{ij}} \tag{3.40}$$

The baseline hazard functions are denoted by $\lambda_{0j}(t)$, while β represents a vector of fixed effect parameters subject to estimation. The frailties Z_i are hypothesized to be identically and independently distributed random variables delineated by a common density function $f(z, \theta)$, wherein θ denotes the parameter associated with the frailty distribution. In this theoretical construct, $\lambda_{0j}(t)$, which represents the baseline hazard function, is addressed in a non-parametric manner within the semi-parametric shared frailty framework.

In the interest of preserving clarity, we confine our examination of frailty models to the bivariate scenario $(n_i = 2)$, given that extensions to the multivariate context are simple. The key idea of a shared frailty model is that both persons in a pair possess the same frailty Z, which is the reason behind its designation as the shared frailty model. It was presented by Clayton (1978) [28], who did not employ the concept of 'frailty', and was thoroughly examined in the works of Hougaard (2000) [83], Therneau and Grambsch (2000) [156], Duchateau et al. (2002, 2003) [39] [40], and Duchateau and Janssen (2004) [41]. The two lifetimes are considered to exhibit conditional independence, predicated upon the existence of a common frailty. We obtain the quantities based on the conditional expression presented below. Conditional on Z, the hazard function for an individual in a

group is expressed as $Z\lambda_0(t)$, where the value of Z is shared by each individual in the group, therefore creating dependence between their lifetimes. The independence of lifetimes within a cluster is reflective of a degenerate frailty distribution, signifying an absence of variability in Z. In all alternative scenarios, the dependence is positive. Independence is presumed among different pairs. If $\mathbf{P}(Z > 0) = 1$ is true, the shared frailty model results in absolutely continuous distributions and so cannot capture dependency due to common events. Therefore, it is inadequate for event-related dependence (shock models), as an event affecting one individual is not significant to the partner; it just modifies information related to the frailty.

The bivariate survival function can be derived. Given the condition on Z, it is

$$S(t_1,t_2 \mid Z) = S_1(t_1)^Z S_2(t_2)^Z = e^{-Z\Lambda_{01}(t_1)} e^{-Z\Lambda_{02}(t_2)} = e^{-Z(\Lambda_{01}(t_1)+\Lambda_{02}(t_2))} \tag{3.41}$$

where $\Lambda_{0i}(t) = \int_0^t \lambda_{0i}(s)ds, i = 1,2$ represent the cumulative baseline hazard functions. Applying (3.41) considering Z yields the marginal bivariate survival function.

$$\begin{aligned} S(t_1,t_2) &= \mathbf{E}S(t_1,t_2 \mid Z) \\ &= \mathbf{E}S_1(t_1)^Z S_2(t_2)^Z \\ &= \mathbf{E}e^{-Z(\Lambda_{01}(t_1)+\Lambda_{02}(t_2))} \\ &= \mathbf{L}(\Lambda_{01}(t_1)+\Lambda_{02}(t_2)) \end{aligned} \tag{3.42}$$

where $\mathbf{L}$ denotes the Laplace transform corresponding to the random variable Z. The bivariate survival function is characterized as the Laplace transform of the frailty distribution, evaluated at the cumulative baseline hazard. The subsequent assertion is considered accurate concerning the marginal survival functions:

$$S_i(t_i) = \mathbf{E}S_i(t_i \mid Z) = \mathbf{E}S_i(t_i)^Z = \mathbf{E}e^{-Z(\Lambda_{0i}(t_i))} = \mathbf{L}(\Lambda_{0i}(t_i)) = p(\Lambda_{0i}(t_i)) \tag{3.43}$$

Let $p = \mathbf{L}$ and $i = 1,2$. Thus, $\Lambda_{0i}(t_i) = q(S_i(t_i))$, where q denotes the inverse function associated with p, and the bivariate unconditional survival function is articulated as

$$S(t_1,t_2) = p(q(S_1(t_1)) + q(S_2(t_2))) \tag{3.44}$$

The Archimedean copula family proposed by Genest and MacKay (1986) [51], in which p is a function that is twice differentiable, fullfils the condition $p(0) = 1$, and exhibits the characteristics that $p'(\cdot) < 0$ and $p''(\cdot) > 0$.

The conventional assumption pertaining to the distribution of frailty is that it adheres to a gamma distribution characterized by a mean of 1 and a variance of σ^2. By averaging equation (3.41) in consideration of Z, one derives the marginal bivariate survival function.

$$\begin{aligned} S(t_1,t_2) &= \mathbf{L}(\Lambda_{01}(t_1) + \Lambda_{02}(t_2)) \\ &= \left(1 + \sigma^2(\Lambda_{01}(t_1) + \Lambda_{02}(t_2))\right)^{-1/\sigma^2} \\ &= \left(S_1(t_1)^{-\sigma^2} + S_2(t_2)^{-\sigma^2} - 1\right)^{-1/\sigma^2} \end{aligned} \tag{3.45}$$

where the final relationship is derived from equation (3.12). The notion of shared frailty diverges from the characterization of individual frailty posited by Vaupel et al. (1979) [160] in their analysis of univariate time data. This distinction has been widely neglected, maybe due to the evident resemblance of the distinct risks associated with both methodologies. The frailty component within the bivariate shared frailty model represents only a fraction of the individual frailty, encompassing just those features of frailty that are prevalent among both individuals.

Clayton (1978) [28], Cox and Oakes (1984) [36], and Yashin and Iachine (1999) [170] elucidated that the bivariate survival function articulated within the framework of the shared gamma frailty model (3.45) can also be obtained by a different methodology. Let the dependent life spans be characterized by T_1 and T_2. Consider the bivariate survival function as $S(t_1,t_2) = \mathbf{P}(T_1 > t_1, T_2 > t_2)$, which exhibits absolute continuity and possesses

marginal survival functions $S_1(t_1) = S(t_1, 0)$ and $S_2(t_2) = S(0, t_2)$.

Consequently, the conditional survival function of T_1 given the condition that $T_2 > t_2$ is $S(t_1 \mid T_2 > t_2) = \frac{S(t_1,t_2)}{S(t_2)}$ and that of T_1 given $T_2 = t_2$ is

$$S(t_1 \mid T_2 = t_2) = \frac{\frac{\partial S(t_1,t_2)}{\partial t_2}}{\frac{\partial S(t_2)}{\partial t_2}} \tag{3.46}$$

The associated conditional hazards hold significant relevance for the subsequent analyses. The utilization of the relationship $\lambda(t) = -\frac{S'(t)}{S(t)}$ suggests

$$\lambda(t_1 \mid T_2 > t_2) = -\frac{\partial}{\partial t_1} \ln(S(t_1, t_2)) \tag{3.47}$$

and

$$\lambda(t_1 \mid T_2 = t_2) = -\frac{\partial}{\partial t_1} \ln\left(-\frac{\partial}{\partial t_2} S(t_1, t_2)\right) \tag{3.48}$$

These hazards represent the probability of encountering failure at age t_i for the i-th individual, dependent upon the state of the second individual. The initial hazard (3.47) relies on the condition $\{T_j > t_j\}$, while the following hazard (3.48) is dependent on $\{T_j = t_j\}$. Oakes (1989) [135] used a deviation of the hazard ratio from 1 as an indicator of the reciprocal dependency of the corresponding marginal lifetimes. The shared gamma frailty model may now be established based on the introduction of the subsequent relationship among the previous hazards:

$$\lambda(t_1 \mid T_2 = t_2) = \left(1 + \sigma^2\right) \lambda(t_1 \mid T_2 > t_2) \tag{3.49}$$

It is evident that (3.49) is equal to a corresponding condition with the roles of T_1 and T_2 reversed. This relationship uniquely defines the bivariate survival function (3.45), according to the marginal distributions.

$$
\begin{aligned}
\lambda\left(t_1 \mid T_2 = t_2\right) &= \left(1+\sigma^2\right)\lambda\left(t_1 \mid T_2 > t_2\right) \\
\frac{\partial}{\partial t_1}\ln\left(-\frac{\partial}{\partial t_2}S\left(t_1,t_2\right)\right) &= \left(1+\sigma^2\right)\frac{\partial}{\partial t_1}\ln\left(S\left(t_1,t_2\right)\right) \\
\int_0^{t_1}\frac{\partial}{\partial t}\ln\left(-\frac{\partial}{\partial t_2}S\left(t,t_2\right)\right)dt &= \left(1+\sigma^2\right)\int_0^{t_1}\frac{\partial}{\partial t}\ln\left(S\left(t,t_2\right)\right)dt \\
\ln\left(-\frac{\partial}{\partial t_2}S\left(t_1,t_2\right)\right) - \ln\left(-\frac{\partial}{\partial t_2}S_2\left(t_2\right)\right) &= \left(1+\sigma^2\right)\left(\ln\left(S\left(t_1,t_2\right)\right) - \ln\left(S_2\left(t_2\right)\right)\right) \\
\ln\left(-\frac{\partial}{\partial t_2}S\left(t_1,t_2\right)\right) - \left(1+\sigma^2\right)\ln\left(S\left(t_1,t_2\right)\right) &= \ln\left(-\frac{\partial}{\partial t_2}S_2\left(t_2\right)\right) - \left(1+\sigma^2\right)\ln\left(S_2\left(t_2\right)\right) \\
\frac{\frac{\partial}{\partial t_2}S\left(t_1,t_2\right)}{S\left(t_1,t_2\right)^{1+\sigma^2}} &= \frac{\frac{\partial}{\partial t_2}S_2\left(t_2\right)}{S_2\left(t_2\right)^{1+\sigma^2}} \\
\int_0^{t_2}\frac{\frac{\partial}{\partial t}S\left(t_1,t\right)}{S\left(t_1,t\right)^{1+\sigma^2}}dt &= \int_0^{t_2}\frac{\frac{\partial}{\partial t}S_2(t)}{S_2(t)^{1+\sigma^2}}dt \\
S\left(t_1,t_2\right)^{-\sigma^2} - S_1\left(t_1\right)^{-\sigma^2} &= S_2\left(t_2\right)^{-\sigma^2} - 1 \\
S\left(t_1,t_2\right) &= \left(S_1\left(t_1\right)^{-\sigma^2} + S_2\left(t_2\right)^{-\sigma^2} - 1\right)^{-\frac{1}{\sigma^2}} \qquad (3.50)
\end{aligned}
$$

The benefit of model (3.49) lies in the clear understanding of $\left(1+\sigma^2\right)$ as the relative risk linked to a non-surviving relative. Consequently, there are two methods for deriving (3.45) based on fundamentally distinct concepts: one employs the assumption (3.49), which posits proportional conditional risks, whereas the latter incorporates random hazards through a gamma-distributed shared frailty.

The bivariate shared frailty model may be extended to encompass the multivariate context, involving p linked failure times, yielding the unconditional multivariate survival function for gamma distributed frailty as follows:

$$
S\left(t_1,\ldots,t_p\right) = \left(\sum_{i=1}^{p} S_i\left(t_i\right)^{-\sigma^2} - p + 1\right)^{-1/\sigma^2} \qquad (3.51)
$$

which was demonstrated by Cook and Johnson (1981) [33]. Guo and Rodriguez (1992)

[60], later advanced by Guo (1993) [61], refined the characterisation (3.49) to integrate the multivariate context. The dependence between lifetimes among randomly selected pairs remains constant; this constraint diminishes the model's efficacy in scrutinising correlations within familial research that encompasses a diverse array of relational dyads, including parent–child, grandparent–grandchild, and sibling associations.

In the context of the shared gamma frailty model incorporating observed covariates, the frailty term $Z_i, (i = 1, \ldots, n)$ specific to each group can be estimated (Nielsen et al., 1992) [131] through the following methodology

$$\hat{Z}_i = \frac{1/\sigma^2 + \sum_{j=1}^{n_i} \delta_{ij}}{1/\sigma^2 + \sum_{j=1}^{n_i} e^{\beta X_{ij}} \Lambda\left(t_{ij}\right)} \tag{3.52}$$

This is attainable due to the repetitive observations within each group, where all observations in a given group are derived from the same value of the frailty variable. Employing the methodology delineated previously, Carvalho et al. (2003) [155] conducted an examination of the quality metrics associated with dialysis centres in Brazil, wherein centre-specific frailties served as indicators of unobserved heterogeneity.

The asymptotic properties of the nonparametric maximum likelihood estimates within the shared gamma frailty model are comprehensively elucidated. Murphy elucidates the properties of consistency (Murphy, 1994) [127] and asymptotic normality (Murphy, 1995) [128] in the context of the shared gamma frailty model devoid of covariates. These results were subsequently generalized to encompass the correlated gamma frailty model with the inclusion of observed covariates by Parner (1998) [137]. Shih and Louis (1995) [152] introduced a graphical approach for evaluating the assumption of a gamma distribution when the baseline hazard is parametric and covariates are absent. Glidden (1999) [53] proposed a test for the gamma frailty model that does not involve parameterising the baseline hazard or including covariates in the model. Cui and Sun (2004) [37] propose both graphical and numerical approaches for evaluating the proper fit of the gamma distribution within a shared frailty model. The examination they conducted is predicated upon the posterior expectation of the frailties derived from the observable data across

temporal times, thereby extending the research contributions of Glidden (1999) [53].

Restrictions of Shared Frailty Models

Shared frailty refers to the statistical associations observed within clusters, wherein a cluster may encompass individuals belonging to the same cohort, such as a family unit, a litter, a clinical setting, or a community, or may pertain to multiple or recurring events originating from the same individual. Nonetheless, it possesses certain limitations. In the following, we employ several arguments presented by Xue and Brookmeyer (1996) [166].

At first, it assumes that the unobserved factors are identical within the cluster, which is typically an inappropriate assumption. For instance, it is typically considered unsuitable to presuppose that both individuals within a twin pair possess identical unknown risk factors, especially in the context of complex, multi-stage diseases. The variability is generally greater during the initial stages, as individuals advance into the later stages of their development, there is a tendency for them to display heightened homogeneity.

Second, in the majority of circumstances, the context of shared frailty will predominantly produce positive correlations within the specified cluster (for notable exceptions, refer to Joe 1993) [90]. Despite this, in specific instances, the durations of survival of subjects within the same group show a negative correlation. For instance, the growth rates of animals sharing the same litter and exposed to poor food resources are likely to be negatively correlated. Another example is transplantation research has elucidated that, as a general principle, an extended duration of waiting for a transplant correlates with a diminished probability of survival following the procedure. Therefore, the duration of waiting and survival time may be oppositely associated. As a further example, consider a patient who is repeatedly hospitalised for the same illness. Patients exhibiting more severe conditions demonstrate a heightened risk of readmission alongside a diminished probability of successful discharge. Consequently, the durations of stay within the hospital and outside the hospital are anticipated to be negatively correlated. An additional example is illustrated by competing risk scenarios, in which a reduction in the risk of

mortality from one disease results in a greater risk of death from another disease.

Third, the interdependence of survival durations within the cohort is represented through a model that utilises the marginal distributions of the survival durations. To elucidate this, it is essential to recognise that the integration of covariates into a proportional hazards model characterised by gamma-distributed frailty results in a combination of the dependence parameter with population heterogeneity (Clayton and Cuzick 1985) [29], implying that the joint probability distribution can be separated from the marginal probability distributions (Hougaard 1986b) [77]. Elbers and Ridder (1982) [45] elucidate that this concern arises for any univariate frailty distribution possessing a finite mean. Nonetheless, shared frailty in bivariate models varies with individual frailty employed in univariate data analysis. The initial distinction between the concepts of frailty was not well comprehended. The value of σ^2 derived from the univariate data may be unrelated to association. Consider two hypothetical bivariate data sets showing distinct relationships between the life durations of related individuals, specifically the lifetimes of men and their sons, as well as those men and their grandsons. Across both experimental conditions, the parameter, denoted as σ^2, is assessed to be equivalent, utilising the dataset pertaining to the male participants (specifically, grandfathers), regardless of the varying correlations observed between the durations of existence of grandfathers in relation to their grandsons, as well as those that pertain to the relationships between grandfathers and their sons. This last and potentially biggest limitation of shared frailty models comes from the issue of identifiability within the univariate frailty model when covariates are observed. Therefore, it is a fundamental feature of all shared frailty models possessing a finite mean of the frailty distribution. To address this issue, Hougaard (1986 [77], 1987 [79]) proposes the shared positive stable frailty model. In this situation, the univariate model that includes observed covariates lacks identifiability as a result of the characteristics of the mean within the positive stable distribution being infinite. Therefore, one can expect greater flexibility from a shared frailty model characterised by a positive stable distribution compared to models utilising gamma frailty. The bivariate survival function encompassed within the shared positive stable frailty model is

$$S(t_1,t_2) = \exp\left\{-\left((-\ln(S_1(t_1)))^{1/\gamma} + (-\ln(S_2(t_2)))^{1/\gamma}\right)^{\gamma}\right\} \tag{3.53}$$

Upon the observation of covariates $X_i, (i = 1,2)$, and under the assumption that the conditional hazard is specified by equation (3.5), the univariate survival functions are subsequently delineated as

$$S_i(t) = \exp\left\{-\Lambda_i^{\gamma}(t)e^{\gamma\beta_i X_i}\right\} = \exp\left\{-\Lambda_i^{*}(t)e^{\beta_i^* X_i}\right\} \tag{3.54}$$

where $\beta_i^* = \gamma\beta_i$ and $\Lambda_i^*(t) = \Lambda_i^{\gamma}(t)$. Implementing a shared positive stable frailty model permits the estimation of both the association parameter γ and the regression coefficients β_i derived from the dataset belonging to related people. A persistent issue remains: the interpretation of regression parameters β_i. To demonstrate this issue, consider $\beta = \beta_1 = \beta_2$ for two categories of relatives, such as monozygotic (MZ) and dizygotic (DZ) twins. The association parameter γ distinctly varies between **MZ** and **DZ** twins. Consequently, the values of the parameter $\beta^* = \gamma\beta$ and $H^*(t)$ in equation (3.53) exhibit variations when contrasting MZ and DZ twins, contravening the basic assertion that the duration of such individuals, based on observable indicators, follows the identical Cox model. Assuming the parameters $\beta^* = \gamma\beta$ are identical for both **MZ** and **DZ** twins suggests that the coefficients β and the baseline hazard functions $\lambda_0(t)$ must exhibit variation among these subjects, therefore, complicating the interpretation of the conditional hazard (3.5) for this model.

To prevent such complications, correlated frailty models have been advanced for the examination of multivariate failure time data. These theoretical frameworks incorporate two interrelated random variables that delineate the frailty effect pertinent to each cluster. As an illustration, a single random variable is given to twin 1 and another to twin 2, thereby removing the constraint of common frailty. The two variables are linked and possess a joint distribution; therefore, knowledge of one does not necessarily indicate information about the other. Additionally, the two variables may indeed be negatively

related, which would result in a negative association between survival times.

There is an intermediate approach between shared and correlated frailty models (often referred to as univariate and bivariate frailty models, respectively, based on the frailty dimension), which addresses bivariate time-to-event data. In this approach, two variables are given to each cluster to account for heterogeneity, yet both are derived from a single common random variable. For instance, a bivariate survival time (T_1, T_2) incorporates Z_1 in order to manage the variability inherent in T_1 and Z_2 pertaining to T_2, with Z_1 and Z_2 delineated as follows:

$$Z_1 = e^{\alpha W} \quad \text{and} \quad Z_2 = e^{\beta W} \tag{3.55}$$

where W denotes a random variable and α and β represent parameters. Researchers have extensively employed this unique factor error formulation (Flinn and Heckman, 1982 [50]; Clayton and Cuzick, 1985 [29]; Heckman and Walker, 1990 [70]; Huang and Wolfe, 2002 [85]). This approach exhibits a greater degree of flexibility compared to the presumption of shared frailty concerning T_1 and T_2, and it enables a negative correlation to a certain point by accepting different signs of α and β. Nevertheless, it continues to establish a connection between variance and correlation, as well as between mean and variance. The average risk to the population, as opposed to individual risk, is hence restricted to fluctuate within limited parameters. Lindeboom and Van Den Berg (1994) [110] concluded that estimating bivariate survival models with a univariate parameterisation of the mixing distribution is challenging, as a univariate random variable may fail to capture both the dependence of survival times and the modifications in sample composition resulting from unobserved heterogeneity.

Clearly, such issues do not occur in a proper bivariate context where the dependence between T_1 and T_2 may be adjusted independently of the marginal distributions associated with T_1 and T_2. Aalen (1987) [2] investigated the implications of multivariate mixing distributions in the context of a Markov Chain. Marshall and Olkin (1988) [118] examined a range of multivariate correlated frailty distributions. Their research was further

advanced by Yashin et al. (1995) [169], who examined the bivariate correlated gamma frailty model and employed it in the analysis of twin survival data. This methodology serves as the foundation for various extensions examined in this thesis, which will be discussed in greater detail following the subsequent section.

Correlated Frailty Model

Examine some bivariate data, such as the lifetimes of twins, the age of disease beginning in spouses, or the time interval preceding the presentation of disease in paired organs such as the kidneys or the eyes. In the context of the bivariate correlated frailty model, the frailty exhibited by each individual within a pair is delineated by a quantification of relative risk, analogous to its definition in the univariate case. In a pair of individuals, frailties may vary, in contrast to the uniformity observed in the shared frailty model. It is posited that the frailties exert a multiplicative influence on the baseline hazard function and that the observations contained within a given pair are conditionally independent, considering the frailties inherent in the analysis, the risk associated with individual $j(j=1,2)$ within pair $i(i=1,\ldots,n)$ is articulated as

$$\lambda(t) = Z_{ij}\lambda_{0j}(t)e^{\beta^T X_{ij}} \tag{3.56}$$

where t denotes either age or temporal progression, X_{ij} signifies a vector comprising observable covariates, and β represents a vector of unobserved regression parameters that characterise the impact of the covariates X_{ij} and $\lambda_{0j}(t)$ represent baseline hazard functions while Z_{ij} represents latent random effects or frailty. The defining feature of bivariate correlated frailty models lies in the joint distribution of the two-dimensional vector of frailties (Z_{i1}, Z_{i2}).

Any methodology in this situation will depend on likelihood functions. The construction of a marginal likelihood function is predicated upon the postulation of conditional independence of the lifetime, conditional upon the frailty. Let δ_{ij} represent a censoring indicator for the individual denoted as j (where $j=1,2$) within the pair identified as i

(where $i = 1, \ldots, n$). The indicator δ_{ij} assumes the value of 1 if the individual has undergone the event of interest, and takes the value of 0 otherwise. The conditional survival function pertaining to the j-th individual within the i-th pair is articulated

$$S\left(t \mid Z_{ij}, X_{ij}\right) = e^{Z_{ij}\Lambda_{0j}(t)e^{\beta^T X_{ij}}} \tag{3.57}$$

where $\Lambda_{0j}(t)$ represents the cumulative baseline hazard function. The involvement of the j^{th} participant within the i^{th} pair of the conditional likelihood is expressed as

$$L\left(t_{ij}, \delta_{ij} \mid Z_{ij}, X_{ij}\right) = \left(Z_{ij}\lambda_{0j}\left(t_{ij}\right)e^{\beta^T X_{ij}}\right)^{\delta_{ij}} e^{Z_{ij}\Lambda_{0j}\left(t_{ij}\right)e^{\beta^T X_{ij}}} \tag{3.58}$$

where t_{ij} represents the age at mortality or the time of censoring for individual j within pair i. By presuming the conditional independence of longevity given the frailty and subsequently integrating the frailty, we obtain the marginal likelihood function:

$$\begin{aligned} L(t, \delta \mid X) = \prod_{i=1}^{n} \iint_{R^+ \times R^+} & \left(z_{i1}\lambda_{01}\left(t_{i1}\right)e^{\beta^T X_{i1}}\right)^{\delta_{i1}} e^{z_{i1}\Lambda_{01}(t_{i1})e^{\beta^T X_{i1}}} \\ * & \left(z_{i2}\lambda_{02}\left(t_{i2}\right)e^{\beta^T X_{i2}}\right)^{\delta_{i2}} e^{z_{i2}\Lambda_{02}(t_{i2})e^{\beta^T X_{i2}}} f_Z\left(z_{i1}, z_{i2}\right) dz_{i1} dz_{i2} \end{aligned} \tag{3.59}$$

where $t = (t_1, \ldots, t_n), t_i = (t_{i1}, t_{i2}), \delta = (\delta_1, \ldots, \delta_n), \delta_i = (\delta_{i1}, \delta_{i2}), X = (X_1, \ldots, X_n)$, $X_i = (X_{i1}, X_{i2})$, and $f_Z(\cdot, \cdot)$ denotes the probability density function corresponding to the relevant frailty distribution.

Bandyopadhyay and Basu (1990) [21], as well as Gupta and Gupta (1990) [62], utilized a differing methodological approach for bivariate frailty modeling. The basic concept of their elaborated framework is a bivariate hazard model.

$$\lambda\left(t_1, t_2, Z\right) = Z\lambda_0\left(t_1, t_2\right) \tag{3.60}$$

The two times do not exhibit conditional independence when considering the frailty. Dependence arises from the bivariate model and by integrating out the frailty Z.

Chapter 4

Two-parameter Lindley distribution

4.1 Lindley Distribution

The Lindley distribution can be characterized and comprehensively understood through its mathematical formulation, which is expressed in the form of a probability density function (p.d.f.), thereby providing a foundational framework for analyzing the behavior of random variables that adhere to this particular statistical distribution

$$f(x) = \frac{\theta^2}{\theta+1}(1+x)\mathrm{e}^{-\theta x}, \quad x>0, \theta>0 \tag{4.1}$$

the conceptual framework for this particular statistical methodology, which was initially presented to the academic community by Lindley during the years 1958 and 1965, marks a significant advancement in the field of probability theory. Furthermore, it is pertinent to note that the cumulative distribution function (c.d.f.), which is intrinsically linked to this theoretical construct, serves a crucial role in elucidating the behaviors and characteristics of the associated probability distribution and is expressed as follow

$$F(x) = 1 - \frac{\theta+1+\theta x}{\theta+1}\mathrm{e}^{-\theta x}, \quad x>0, \theta>0 \tag{4.2}$$

A probability distribution that holds a notable resemblance to the one identified in equation (4.1) is the widely recognized and extensively studied exponential distribution, which is distinctly characterized by its specific probability density function

$$f(x) = \theta\mathrm{e}^{-\theta x}, \quad x>0, \theta>0. \tag{4.3}$$

Nevertheless, owing to the broad application of the exponential distribution in statis-

tics and numerous applied fields, the Lindley distribution presented in (4.1) has been largely neglected in the literature.

Sankaran (1970) [146] employed (4.1) as a mixture model for the Poisson parameter to derive a mixed Poisson distribution, referred to as the discrete Poisson-Lindley distribution, with the probability mass function (p.m.f.):

$$P_\theta(Z=z)=\int_0^\infty \mathrm{e}^{-\lambda}\frac{\lambda^z}{z!}\cdot f(\lambda;\theta)\mathrm{d}\lambda=\frac{\theta^2(\theta+2+z)}{(\theta+1)^{3+z}},\quad z=0,1,2,\ldots,\theta>0 \tag{4.4}$$

It is well established that numerous characteristics of a continuous mixing distribution are preserved in the associated discrete mixed Poisson distribution, as illustrated by Holgate (1970) [74] and Grandell (1997) [58]. Therefore, many characteristics of (4.4), including shape and failure rate, can be directly derived from those of (4.1).

The first derivative of equation (4.1) is:

$$\frac{\mathrm{d}}{\mathrm{d}x}f(x)=\frac{\theta^2}{\theta+1}(1-\theta-\theta x)\mathrm{e}^{-\theta x} \tag{4.5}$$

Consequently,

For $\theta<1$, $\frac{\mathrm{d}}{\mathrm{d}x}f(x)=0$ implies that $x_0=((1-\theta)/\theta)$ is the only critical point at which $f(x)$ attains its maximum value.

For $\theta\geq 1$, it holds that $\frac{\mathrm{d}}{\mathrm{d}x}f(x)\leq 0$, indicating that the function $f(x)$ exhibits a continuous decline behavior with respect to the variable x.

The mode, which represents the value that appears most frequently within this probability distribution, is provided as

$$\mathrm{mode}(X)=\begin{cases}\frac{1-\theta}{\theta}, & \text{if } 0<\theta<1\\ 0, & \text{otherwise}\end{cases} \tag{4.6}$$

Let the notation $X\sim\mathrm{Lindley}(\theta)$ represent a continuous random variable whose probability density function is specified by (4.1).

Theorem 1. Let $X \sim \text{Lindley}(\theta)$. Then

$\text{mode}(X) < \text{median}\ (X) < E(X)$.

The conclusion of Theorem 1 is similarly applicable to the exponential distribution.

Abadir (2005) [6] presented alternatives demonstrating that the inequality mode $\leq$ median $\leq$ mean does not necessarily apply to unimodal and positively skewed distributions with existing first three moments.

4.1.1 Moments and Respective Quantifications

The mathematical representation of the rth moment in relation to the foundational points or origins of the Lindley distribution can be articulated as

$$\mu'_r = E\left(X^r\right) = \frac{r!(\theta + r + 1)}{\theta^r(\theta + 1)}, \quad r = 1, 2, \ldots \tag{4.7}$$

Specifically, we possess

$$\mu'_1 = \frac{\theta + 2}{\theta(\theta + 1)} = \mu \tag{4.8}$$

$$\mu'_2 = \frac{2(\theta + 3)}{\theta^2(\theta + 1)} \tag{4.9}$$

$$\mu'_3 = \frac{6(\theta + 4)}{\theta^3(\theta + 1)} \tag{4.10}$$

$$\mu'_4 = \frac{24(\theta + 5)}{\theta^4(\theta + 1)} \tag{4.11}$$

The mathematical representation of the rth moment, which is calculated in proximity to the origin for the statistical framework of the exponential distribution, is expressed by $\mu'_r = r!/\theta^r$.

Ottestad (1944) [136] articulates the notion that, under the condition wherein the random variable Z is conditionally distributed as a Poisson distribution with parameter x, denoted as $Z \mid X = x \sim \text{Poisson}(x)$, it follows that the rth factorial moment of the variable Z can be mathematically expressed through the expectation of the product

$E[Z(Z-1)\dots(Z-r+1)]$, which is in turn proportional to the rth moment of the random variable X, represented as $E(X^r)$, where the index r takes on the values $r = 1,2,\dots$.

Consequently, the rth factorial moment of the discrete Poisson-Lindley distribution, characterized by the probability mass function (4.4), is

$$E[Z(Z-1)\dots(Z-r+1)] = \frac{r!(\theta+r+1)}{\theta^r(\theta+1)}, \quad r = 1,2,\dots \tag{4.12}$$

The principal statistical measures, often referred to as the central moments, which are integral to the comprehensive understanding of the Lindley distribution, are as follows

$$\mu_k = E\left\{(X-\mu)^k\right\} = \sum_{r=0}^{k}\binom{k}{r}\mu_r'(-\mu)^{k-r} \tag{4.13}$$

Specifically, we possess

$$\mu_2 = \frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2} = \sigma^2 \tag{4.14}$$

$$\mu_3 = \frac{2\left(\theta^3+6\theta^2+6\theta+2\right)}{\theta^3(\theta+1)^3} \tag{4.15}$$

$$\mu_4 = \frac{3\left(3\theta^4+24\theta^3+44\theta^2+32\theta+8\right)}{\theta^4(\theta+1)^4} \tag{4.16}$$

The coefficient of variation (γ), skewness $\left(\sqrt{\beta_1}\right)$, and kurtosis ($\beta_2$) are as follows:

$$\gamma = \frac{\sqrt{\theta^2+4\theta+2}}{\theta+2} \tag{4.17}$$

$$\sqrt{\beta_1} = \frac{2\left(\theta^3+6\theta^2+6\theta+2\right)}{\left(\theta^2+4\theta+2\right)^{3/2}} \tag{4.18}$$

$$\beta_2 = \frac{3\left(3\theta^4+24\theta^3+44\theta^2+32\theta+8\right)}{\left(\theta^2+4\theta+2\right)^2} \tag{4.19}$$

4.1.2 Hazard Rate and Mean Residual Life Functions

In the context of a continuous probability distribution delineated by the probability density function (p.d.f.) denoted as $f(x)$ and the cumulative distribution function (c.d.f.) represented as $F(x)$, the failure rate function, which is frequently termed the hazard rate function, is articulated as

$$h(x) = \lim_{\Delta x \to 0} \frac{P(X < x + \Delta x \mid X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \tag{4.20}$$

The hazard rate function for the Lindley distribution is

$$h(x) = \frac{\theta^2(1+x)}{\theta + 1 + \theta x} \tag{4.21}$$

$h(0) = f(0) = \theta^2/(\theta+1)$.

$h(x)$ is a monotonically growing function in terms of x and θ, satisfying $\frac{\theta^2}{\theta+1} < h(x) < \theta$.

This sequence of implications is widely recognised

$$\text{IFR} \Rightarrow \text{IFRA} \Rightarrow \text{NBU} \Rightarrow \text{NBUE}$$

where IFR, IFRA, NBU, NBUE denote increasing failure rate, increasing failure rate average, new better than used, and new better than used in expectation, respectively. For further information regarding the definitions of these ageing features, refer to Barlow and Proschan (1975) [22].

Page 135 of Grandell (1997) [58] states that an IFR mixing distribution indicates that the associated mixed Poisson distribution is also IFR. Consequently, the discrete Poisson-Lindley distribution exhibits the increasing failure rate property.

In the case of the exponential distribution, $h(x) = \theta$, so (4.21) once more illustrates the flexibility of the Lindley distribution compared to the exponential distribution.

The mean residual life function, which serves as a significant statistical measure for a continuous probability distribution and is intricately characterized by its associated prob-

ability density function, denoted as $f(x)$, alongside the cumulative distribution function represented as $F(x)$, is formally expressed as

$$m(x) = E(X - x \mid X > x) = \frac{1}{1 - F(x)} \int_x^\infty [1 - F(t)] \mathrm{d}t \tag{4.22}$$

The mean residual life function for the Lindley distribution is

$$m(x) = \frac{1}{(\theta + 1 + \theta x)\mathrm{e}^{-\theta x}} \int_x^\infty (\theta + 1 + t)\mathrm{e}^{-\theta t} \, \mathrm{d}t = \frac{\theta + 2 + \theta x}{\theta(\theta + 1 + \theta x)} \tag{4.23}$$

$m(0) = \mu$.

$m(x)$ is a monotonically falling function in x and θ, satisfying

$1/\theta < m(x) < (\theta + 2)/(\theta(\theta + 1)) = \mu$.

The IFR discrete Poisson-Lindley distribution possesses a declining mean residual life function.

In the case of the exponential distribution, $m(x) = 1/\theta$, so (4.23) once more illustrates the flexibility of the Lindley distribution in comparison to the exponential distribution.

4.1.3 Estimation

For a randomly selected sample consisting of the observations denoted as $X_1, X_2, \ldots, X_n$, which have been extracted from the Lindley distribution as specified in equation (4.1), it is noteworthy to mention that both the method of moments and the maximum likelihood estimators, which are employed to deduce the parameter θ, yield identical results and are represented as follow

$$\hat{\theta} = \frac{-(\bar{X} - 1) + \sqrt{(\bar{X} - 1)^2 + 8\bar{X}}}{2\bar{X}}, \quad \bar{X} > 0 \tag{4.24}$$

The theorem that follows subsequently establishes and provides a rigorous proof that the estimator utilized for the parameter denoted as θ exhibits a systematic bias, indicating that its expected value does not align with the true value of the parameter being estimated.

Theorem 2. The estimator $\hat{\theta}$ of θ exhibits positive bias, specifically $E\{\hat{\theta}\} - \theta > 0$.

Proof. Let

$$\hat{\theta} = g(\bar{X})$$

and

$$g(t) = \frac{-(t-1) + \sqrt{(t-1)^2 + 8t}}{2t}$$

for $t > 0$. Since

$$g''(t) = \frac{1}{t^3}\left[1 + \frac{3t^3 + 15t^2 + 9t + 1}{[(t-1)^2 + 8t]^{3/2}}\right] > 0$$

$g(t)$ exhibits strict convexity. Consequently, through Jensen's inequality, we have $E\{g(\bar{X})\} > g\{E(\bar{X})\}$. subsequently, given $g\{E(\bar{X})\} = g(\mu) = g((\theta+2)/(\theta(\theta+1))) = \theta$, we conclude that $E(\hat{\theta}) > \theta$.

The subsequent theorem presents the limiting distribution of $\hat{\theta}$.

Theorem 3. The estimator $\hat{\theta}$ representing the parameter θ possesses the properties of both consistency and asymptotic normality:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N\left(0, \frac{1}{\sigma^2}\right) \tag{4.25}$$

The confidence interval for θ with a high sample size, denoted as $100(1-\alpha)\%$, is expressed as follows:

$$\hat{\theta} \pm z_{\alpha/2} \cdot \frac{1}{\sqrt{n\hat{\sigma}^2}} \tag{4.26}$$

where $z_{\alpha/2}$ denotes the $1-(\alpha/2)$ percentile of the standard normal distribution.

Proof. Given that μ is finite, it follows that $\bar{X} \xrightarrow{P} \mu$. considering that $g(t)$ exhibits continuity at the point $t = \mu$, it follows that $g(\bar{X}) \xrightarrow{P} g(\mu)$, which implies $\hat{\theta} \xrightarrow{P} \theta$. Since, $\sigma^2 < \infty$, the central limit theorem implies that we have

$$\sqrt{n}(\bar{X}-\mu) \xrightarrow{D} N\left(0,\sigma^2\right).$$

Furthermore, given that $g(\mu)$ is differentiable and $g'(\mu) \neq 0$, the delta technique yields

$$\sqrt{n}(g(\bar{X})-g(\mu)) \xrightarrow{D} N\left(0,\left[g'(\mu)\right]^2\sigma^2\right)$$

As a result, given that $g(\bar{X}) = \hat{\theta}, g(\mu) = \theta$ and
$g'(\mu) = -\left(1/2\mu^2\right)\left[1+(1+3\mu)/\left(\sqrt{(\mu-1)^2+8\mu}\right)\right] = -\left(1/\sigma^2\right)$,
the theorem follows.

In the framework of the exponential distribution, the estimator $\hat{\theta} = 1/\bar{X}$ operates as both the maximum likelihood estimator and the method of moments estimator for the parameter θ. This estimator exhibits bias, consistency, and asymptotic normality.

4.2 Construction of Two-Parameter Lindley Distribution

The Two-Parameter Lindley distribution, defined by the parameters α and θ, is elucidated through its probability density function (p.d.f).

$$f(x;\alpha,\theta) = \frac{\theta^2}{\alpha\theta+1}(\alpha+x)\exp(-\theta x); x>0, \theta>0, \alpha\theta>-1 \tag{4.27}$$

At the point where the parameter α is equal to 1, the mathematical distribution in question undergoes a significant simplification, ultimately resulting in what is known as the Lindley distribution, which possesses distinctive characteristics that set it apart from other distributions, whereas in the contrasting scenario where the parameter α takes on the value of 0, this distribution further simplifies to the well-known gamma distribution, which is defined by its specific parameters, namely $(2,\theta)$. The probability density function (4.27) can be articulated as a fusion of exponential (θ) and gamma $(2,\theta)$ distributions as subsequent:

$$f(x;\alpha,\theta) = pf_1(x) + (1-p)f_2(x) \tag{4.28}$$

where $p = \dfrac{\theta}{\alpha\theta+1}, f_1(x) = \theta\exp(-\theta x)$ and $f_2(x) = \theta^2 x\exp(-\theta x)$

The first derivative of (4.27) is determined as

$$f^{'}(x) = \frac{\theta^2}{\alpha\theta+1}(1-\alpha\theta-x\theta)\exp(-\theta x) \tag{4.29}$$

thus, $f^{'}(x) = 0$ gives $x = \dfrac{1-\alpha\theta}{\theta}$. Consequently, it may be concluded that

For $|\alpha\theta| < 1, x_0 = \dfrac{1-\alpha\theta}{\theta}$ is the only critical point at which $f_1(x)$ attains its maximum value.

For $\alpha \succeq 1$, the derivative $f'(x)$ is constrained to be non-positive, thereby signifying that the function $f(x)$ exhibits a monotonically declining behavior with respect to the variable x.

As a result, the distribution's mode is determined by

$$Mode = \left\{ \begin{array}{c} \dfrac{1-\alpha\theta}{\theta}, |\alpha\theta| < 1 \\ 0, \quad \text{otherwise} \end{array} \right\} \tag{4.30}$$

The distribution's cumulative distribution function can be written in the form

$$F(x) = 1 - \frac{1+\alpha\theta+x\theta}{\alpha\theta+1}\exp(-\theta x); x > 0, \theta > 0, \alpha\theta > -1 \tag{4.31}$$

4.2.1 Moments and Respective Quantifications

The mathematical expression representing the r^{th} moment in relation to the origin for the two-parameter Lindley distribution has been thoroughly obtained and articulated as follows.

$$\mu_r^{'} = \frac{\Gamma(r+1)(\alpha\theta+r+1)}{\theta^r(\alpha\theta+1)}; r = 1,2,... \tag{4.32}$$

For $r = 1,2,3$, and 4, the initial four moments about the origin are derived as follows:

$$\mu_1^{'} = \frac{\alpha\theta+2}{\theta(\alpha\theta+1)}, \mu_2^{'} = \frac{2(\alpha\theta+3)}{\theta^2(\alpha\theta+1)}, \mu_3^{'} = \frac{6(\alpha\theta+4)}{\theta^3(\alpha\theta+1)}, \mu_4^{'} = \frac{24(\alpha\theta+5)}{\theta^4(\alpha\theta+1)} \tag{4.33}$$

It is clearly illustrated that for $\alpha = 1$, the moments about the origin of the probability distribution align with the moments inherent to the Lindley distribution. The average value of the distribution invariably surpasses the mode, signifying a positive skewness. The central moments associated with this distribution have been thoroughly calculated as

$$\mu_2^{'} = \frac{\alpha^2\theta^2+4\alpha\theta+2}{\theta(\alpha\theta+1)^2}, \tag{4.34}$$

$$\mu_3^{'} = \frac{2(\alpha^3\theta^3+6\alpha^2\theta^2+6\alpha\theta+2)}{\theta^3(\alpha\theta+1)^3}, \tag{4.35}$$

$$\mu_4^{'} = \frac{3(3\alpha^4\theta^4+24\alpha^3\theta^3+44\alpha^2\theta^2+32\alpha\theta+8)}{\theta^4(\alpha\theta+1)^4}, \tag{4.36}$$

It can be straightforwardly demonstrated that when $\alpha = 1$, the central moments of the distribution align with the moments of the Lindley distribution.

The metrics of variability, specifically the coefficients of variation (γ), skewness represented as ($\sqrt{\beta_1}$), and kurtosis denoted by (β_2), associated with the two-parameter Lindley distribution are articulated as follow

$$\gamma = \frac{\sigma}{\mu_1^{'}} = \frac{\sqrt{\alpha^2\theta^2+4\alpha\theta+2}}{(\alpha\theta+2)} \tag{4.37}$$

$$\sqrt{\beta_1} = \frac{2(\alpha^3\theta^3+6\alpha^2\theta^2+6\alpha\theta+2)}{(\alpha^2\theta^2+4\alpha\theta+2)^{3/2}} \tag{4.38}$$

$$\beta_2 = \frac{3(3\alpha^4\theta^4+24\alpha^3\theta^3+44\alpha^2\theta^2+32\alpha\theta+8)}{(\alpha^2\theta^2+4\alpha\theta+2)^2} \tag{4.39}$$

4.2.2 Hazard Rate and Mean Residual Life Functions

The function representing the failure rate, commonly designated as the hazard rate, in conjunction with the mean residual life function, can be articulated for any continuous probability distribution once its probability density function $f(x)$ and cumulative distribution function $F(x)$ are explicitly delineated and can be expressed as

$$h(x) = \lim_{\Delta x \to 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \tag{4.40}$$

and

$$m(x) = E\left[X - x | X > x\right] = \frac{1}{1 - F(x)} \int_x^\infty \left[1 - F(t)\right] dt \tag{4.41}$$

The functions $h(x)$ and $m(x)$, which are related to the distribution, delineate the failure rate and the mean residual life, respectively, and can be articulated as:

$$h(x) = \frac{\theta^2(\alpha + x)}{1 + \alpha\theta + x\theta} \tag{4.42}$$

and

$$m(x) = \frac{1}{(1 + \alpha\theta + \theta x)\exp(-\theta x)} \int_x^\infty (1 + \alpha\theta + \theta t)\exp(-\theta t) dt = \frac{2 + \alpha\theta + \theta x}{\theta\,(1 + \alpha\theta + \theta x)} \tag{4.43}$$

It can be simply determined that $h(0) = \dfrac{\theta^2\alpha}{\alpha\theta + 1} = f(0)$ and $m(0) = \dfrac{\alpha\theta + 2}{\theta(\alpha\theta + 1)} = \mu_1'$. It is apparent that the function $h(x)$ exhibits monotonic increasing behavior with respect to the variables x, α, and θ, whereas the function $m(x)$ demonstrates a decreasing trend in relation to the variables x and α, while concurrently displaying an increasing trend in relation to θ. For the specific case where $\alpha = 1$, equations (4.42) and (4.43) are reduced to the corresponding metrics of the Lindley distribution. The failure rate function, alongside the mean residual life function of the distribution, illustrates its adaptability in comparison to both the Lindley distribution and the exponential distribution.

4.2.3 Estimation of Parameters

Maximum Likelihood Approach

Let the variables $x_1, x_2, \ldots, x_n$ be indicative of a random sample that has been extracted from a two-parameter Lindley distribution, as identified in equation (4.27), and in this context, let us denote by f_x the frequency that is observed within the sample that corresponds to the value $X = x$ for each value of x ranging from 1 to k, with k representing the maximum observed value that possesses a non-zero frequency, thereby satisfying the condition that the summation of the observed frequencies from $x = 1$ to k equals the total sample size n. The likelihood function, denoted as L, which reflects the statistical characteristics of the two-parameter Lindley distribution as described in equation (4.27), can be articulated in a manner that expresses the comprehensive relationship between the parameters of the distribution and the observed data, thus enabling us to derive further statistical inferences. Therefore, the formulation of the likelihood function is given by

$$L = \left(\frac{\theta^2}{\alpha\theta + 1}\right)^n \prod_{x=1}^{k} (\alpha + x)^{f_x} \exp(-n\theta\overline{X}) \tag{4.44}$$

Hence, the formulation of the log-likelihood function is established as

$$\log L = n\log\theta^2 - n\log(\alpha\theta + 1) + \sum_{x=1}^{k} f_x \log(\alpha + x) - n\theta\overline{X} \tag{4.45}$$

The subsequent log-likelihood equations are thus formulated as

$$\frac{\partial \log L}{\partial \theta} = \frac{2n}{\theta} - \frac{n\alpha}{\alpha\theta + 1} - n\overline{X} = 0 \tag{4.46}$$

$$\frac{\partial \log L}{\partial \alpha} = -\frac{n\theta}{\alpha\theta + 1} + \sum_{x=1}^{k} \frac{f_x}{\alpha + x} = 0 \tag{4.47}$$

Equation (4.46) provides the expression $\overline{X} = \dfrac{\alpha\theta + 2}{\theta(\alpha\theta + 1)}$, which represents the mean value of the two-parameter Lindley distribution. The expressions shown in (4.46) and

(4.47) appear to be unsolvable directly. Nonetheless, Fisher's scoring method is applicable for resolving these equations. We possess

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{2n}{\theta^2} + \frac{n\alpha^2}{(\alpha\theta+1)^2} \tag{4.48}$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = \frac{n}{(\alpha\theta+1)^2} \tag{4.49}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{n\theta^2}{(\alpha\theta+1)^2} - \sum_{x=1}^{k} \frac{f_x}{(\alpha+x)^2} \tag{4.50}$$

The subsequent equations for $\widehat{\theta}$ and $\widehat{\alpha}$ are solvable are follow

$$\left[\begin{array}{cc} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \log L}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2} \end{array}\right]_{\substack{\widehat{\theta}=\theta_0 \\ \widehat{\alpha}=\alpha_0}} \left[\begin{array}{c} \widehat{\theta}-\theta_0 \\ \widehat{\alpha}-\alpha_0 \end{array}\right] = \left[\begin{array}{c} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \alpha} \end{array}\right]_{\substack{\widehat{\theta}=\theta_0 \\ \widehat{\alpha}=\alpha_0}} \tag{4.51}$$

where θ_0 and α_0 represent the preliminary values of θ and α, respectively. The equations are approached through an iterative solution process until sufficiently approximate estimations of $\widehat{\theta}$ and $\widehat{\alpha}$ are achieved.

Moments Approach Estimations

Employing the initial two moments about the origin, we obtain

$$\frac{\mu_2^{'}}{\mu_1^{'2}} = k(say) = \frac{2(\alpha\theta+3)(\alpha\theta+1)}{(\alpha\theta+2)^2} \tag{4.52}$$

By setting $b = \alpha\theta$, we obtain

$$\frac{\mu_2^{'}}{\mu_1^{'2}} = \frac{2(b+3)(b+1)}{(b+2)^2} = \frac{2b^2+8b+6}{b^2+4b+4} = k \tag{4.53}$$

This provides

$$(2-k)b^2 + 4(2-k)b + 2(3-2k) = 0 \tag{4.54}$$

This constitutes a quadratic equation in b. By replacing the first and second moments $\mu_1^{'}$ and $\mu_2^{'}$ with their respective empirical moments, an estimate of k may be derived from $\overline{X}$ and $m_2^{'}$, which can then be utilised to solve equation (4.54) and yield an estimate of b.

By incorporating this estimation of b into the computation for the mean of the two-parameter Lindley distribution, a corresponding estimation of θ may be obtained.

$$\widehat{\theta} = \left(\frac{b+2}{b+1}\right)\frac{1}{\overline{\overline{X}}} \tag{4.55}$$

In order to derive an estimation of α, we substitute b along with the estimation of θ into the equation $b = \alpha\theta$, resulting an estimate for α as

$$\widehat{\alpha} = \frac{b}{\widehat{\theta}} \tag{4.56}$$

Chapter 5

Two-parameter Lindley frailty model

5.1 Construction of The Model

In accordance with the findings of Shanker and Mishra (2013) [151], the probability density function (PDF) associated with the Two Parameter Lindley (TPL) model can be expressed as

$$f_{\alpha,\theta}(y) = \frac{1}{\alpha\theta+1}\theta^2(\alpha+y)\exp(-\theta y)|y>0,\theta>0,\alpha\theta>-1, \tag{5.1}$$

Let us postulate that the frailty variable Z within the conditional model is characterized by a TPL distribution (5.1) that possesses a mean value of one, or equivalently, $E[Z]=1$. This presumption is crucial for the formulation of the consequent frailty model (refer to Elbers and Ridder (1982)) [45]. Consequently, through the implementation of the alternative parameterization of the TPL model articulated in terms of the mean as proposed by Mazucheli et al. (2016) [120], the probability density function (PDF) of the Two Parameter Lindley (TPLF) model is expressed as follows

$$f_\theta(z) = \frac{1}{3}\theta^2(1-\theta)\left[\frac{\theta+2}{\theta(1-\theta)}+z\right]\exp(-\theta z)|\ z>0, \tag{5.2}$$

where the unidentified shape parameter is represented by $\theta>0$. Generally, the variance associated with the frailty distribution serves as a metric to assess the extent of unobserved heterogeneity that exists within the population under investigation in a given study. Taking into account that the probability density function (PDF) (5.2) represents

a distribution of frailty. The variance is articulated as:

$$\sigma^2 = \frac{1}{9\theta^2}\left[(1-\theta)(2\theta+10)+(\theta+2)^2\right], \tag{5.3}$$

In accordance with its variance, the Laplace transform of the frailty probability density function (PDF) (6) is articulated as follows:

$$L_f(s) = \frac{\tau(\sigma^2)-2}{3[s(\sigma^2)-2]}\left\{\frac{D(\sigma^2)}{(1+9\sigma^2)}+\frac{[\tau(\sigma^2)-2]C(\sigma^2)}{(1+9\sigma^2)[s(\sigma^2)-2]}\right\}|s\in\mathbb{R}, \tag{5.4}$$

where $\tau(\sigma^2) = 3\sqrt{2(1+7\sigma^2)}$, $D(\sigma^2) = 18\sigma^2+\tau(\sigma^2), s(\sigma^2) = s(1+9\sigma^2)+\tau(\sigma^2)$ and $C(\sigma^2) = 3+9\sigma^2-\tau(\sigma^2)$. In the interest of clarity, we assess equation (5.4) at the point $s = \Lambda_0(t_i)\xi_i$, where ξ_i is defined as $\exp(x_i^{\intercal}\beta)$. Furthermore, it can be ascertained that the marginal survival function within the framework of the TPLF model can be derived by

$$S(t_i|x_i) = \frac{\tau(\sigma^2)-2}{3\Lambda_0(\xi_i,\sigma^2)}\left\{\frac{D(\sigma^2)}{1+9\sigma^2}+\frac{[\tau(\sigma^2)-2]C(\sigma^2)}{(1+9\sigma^2)\Lambda_0(\xi_i,\sigma^2)}\right\}, \tag{5.5}$$

where

$$\Lambda_0(\xi_i,\sigma^2) = \Lambda_0(t_i)\xi_i(1+9\sigma^2)+\tau(\sigma^2)-2.$$

The derived marginal hazard function consequently attains the following form:

$$\begin{aligned}\lambda(t_i|x_i) &= \left[\frac{\lambda_0(t_i)\xi_i(1+9\sigma^2)}{\Lambda_0(t_i)\xi_i(1+9\sigma^2)+\tau(\sigma^2)-2}\right] \\ &\quad\times\left\{1+\frac{[\tau(\sigma^2)-2]C(\sigma^2)}{\Lambda_0(\xi_i,\sigma^2)D(\sigma^2)+[\tau(\sigma^2)-2]C(\sigma^2)}\right\}.\end{aligned} \tag{5.6}$$

The TPLF framework is subjected to rigorous evaluation and analysis utilizing the Weibull baseline hazard function (WBLHF), the exponential baseline hazard function (EBLHF), the Gompertz baseline hazard function (GBLHF), and the Pareto baseline hazard function (PBLHF).

5.1.1 Two Parameter Lindley Frailty Model with Weibull Baseline Hazard Function

The fundamental characteristics of the Weibull distribution, specifically its baseline hazard function and its cumulative hazard function, are precisely defined as follows

$$\lambda_0(t_i) = \frac{\kappa}{\rho}\left(\frac{t_i}{\rho}\right)^{\kappa-1} | t_i > 0 \quad \text{and} \quad \Lambda_0(t_i) = \left(\frac{t_i}{\rho}\right)^{\kappa} | t_i > 0, \tag{5.7}$$

where $\kappa > 0$ and $\rho > 0$ signify, correspondingly, the shape parameter and the scale parameter. The hazard function of the Weibull distribution exhibits a monotonically decreasing trend when $\kappa < 1$; it remains constant over time when $\kappa = 1$ (representing the exponential distribution); and it demonstrates a monotonically increasing pattern when $\kappa > 1$ (Wienke, 2010) [165]. By integrating equation (5.6) into equation (5.5), the marginal survival and hazard functions pertaining to the TPLF model incorporating the WBLHF are, accordingly, derived as follows

$$S(t_i|x_i) = \left\{\frac{\rho^k\left[\tau\left(\sigma^2\right)-2\right]}{3\rho\left(\sigma^2|t_i^k\xi_i\right)}\right\}\left\{\frac{D\left(\sigma^2\right)}{1+9\sigma^2}+\frac{\rho^k\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\left(1+9\sigma^2\right)\rho\left(\sigma^2|t_i^k\xi_i\right)}\right\}, \tag{5.8}$$

where

$$\rho\left(\sigma^2|t_i^k\xi_i\right) = t_i^k\xi_i\left(1+9\sigma^2\right)+\rho^k\left[\tau\left(\sigma^2\right)-2\right],$$

and

$$\begin{aligned}\lambda\left(t_i|x_i\right) &= \left[\frac{kt_i^{k-1}\xi_i\left(1+9\sigma^2\right)}{t_i^k\xi_i\left(1+9\sigma^2\right)+\rho^k\left[\tau\left(\sigma^2\right)-2\right]}\right] \\ &\quad\times\left\{1+\frac{\rho^k\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\rho\left(\sigma^2|t_i^k\xi_i\right)D\left(\sigma^2\right)+\rho^k\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}\right\}.\end{aligned} \tag{5.9}$$

Figures 5.1 and 5.2 illustrate various curves of the survival and hazard functions associated with the TPLF model, employing a Weibull baseline hazard function characterized

by defined parameter values.

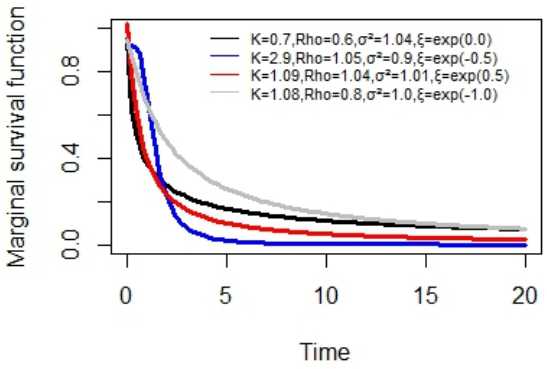

Figure 5.1: Graphical Representation of the Marginal Survival Function for the TPLF model with Weibull Baseline Hazard Function

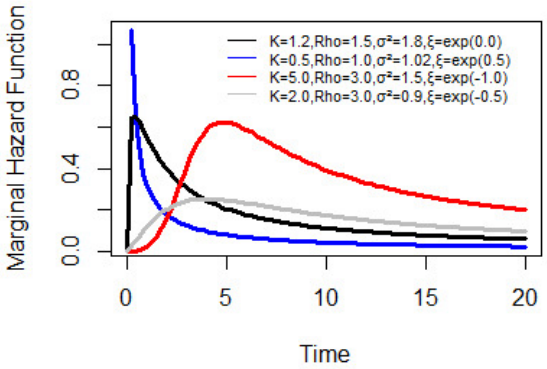

Figure 5.2: Graphical Representation of the Marginal Hazard Function for the TPLF Model with a Weibull Baseline Hazard Function

5.1.2 Two-parameter Lindley Frailty Model with Exponential Baseline Hazard Function

The key constituents of the exponential distribution, notably its baseline hazard function together with its cumulative hazard function, are systematically detailed and illustrated

through the succeeding expressions

$$\lambda_0(t_i) = \lambda | t_i > 0 \quad \text{and} \quad \Lambda_0(t_i) = \lambda t_i | t_i > 0, \tag{5.10}$$

where $\lambda > 0$ denotes the parameter associated with the rate. The hazard function associated with the exponential distribution remains invariant throughout the temporal intervals. This characteristic is referred to as the memoryless property. By substituting equation (5.9) into equation (5.5), one can derive the marginal survival and hazard functions pertinent to the TPLF model that incorporates the EBLHF, which are subsequently articulated as follows

$$\begin{aligned} S(t_i|x_i) &= \left\{ \frac{\tau(\sigma^2) - 2}{3\{\lambda t_i \xi_i (1+9\sigma^2) + \tau(\sigma^2) - 2\}} \right\} \\ &\quad \times \left\{ \frac{D(\sigma^2)}{1+9\sigma^2} + \frac{[\tau(\sigma^2) - 2] C(\sigma^2)}{(1+9\sigma^2)\{\lambda t_i \xi_i (1+9\sigma^2) + \tau(\sigma^2) - 2\}} \right\}, \end{aligned} \tag{5.11}$$

and

$$\begin{aligned} \lambda(t_i|x_i) &= \left[\frac{\lambda \xi_i (1+9\sigma^2)}{\lambda t_i \xi_i (1+9\sigma^2) + \tau(\sigma^2) - 2} \right] \\ &\quad \times \left\{ 1 + \frac{[\tau(\sigma^2) - 2] C(\sigma^2)}{\{\lambda t_i \xi_i (1+9\sigma^2) + \tau(\sigma^2) - 2\} D(\sigma^2) + [\tau(\sigma^2) - 2] C(\sigma^2)} \right\}. \end{aligned} \tag{5.12}$$

Figures 5.3and 5.4 show diverse representations of the survival and hazard functions pertinent to the TPLF model, utilizing an Exponential baseline hazard function with specified parameter values.

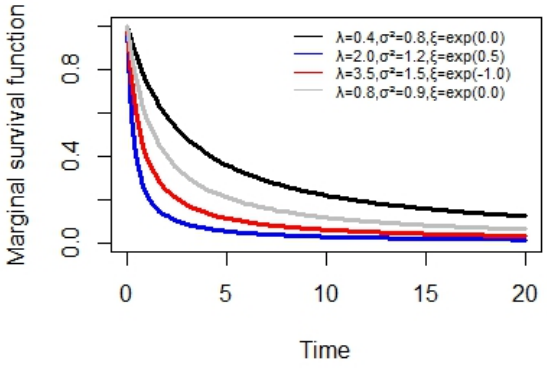

Figure 5.3: Graphical Representation of the Marginal Survival Function for the TPLF model with Exponential Baseline Hazard Function

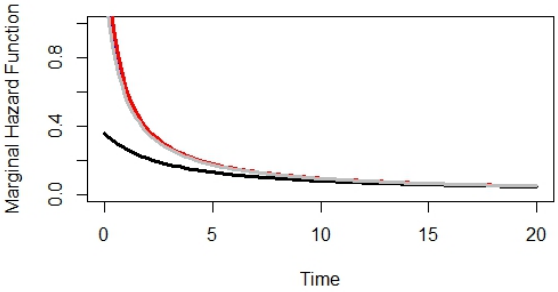

Figure 5.4: Graphical Representation of the Marginal Hazard Function for the TPLF model with Exponential Baseline Hazard Function

5.1.3 Two-Parameter Lindley Frailty Model with Gompertz Baseline Hazard Function

The principal features that characterize the Gompertz distribution, particularly its baseline hazard function and its cumulative hazard function, are comprehensively described as follows:

$$\lambda_0(t_i) = \varphi \exp(\gamma t_i) | t_i > 0 \quad \text{and} \quad \Lambda_0(t_i) = \frac{\varphi}{\gamma} [\exp(\varphi t_i) - 1] | t_i > 0, \tag{5.13}$$

where $\gamma > 0$ denotes the shape parameter and $\varphi > 0$ represents the scale parameter. In instances where $\gamma < 0$, the Gompertz distribution becomes invalid attributable to the convergence of its cumulative hazard function to the constant $-\varphi/\gamma$, resulting in a proportion of individuals achieving either a cure or long-term survival, denoted as $p_0 = \exp(\varphi/\gamma)$, within the examined population. The specific scenario in which $\gamma = 0$ corresponds to the exponential probability distribution. Therefore, the hazard function associated with the Gompertz distribution may exhibit a decreasing trend ($\gamma < 0$), remain constant ($\gamma = 0$), or display an increasing pattern ($\gamma > 0$). By methodically integrating equation (5.12) into equation (5.5), one can derive, in a systematic manner, the marginal survival function as well as the hazard functions specifically associated with the TPLF model, which utilizes the Gompertz baseline hazard function, and these functions can be articulated as follows

$$\begin{aligned} S(t_i|x_i) &= \left\{ \frac{\gamma\left[\tau\left(\sigma^2\right)-2\right]}{3\left\{\varphi\left[\exp\left(\gamma t_i\right)-1\right]\xi_i\left(1+9\sigma^2\right)+\gamma\left[\tau\left(\sigma^2\right)-2\right]\right\}} \right\} \\ &\times \left\{ \frac{D\left(\sigma^2\right)}{\left(1+9\sigma^2\right)} + \frac{\gamma\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\left(1+9\sigma^2\right)\left[\frac{\varphi}{\gamma}\left[\exp\left(\gamma t_i\right)-1\right]\xi_i\left(1+9\sigma^2\right)+\gamma\left[\tau\left(\sigma^2\right)-2\right]\right]} \right\}, \end{aligned} \tag{5.14}$$

and

$$\begin{aligned} \lambda\left(t_i|x_i\right) &= \left[\frac{\varphi\exp\left(\gamma t_i\right)\xi_i\left(1+9\sigma^2\right)}{\varphi\left[\exp\left(\gamma t_i\right)-1\right]\xi_i\left(1+9\sigma^2\right)+\gamma\left[\tau\left(\sigma^2\right)-2\right]} \right] \\ &\times \left(1+\frac{\gamma\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\left\{ \begin{array}{c} \left\{\varphi\left[\exp\left(\gamma t_i\right)-1\right]\xi_i\left(1+9\sigma^2\right)+\gamma\left[\tau\left(\sigma^2\right)-2\right]\right\}D\left(\sigma^2\right) \\ +\gamma\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right) \end{array} \right\}} \right). \end{aligned} \tag{5.15}$$

Under fixed Gompertz baseline parameters, Figures 5.5 and 5.6 demonstrate multiple shapes of the survival and hazard functions for the TPLF model.

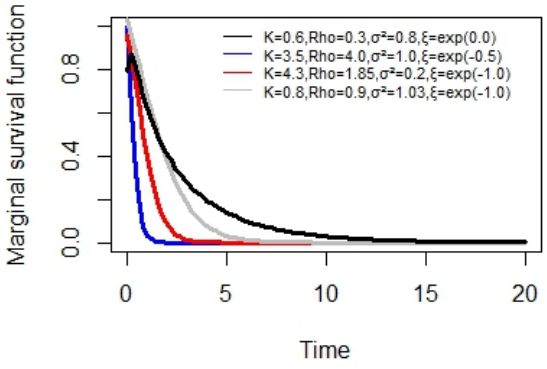

Figure 5.5: Graphical Representation of the Marginal Survival Function for the TPLF model with Gompertz Baseline Hazard Function

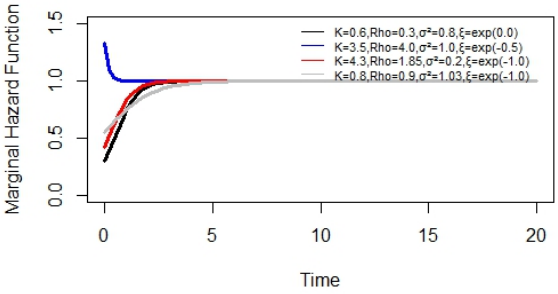

Figure 5.6: Graphical Representation of the Marginal Hazard Function for the TPLF model with Gompertz Baseline Hazard Function

5.1.4 Two-parameter Lindley Frailty Model with Pareto Baseline Hazard Function

The essential characteristics of the Pareto distribution, notably the functions that specify both the baseline hazard and the cumulative hazard, are precisely articulated as follows

$$\lambda_0(t_i) = \frac{\eta}{\alpha+t_i}|t_i>0 \quad \text{and} \quad \Lambda_0(t_i) = -\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)|t_i>0, \tag{5.16}$$

This distribution exhibits a skewness and is characterized by a heavy tail, dependent on two parameters $\alpha>0$ and $\eta>0$. The hazard function demonstrates a monotonically decreasing trend. By incorporating the equation denoted as (5.15) into the model established by equation (5.5), it becomes feasible to derive the marginal survival function, along with the corresponding hazard functions, which are pertinent to the TPLF model that with the PBLHF, and these results are subsequently articulated as follows:

$$\begin{aligned} S(t_i|x_i) &= \left\{\frac{\tau(\sigma^2)-2}{3\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\xi_i(1+9\sigma^2)+\tau(\sigma^2)-2\right]}\right\} \\ &\times\left\{\frac{D(\sigma^2)}{1+9\sigma^2}+\frac{\left[\tau(\sigma^2)-2\right]C(\sigma^2)}{(1+9\sigma^2)\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\xi_i(1+9\sigma^2)+\tau(\sigma^2)-2\right]}\right\}, \end{aligned} \tag{5.17}$$

and

$$\begin{aligned} \lambda(t_i|x_i) &= \frac{\eta}{\alpha+t_i}\left[\frac{\xi_i(1+9\sigma^2)}{-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\xi_i(1+9\sigma^2)+\tau(\sigma^2)-2}\right] \\ &\times\left\{1+\frac{\left[\tau(\sigma^2)-2\right]C(\sigma^2)}{\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\xi_i(1+9\sigma^2)+\tau(\sigma^2)\right]D(\sigma^2)+\left[\tau(\sigma^2)-2\right]C(\sigma^2)}\right\}. \end{aligned} \tag{5.18}$$

Various shapes of the survival and hazard functions for the TPLF model are illustrated in Figures 5.7 and 5.8, where a Pareto baseline hazard function is adopted with fixed parameter values.

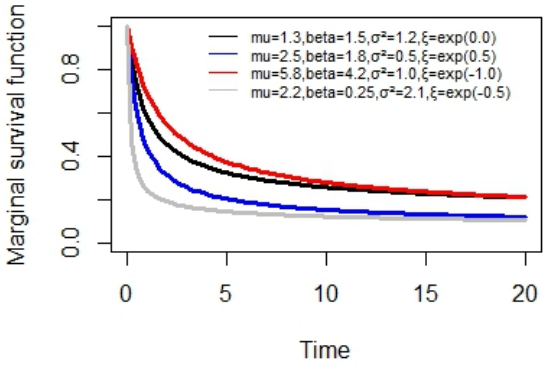

Figure 5.7: Graphical Representation of the Marginal Survival Function for the TPLF model with Pareto Baseline Hazard Function

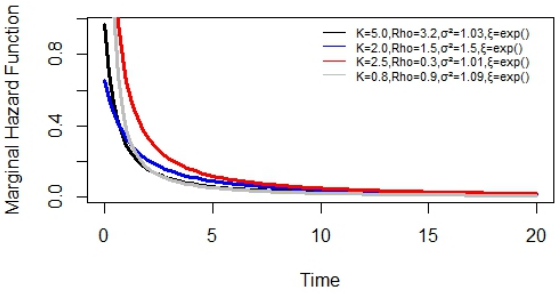

Figure 5.8: Graphical Representation of the Marginal Hazard Function for the TPLF model with Pareto Baseline Hazard Function

5.2 Estimation

In this segment, we delineate the maximum likelihood methodology for the estimation of the parameters associated with the TPLF model, employing Weibull, exponential, Gompertz, and Pareto baseline hazard functions. Maximum Likelihood Estimators (MXLEs) exhibit several distinguished characteristics, encompassing attributes such as consistency, efficiency, asymptotic normality, and various others, depending on specific regularity conditions (Lehmann and Casella, 2006) [106].

In specific instances of research investigation, data pertaining to lifetimes may be unavailable. For example, certain lifetimes may exhibit right-censoring, whereby the only ascertainable information is that these values exceed the documented value. In such a scenario, let T_i and C_i denote the lifetime and censoring time variables associated with the i^{th} subject within the population being examined, respectively. Given the existence of two independent random variables, denoted as T_i and C_i, we further define the censoring indicator δ_i as $\mathbf{I}_{(T_i \leq C_i)}$, where this indicator function takes on the value of one, indicating that the variable T_i represents a lifetime observation, and conversely, it assumes a value of zero in scenarios where the condition is not satisfied. We subsequently assess the variable $t_i = \min\{T_i, C_i\}$. Denote x_i as a $p \times 1$ vector representing the covariates pertinent to the i^{th} participant. Thereafter, utilizing a sample comprising n individuals, the likelihood function associated with the model parameter vector $\underline{\mathbf{P}}$ within the context of non-informative censoring is articulated as follows:

$$L(\underline{\mathbf{P}}) = \prod_{i=1}^{n} \lambda(t_i \mid x_i)^{\delta_i} S(t_i \mid x_i), \tag{5.19}$$

where $S(\cdot \mid x_i)$ and $\lambda(\cdot \mid x_i)$ denote the Marginal survival and hazard functions delineated in (5.5). Subsequently, the associated log likelihood function is derived by employing the natural logarithm of $L(\underline{\mathbf{P}})$.

5.3 Numerical Results from Simulations

5.3.1 Based on Weibull Baseline Hazard Function

Taking into account the WBLHF, the log-likelihood function for the parameter vector $\underline{\mathbf{P}} = (\kappa, \rho, \sigma^2, \beta)$ is delineated as follows:

$$\begin{aligned} \log L(\underline{\mathbf{P}}) \;=\; & r \log\left[k\left(1+9\sigma^2\right)\right] \;+\; (k-1)\sum_{i=1}^{n}\delta_i \log t_i \;+\sum_{i=1}^{n}\delta_i x_i^{\intercal}\beta \\ & -\sum_{i=1}^{n}\delta_i \log\left[t_i^k \exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+3\rho^k\sqrt{2\left(1+7\sigma^2\right)}-\rho^k 2\right] \end{aligned}$$

$$+\sum_{i=1}^{n}\delta_i \log\Psi_i+\sum_{i=1}^{n}\log\chi_i+\sum_{i=1}^{n}\log m_i\ , \tag{5.20}$$

where $r=\sum_{i=1}^{n}\delta_i$ denotes the cumulative total of failures,

$$\chi_i=\frac{\rho^k}{3\left\{t_i^k\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\rho^k\left[\tau\left(\sigma^2\right)-2\right]\right\}}\left[\tau\left(\sigma^2\right)-2\right],$$

$$m_i=\frac{D\left(\sigma^2\right)}{1+9\sigma^2}+\frac{\rho^k\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\left(1+9\sigma^2\right)\left\{t_i^k\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\rho^k\left[\tau\left(\sigma^2\right)-2\right]\right\}},$$

and

$$\begin{aligned}\Psi_i &= \rho^k\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\\ &\quad\times\left\{\begin{array}{c}\left[t_i^k\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\rho^k\tau\left(\sigma^2\right)-\rho^k 2\right]D\left(\sigma^2\right)\\ +\rho^k\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\end{array}\right\}^{-1}+1.\end{aligned}$$

The pertinent Maximum Likelihood Estimators (MLE) denoted as $\widehat{\underline{\mathbf{P}}}$ for the parameter vectors $\underline{\mathbf{P}}$ are derived through the maximization of the logarithmic likelihood functions as indicated in equation (5.19). If $\widehat{\underline{\mathbf{P}}}$ lacks a closed analytical form, it becomes imperative to resort to the application of numerical nonlinear optimization methodologies in order to effectively identify and ascertain a solution. These sophisticated optimization techniques are executed through the utilization of the BBsolve package within the R software packages, as articulated by Ravi (2009). In accordance with the

TPLF model in conjunction with the WBLHF, the dataset was generated through simulation processes totaling $N=12{,}000$ iterations; the parameter values were determined as $\kappa=0.85,\rho=0.85,\sigma^2=0.65,\beta_1=0.7$, with sample sizes specified as $n=20,n=40,n=350$, and $n=1000$, along with censoring proportions of 0%, 15%, 35%, and 55%. We computed the mean values of the simulated outcomes pertaining to the maximum likelihood estimators (MXLEs) $\widehat{\kappa},\widehat{\rho},\widehat{\sigma^2},\widehat{\beta_1}$ parameters, along with their Mean Squared Quantitative

Error (MSQE), utilizing the R programming language and the Barzilai-Borwein (BB) optimization algorithm (refer to Ravi (2009)). The results of the simulation are provided in Table 5.1. The outcomes of the simulation are delineated in Table 5.1. The maximum likelihood estimations for the TPLF model employing WBLHF exhibit convergence, as shown in Table 5.1.

Table 5.1: Bias and MSQE of the MXLEs under the WBLHF

n		Bias	MSQE	Bias	MSQE	Bias	MSQE	Bias	MSQE
		0%cens.		15%cens.		35%cens.		55%cens.	
20	ρ	0.86548	0.0499	0.85999	0.0467	0.86374	0.0437	0.85761	0.0435
	κ	0.89245	0.0519	0.87132	0.0512	0.86371	0.0467	0.87660	0.0486
	σ^2	0.67002	0.0486	0.66814	0.0485	0.66648	0.0419	0.68001	0.0415
	β_1	0.79256	0.0432	0.75324	0.0416	0.74381	0.0431	0.74318	0.0398
40	ρ	0.86215	0.0416	0.85462	0.0413	0.86001	0.0400	0.85346	0.0412
	κ	0.86754	0.0483	0.85116	0.0476	0.85344	0.0422	0.86332	0.0406
	σ^2	0.66532	0.0431	0.66004	0.0401	0.65807	0.0376	0.66341	0.0375
	β_1	0.78361	0.0412	0.74198	0.0358	0.73674	0.0402	0.73165	0.0342
350	ρ	0.85673	0.0400	0.851203	0.0356	0.85749	0.0321	0.84778	0.0346
	κ	0.85421	0.0412	0.85100	0.0354	0.85207	0.0396	0.85341	0.0401
	σ^2	0.65432	0.0364	0.65504	0.0359	0.65291	0.0323	0.65127	0.0288
	β_1	0.75000	0.0351	0.71065	0.0241	0.71205	0.0234	0.71138	0.0222
1000	ρ	0.85201	0.0338	0.84896	0.0248	0.85120	0.0264	0.84986	0.0274
	κ	0.851204	0.0308	0.84998	0.0328	0.85110	0.0315	0.85002	0.0200
	σ^2	0.65213	0.0339	0.65128	0.0311	0.64899	0.0275	0.65002	0.0241
	β_1	0.72101	0.0301	0.71158	0.0222	0.69984	0.0215	0.70025	0.0159

5.3.2 Based on Exponential Baseline Hazard Function

Taking into account (EBLHF), the log-likelihood function corresponding to the vector parameter $\underline{\mathbf{P}} = (\lambda, \sigma^2, \beta)$ is delineated by

$$\begin{aligned} \log L(\underline{\mathbf{P}}) &= r\log\left[\lambda\left(1+9\sigma^2\right)\right] + \sum_{i=1}^{n}\delta_i x_i^{\intercal}\beta \\ &\quad -\sum_{i=1}^{n}\delta_i \log\left[\lambda t_i \exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right) + \tau\left(\sigma^2\right) - 2\right] \end{aligned} \tag{5.21}$$

$$+\sum_{i=1}^{n}\delta_i \log \Upsilon_i + \sum_{i=1}^{n} \log \vartheta_i + \sum_{i=1}^{n} \log \Delta_i,$$

where

$$\vartheta_i = \frac{\tau\left(\sigma^2\right)-2}{3\left[\lambda t_i \exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right]},$$

$$\Delta_i = \frac{D\left(\sigma^2\right)}{1+9\sigma^2} + \frac{\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\left(1+9\sigma^2\right)\left[\lambda t_i \exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right]}$$

and

$$\Upsilon_i = \left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\left\{\begin{array}{c}\left[\lambda t_i \exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right]D\left(\sigma^2\right)\\ +\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\end{array}\right\}^{-1}+1.$$

In the context of the TPLF model when combined with the EBLHF, the dataset was subjected to simulation procedures for $N = 12{,}000$ iterations; the parameter values were determined as $\lambda = 0.6, \sigma^2 = 0.5, \beta_1 = 0.9$, with sample sizes specified as $n = 20, n = 40, n = 350$, and $n = 1000$, alongside censoring proportions of 0%, 15%, 35%, and 55%. The mean values of the simulated parameters associated with the MXLEs, specifically $\widehat{\lambda}, \widehat{\sigma^2}, \widehat{\beta_1}$, along with their Mean Squared Quantitative Error (MSQE), were computed utilizing the R software in conjunction with the Barzilai-Borwein (BB) algorithm (refer to Ravi (2009)). The results derived from the simulations are delineated in Table 5.2. The maximum likelihood estimations associated with the TPLF model that incorporates EBLHF demonstrate convergence, as illustrated in Table 5.2.

5.3.3 Based on Gompertz Baseline Hazard Function

Taking into account the GBLHF, the log-likelihood function for the vector parameter $\underline{\mathbf{P}} = (\gamma, \varphi, \sigma^2, \beta)$ is expressed as follows:

Table 5.2: Bias and MSQE of the MXLEs under the EBLHF

n		Bias	MSQE	Bias	MSQE	Bias	MSQE	Bias	MSQE
		0% cens.		15% cens.		35% cens.		55% cens.	
20	λ	0.63514	0.0496	0.62154	0.0437	0.64831	0.0487	0.62354	0.0438
	σ^2	0.55296	0.0435	0.53333	0.0462	0.55001	0.0438	0.53769	0.0426
	β_1	0.96278	0.0431	0.94271	0.0396	0.93265	0.0421	0.93349	0.0354
40	λ	0.62853	0.0417	0.62003	0.0400	0.63497	0.0427	0.61728	0.0411
	σ^2	0.54862	0.0374	0.52481	0.0402	0.53189	0.0476	0.91638	0.0302
	β_1	0.94371	0.0411	0.92648	0.0324	0.91548	0.0416	0.51221	0.0382
350	λ	0.61705	0.0361	0.61965	0.0296	0.62085	0.0355	0.59537	0.0374
	σ^2	0.53719	0.0309	0.51012	0.0367	0.52247	0.0422	0.50252	0.0318
	β_1	0.91187	0.0412	0.90995	0.0219	0.90678	0.0332	0.91203	0.0284
1000	λ	0.61202	0.0302	0.59834	0.0178	0.61207	0.0273	0.59889	0.0300
	σ^2	0.05108	0.0265	0.50067	0.0288	0.51305	0.0374	0.49896	0.0331
	β_1	0.90506	0.0445	0.90046	0.0222	0.89798	0.0227	0.90010	0.0212

$$\begin{aligned}\log L(\underline{\mathbf{P}}) &= r\log\left[\gamma\varphi\left(1+9\sigma^2\right)\right]+\gamma\sum_{i=1}^{n}\delta_i\log t_i+\sum_{i=1}^{n}\delta_i x_i^{\intercal}\beta \\ &\quad -\sum_{i=1}^{n}\delta_i\log\left[\varphi\left[\exp\left(\gamma t_i\right)-1\right]\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+3\gamma\sqrt{2\left(1+7\sigma^2\right)}-2\gamma\right] \\ &\quad +\sum_{i=1}^{n}\delta_i\log\Phi_i+\sum_{i=1}^{n}\log\Pi_i+\sum_{i=1}^{n}\log\zeta_i\ ,\end{aligned} \tag{5.22}$$

where

$$\zeta_i=\frac{\gamma}{3\varphi\left(\sigma^2|t_i\right)}\left[\tau\left(\sigma^2\right)-2\right],$$

$$\Pi_i=\frac{1}{\left(1+9\sigma^2\right)}D\left(\sigma^2\right)+\frac{\gamma}{\left(1+9\sigma^2\right)\varphi\left(\sigma^2|t_i\right)}\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right),$$

and

$$\begin{aligned}\Phi_i &= \gamma\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\\ \times &\left\{\begin{array}{c}\left[\varphi\left[\exp\left(\gamma t_i\right)-1\right]\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+3\gamma\sqrt{2\left(1+7\sigma^2\right)}-2\gamma\right]D\left(\sigma^2\right)\\ +\ \gamma\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\end{array}\right\}^{-1}+1.\end{aligned}$$

In the context of the TPLF model in conjunction with the GBLHF, a total of $N = 12,000$ simulations were conducted. The parameter values were predetermined and established as $\gamma = 0.6, \varphi = 0.35, \sigma^2 = 0.5, \beta_1 = 1.5$, with sample sizes set at $n = 20, n = 40, n = 350$, and $n = 1000$, alongside censoring proportions of 0%, 15%, 35%, and 55%. We executed a computation of the mean values of the simulated parameters associated with the MXLEs, specifically $\widehat{\gamma}, \widehat{\varphi}, \widehat{\sigma^2}, \widehat{\beta_1}$, alongside their respective Mean Squared Quantile Errors (MSQE), employing the R statistical software in conjunction with the Barzilai-Borwein (BB) algorithm (Ravi (2009)). The results derived from the simulation are presented in Table 5.3. The maximum likelihood estimations for the TPLF model incorporating GBLHF demonstrate convergence, as shown in Table 5.3.

Table 5.3: Bias and MSQE of the MXLEs under the GBLHF

n		Bias	MSQE	Bias	MSQE	Bias	MSQE	Bias	MSQE
		0% cens.		15% cens.		35% cens.		55% cens.	
20	γ	0.65548	0.0435	0.64937	0.0325	0.63418	0.0462	0.63854	0.0481
	φ	0.35945	0.0475	0.35962	0.0387	0.35401	0.0475	0.35719	0.0415
	σ^2	0.54612	0.0421	0.53481	0.0489	0.53841	0.0321	0.52214	0.0358
	β_1	1.56382	0.0437	1.54062	0.0384	1.53846	0.0384	1.53048	0.0485
40	γ	0.64381	0.0395	0.62559	0.0305	0.63084	0.0439	0.62443	0.0392
	φ	0.35512	0.0357	0.35462	0.0265	0.35286	0.0385	0.35608	0.0377
	σ^2	0.53816	0.0381	0.52647	0.0435	0.52937	0.0311	0.51473	0.0267
	β_1	1.54371	0.0381	1.52034	0.0332	1.52739	0.0367	1.52271	0.0432
350	γ	0.63894	0.0312	0.61738	0.2384	0.61608	0.0327	0.61850	0.0316
	φ	0.35334	0.0276	0.35210	0.0213	0.35167	0.0241	0.35224	0.0324
	σ^2	0.52743	0.0314	0.52003	0.0412	0.51784	0.0276	0.50734	0.0233
	β_1	1.52496	0.0341	1.51274	0.0251	1.51172	0.0237	1.51092	0.0255
1000	γ	0.61862	0.0211	0.06522	0.0213	0.69665	0.0300	0.60023	0.0275
	φ	0.35206	0.0213	0.35082	0.0135	0.35044	0.0223	0.34995	0.0281
	σ^2	0.51223	0.0311	0.51302	0.0246	0.50937	0.0214	0.50234	0.0200
	β_1	1.51204	0.0276	1.51006	0.0233	1.51062	0.0219	1.49968	0.0201

5.3.4 Based on Pareto Baseline Hazard Function

Taking into account the PBLHF, the log-likelihood function for the vector parameter $\underline{\mathbf{P}} = (\eta, \alpha, \sigma^2, \beta)$ is expressed as

$$\begin{aligned} \log L(\underline{\mathbf{P}}) &= r\log\left[\eta\left(1+9\sigma^2\right)\right] + \sum_{i=1}^{n}\delta_i x_i^{\intercal}\beta \\ &\quad -\sum_{i=1}^{n}\delta_i \log\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right] \\ &\quad -\sum_{i=1}^{n}\delta_i\log\left(\alpha+t_i\right)+\sum_{i=1}^{n}\log\rho_i+\sum_{i=1}^{n}\log\mu_i+\sum_{i=1}^{n}\delta_i\log\Gamma_i, \end{aligned} \tag{5.23}$$

where

$$\rho_i = \frac{\tau\left(\sigma^2\right)-2}{3\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right]},$$

$$\mu_i = \frac{D\left(\sigma^2\right)}{\left(1+9\sigma^2\right)} + \frac{\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)}{\left(1+9\sigma^2\right)\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right]}$$

and

$$\begin{aligned}\Gamma_i &= \left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right) \\ \times &\left\{\begin{array}{c}\left[-\eta\log\left(\frac{\alpha}{\alpha+t_i}\right)\exp\left(x_i^{\intercal}\beta\right)\left(1+9\sigma^2\right)+\tau\left(\sigma^2\right)-2\right]D\left(\sigma^2\right) \\ +\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\end{array}\right\}^{-1}+1.\end{aligned}$$

In the context of the TPLF model integrated with the PBLHF, the data were subjected to simulation procedures $N = 12{,}000$ iterations; we established the parameter values as $\eta = 0.4, \alpha = 0.6, \sigma^2 = 0.5, \beta_1 = 1.7$, alongside sample sizes $n = 20, n = 40, n = 350$ and $n = 1000$, in addition to censoring proportions of 0%, 15%, 35%, and 55%. The mean values of the simulated parameters of the MXLEs, specifically $\widehat{\eta}, \widehat{\alpha}, \widehat{\sigma^2}, \widehat{\beta}_1$, along with their corresponding Mean Squared Error (MSQE), were computed utilizing the R statistical software in conjunction with the Barzilai-Borwein (BB) algorithm as delineated by Ravi (2009). The outcomes derived from the simulation are delineated in Table 5.4. The maximum likelihood estimations for the TPLF model incorporating PBLHF exhibit convergence, as illustrated in Table 5.4.

Table 5.4: Bias and MSQE of the MXLEs under the PBLHF

n		Bias	MSQE	Bias	MSQE	Bias	MSQE	Bias	MSQE
		0% cens.		15% cens.		35% cens.		55% cens.	
20	η	0.46075	0.0466	0.45137	0.0432	0.45002	0.0398	0.44521	0.0452
	α	0.64015	0.0461	0.63174	0.0319	0.64001	0.0318	0.63462	0.0437
	σ^2	0.54832	0.0486	0.55104	0.0482	0.53819	0.0399	0.54468	0.0437
	β_1	1.76034	0.0477	1.74392	0.0451	1.73005	0.0334	1.72938	0.0468
40	η	0.44381	0.0376	0.44382	0.0396	0.43185	0.0321	0.43719	0.0427
	α	0.63176	0.0367	0.62638	0.0278	0.63591	0.0237	0.62649	0.0348
	σ^2	0.53714	0.0431	0.53192	0.0432	0.52619	0.0287	0.52731	0.0316
	β_1	1.75123	0.0395	1.72154	0.0349	1.72419	0.0267	1.71673	0.0427
350	η	0.42864	0.0324	0.42658	0.0324	0.41300	0.0311	0.42635	0.0316
	α	0.62574	0.0324	0.61674	0.0126	0.62649	0.0173	0.61873	0.0222
	σ^2	0.52067	0.0325	0.52230	0.0325	0.52043	0.0125	0.51473	0.0247
	β_1	1.73198	0.0351	1.71708	0.0243	1.71067	0.0213	1.70937	0.0357
1000	η	0.42100	0.0261	0.41873	0.0284	0.41074	0.0284	0.41986	0.0294
	α	0.61346	0.0300	0.60261	0.0124	0.61649	0.0122	0.60936	0.0162
	σ^2	0.51003	0.0301	0.49852	0.0202	0.51170	0.0100	0.50017	0.0233
	β_1	1.71103	0.0307	1.70688	0.0201	1.70032	0.0187	1.70634	0.0251

5.4 Validation of Two-parameter Lindley Frailty Model for Uncensored Data Based on Nikulin–Rao–Robson Test

The N.RR test statistic evaluates the degree to which the statistical model aligns with a specific collection of observations. An extensive assessment known as the N-RR test is applicable for evaluating the predictive adequacy of diverse statistical models, including time series, regression, and survival frameworks. In general, the N.RR test statistic serves as an invaluable resource for conducting statistical analyses and possesses extensive applicability. Its utility is particularly pronounced in the context of model selection, assessment of a model's goodness of fit, and the detection of potential issues associated with a model (for further elucidation, refer to Nikulin (1973a), Nikulin (1973b), Nikulin (1973c), [132] [133] [134] and Rao and Robson (1974) [140]). One of the primary merits of the N-RR test statistic resides in its capacity to identify variations from normality that

alternative statistical methodologies may fail to recognize. Notably, the N.RR test demonstrates an impressive robustness to outliers, which positions it as an effective method for identifying and analyzing data sets that feature extreme values. This positions it as especially useful in a financial context, in which it is vital to recognize and analyze critical circumstances like market recessions and considerable price changes. The following delineates several applications and the significance of the N.RR test statistic:

- The N.RR test statistic serves as a tool for evaluating the adequacy of various statistical models in relation to a common dataset. This methodology aids in the process of model selection by determining the model that yields the most optimal fit to the dataset.

- The N.RR test statistic serves as a tool for evaluating the adequacy of a statistical model in relation to the empirical data. A low N-RR test statistic signifies a beneficial association between the model and the observed data. In contrast, a high value of the N-RR test statistic indicates an inadequate alignment between the model and the empirical data.

- The N.RR test statistic serves as a valuable tool for the identification of outliers within the dataset. Outliers are characterized as data points that exhibit significant deviation from the primary trend of the dataset and may exert a potential impact on the model's alignment. The N.RR test has the capacity to recognize these outliers, thereby facilitating enhancements in the model's overall fit.

- The N.RR test statistic serves as an analytical tool for evaluating deficiencies within a statistical model. A significant value of the N.RR test statistic may suggest that the model is incorrectly specified or that there are underlying issues with the assumptions inherent to the model.

In accordance with the N.RR statistical framework, it is imperative to evaluate the

subsequent null hypothesis.

$$H_0 : \Pr\{z_i \leq z\} = F_{\underline{\mathbf{P}}}(z),\ \ z \in \mathbb{R},\ \ \underline{\mathbf{P}} = (\underline{\mathbf{P}}_1, \underline{\mathbf{P}}_2, \cdots, \underline{\mathbf{P}}_s)^T,$$

Consequently, the N.RR statistic may be formulated as

$$Y^2(\widehat{\underline{\mathbf{P}}}_n) = X_n^2(\widehat{\underline{\mathbf{P}}}_n) + \frac{1}{n}\ell^T(\widehat{\underline{\mathbf{P}}}_n)(\mathbf{I}(\widehat{\underline{\mathbf{P}}}_n) - \mathbf{J}(\widehat{\underline{\mathbf{P}}}_n))^{-1}\ell(\widehat{\underline{\mathbf{P}}}_n),$$

where

$$X_n^2(\underline{\mathbf{P}}) = \left([np_1(\underline{\mathbf{P}})]^{-\frac{1}{2}}[-np_1(\underline{\mathbf{P}}) + \underline{\mathbf{P}}_1], \cdots, [np_b(\underline{\mathbf{P}})]^{-\frac{1}{2}}[-np_b(\underline{\mathbf{P}}) + \underline{\mathbf{P}}_b]\right)^T,$$

and

$$\mathbf{J}(\underline{\mathbf{P}}) = B(\underline{\mathbf{P}})^T B(\underline{\mathbf{P}}),$$

relates to the matrix of information in which

$$B(\underline{\mathbf{P}}) = \left[\frac{1}{\sqrt{p}_i}\frac{\partial}{\partial\mu}(\underline{\mathbf{P}})\right]_{r\times s}|_{(i=1,2,\cdots,b \text{ and } \kappa=1,2,\cdots,s)},$$

and

$$\ell(\underline{\mathbf{P}}) = (\ell_1(\underline{\mathbf{P}}), ..., \ell_s(\underline{\mathbf{P}}))^T \text{ with } \ell_\kappa(\underline{\mathbf{P}}) = \sum_{i=1}^{r}\frac{\underline{\mathbf{P}}_i}{p_i}\frac{\partial p_i(\underline{\mathbf{P}})}{\partial\underline{\mathbf{P}}_\kappa},$$

The $Y^2(\widehat{\underline{\mathbf{P}}}_n)$ statistic possesses $(b-1)$ degrees of freedom (DF) and is associated with the χ^2_{b-1} distribution, wherein the observations are considered. Let $x_1, x_2, \cdots, x_n$ be the data points organized within the disjoint subintervals $\mathbf{I}_1, \mathbf{I}_2, \cdots, \mathbf{I}_b$, where each subinterval is defined as $\mathbf{I}_j =]a_{j,b} - 1; a_{j,b}]$. The limit values for $a_{j,b}$ associated with the intervals $\mathbf{I}_j$ are established through the following methodology.

$$p_j(\underline{\mathbf{P}}) = \int_{a_{j,b}-1}^{a_{j,b}} f_{\underline{\mathbf{P}}}(x)\,dx|_{(\ j=1,2,\cdots,b)},$$

and

$$a_{j,b} = F^{-1}\left(\frac{j}{b}\right)|_{(j=1,\cdots,b-1)}.$$

In numerous instances, the objective of a goodness-of-fit assessment extends beyond ascertaining the compatibility of a specific distribution with the data; it also encompasses the estimation of its parameter values. Simulation analyses can yield valuable understanding regarding the accuracy of parameter estimations across various contexts, thereby guiding the selection of appropriate distributions for future investigations. In conclusion, uncensored simulation analyses conducted within the framework of N.RR statistics serve as a crucial method for assessing and contrasting various probability distributions under controlled conditions. These investigations can yield significant understanding regarding the efficacy of the N.RR assessments across various conditions, and can guide the selection of appropriate distributions for future analytical procedures. Through the application of numerical simulations, we executed a comprehensive examination to validate the assertions present in this study.

In order to validate the null hypothesis H_0, we consequently generated the N.RR statistics pertaining to the TPLF model to ascertain that the sample size is 13000, utilizing simulated samples of sizes $n = 26, n = 40, n = 140, n = 250, n = 600$, and $n = 1200$. Concerning different theoretical levels ($\varepsilon = 0.01, 0.02, 0.04, 0.09$), we calculate the mean of the instances of non-rejection for the null hypothesis. $Y^2 \leq \chi^2_{\varepsilon}(b-1)$. The suitable empirical and theoretical levels are delineated in Table 5.5. It is apparent that a strong correspondence exists between the computed empirical level value and its corresponding theoretical level value. Consequently, we deduce that the suggested test is highly effective for the TPLF distribution.

Table 5.5: Evaluation of the N.RR statistic based on uncensored data for $\varepsilon = 0.01, 0.02, 0.04, 0.09$ and $N = 13000$.

$n \downarrow$ & $\varepsilon \longrightarrow$	$\varepsilon = 0.01$	$\varepsilon = 0.02$	$\varepsilon = 0.04$	$\varepsilon = 0.09$
$n = 26$	0.9924	0.9822	0.9631	0.9120
$n = 40$	0.9916	0.9817	0.9627	0.9116
$n = 140$	0.9914	0.9815	0.9622	0.9110
$n = 250$	0.9906	0.9811	0.9616	0.9108
$n = 600$	0.9904	0.9807	0.9611	0.9104
$n = 1200$	0.9902	0.9804	0.9606	0.9101

5.5 Validation of Two-Parameter Lindley Frailty Model for Censored Data Based on Bagdonavicius–Nikulin Test

As established by Bagdonavicius and Nikulin (2011) [17] and further elaborated by Bagdonavicius et al. (2013) [18], it is feasible to assess the applicability of the TPLF model in scenarios where the parameters remain indeterminate, and the data are censored, in which case the null hypothesis can be articulated as

$$H_0 : F(x) \in F_0 = \left\{ F_0(x, \underline{\mathbf{P}}), x \in R^1,\ \underline{\mathbf{P}} \in \underline{\mathbf{P}} \subset R^s \right\},$$

Let us partition the constrained temporal interval $[0, \tau(\sigma^2)]$ into κ segments, where $\kappa = 1, 2, \cdots, s$, each representing a briefer duration. This delineates the maximum operational timeframe of the investigation and $\mathbf{I}_j = (a_{j-1}, a_{j,b}]$; $0 =< a_{0,b} < a_{1,b} ... < a_{\kappa-1,b} < a_{\kappa,b} = +\infty$. The expected value of $\widehat{a_{j,b}}$ can be articulated as follows, where $x_{(i)}$ denotes the i^{th} element within the sequence of ordered statistics $(x_{(1)},,,x_{(n)})$, and Λ^{-1} signifies the cumulative hazard function and

$$\widehat{a_{j,b}} = \Lambda^{-1}\left((E_{j,X} - \sum_{l=1}^{i-1} \Lambda(x_{(l)}, \underline{\widehat{\mathbf{P}}}))/(n-i+1), \underline{\widehat{\mathbf{P}}} \right), \quad \widehat{a_\kappa} = x_{(n)}|_{(j=1,\ldots,\kappa)},$$

where

$$e_{j,Z} = E_{\kappa}/\kappa \text{ for every } j.$$

$$\Lambda(x,\underline{\mathbf{P}}) = -\ln\left\{\begin{array}{c}\left[\dfrac{\lambda_0(t_i)\,\xi_i\left(1+9\sigma^2\right)}{\Lambda_0(t_i)\,\xi_i\,(1+9\sigma^2)+\tau\left(\sigma^2\right)-2}\right]\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\\ \times\left\{\Lambda_0\left(\xi_i,\sigma^2\right)D\left(\sigma^2\right)+\left[\tau\left(\sigma^2\right)-2\right]C\left(\sigma^2\right)\right\}^{-1}+1\end{array}\right\},$$

and

$$E_{j,Z} = (n-i+1)\Lambda(\widehat{a_{j,b}},\widehat{\underline{\mathbf{P}}})+\sum_{l=1}^{i-1}\Lambda(x_{(l)},\widehat{\underline{\mathbf{P}}}) = \sum_{i:z_i>a_{j,b}}(\Lambda(a_{j,b}\wedge z_i,\widehat{\underline{\mathbf{P}}})-\Lambda(a_{j-1},\widehat{\underline{\mathbf{P}}}), E_{\kappa} = \sum_{i=1}^{n}\Lambda(z_i,\widehat{\underline{\mathbf{P}}}).$$

The functions $a_{j,b}$ pertaining to random datasets and the $e_{j,Z}$ corresponding to the specified periods κ exhibit equivalence in the predicted failure rates. statistically observed data $Y_n^2 = \mathbf{Z}^T\widehat{\mathbf{S}}^{-1}\mathbf{Z}$, where $\mathbf{Z} = (Z_1,...,Z_{\kappa})^x$, $Z_j = \frac{1}{\sqrt{n}}(\mathbf{W}_{j,Z}-e_{j,Z})|_{(\ j=1,2,...,\kappa)}$ and $\mathbf{W}_{j,Z}$ may be employed in the evaluation of a hypothesis, as it indicates the cumulative total of failures registered within these particular durations. The components of the Bg.N test statistic

$$Y_n^2 = \sum_{j=1}^{\kappa}\frac{1}{\mathbf{W}_{j,Z}}(\mathbf{W}_{j,Z}-e_{j,Z})^2+\mathbf{D}_{W,G},$$

where

$$\begin{aligned}
\mathbf{D}_{W,G} &= \widehat{\mathbf{V}}^T\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{V}},\widehat{\mathbf{S}}^{-1} = \widehat{\mathbf{B}}^{-1}+\widehat{\mathbf{M}}^{-1}\widehat{\mathbf{B}}^T\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{M}}\widehat{\mathbf{B}}^{-1},\\
\widehat{\mathbf{G}} &= [\widehat{g}_{ll'}]_{s\times s} = \widehat{i}-\widehat{\mathbf{M}}\widehat{\mathbf{B}}^{-1}\widehat{\mathbf{M}}^x,\\
\widehat{\mathbf{M}}_{lj} &= \frac{1}{n}\sum_{i:z_i\in\mathbf{I}_j}\rho_i\frac{\partial}{\partial\underline{\mathbf{P}}}\ln\left[\lambda_{i,\underline{\mathbf{P}}}(z_i)\right], \mathbf{W}_{j,Z} = \sum_{i:z_i\in\mathbf{I}_j}\rho_i,\ \widehat{\mathbf{B}}_j = n^{-1}\mathbf{W}_{j,Z},\\
\widehat{\mathbf{V}}_l &= \sum_{j=1}^{\kappa}\widehat{\mathbf{M}}_{lj}\widehat{\mathbf{B}}_j^{-1}\mathbf{Z}_j|l,l' = 1,...,s,\\
\widehat{i}_{ll'} &= n^{-1}\sum_{i=1}^{n}\rho_i\frac{\partial}{\partial\underline{\mathbf{P}}_l}\ln\left[\lambda_{i,\underline{\mathbf{P}}}(z_i)\right]\frac{\partial}{\partial\underline{\mathbf{P}}_{l'}}\ln\left[\lambda_{i,\widehat{\underline{\mathbf{P}}}}(z_i)\right],
\end{aligned}$$

and

$$\widehat{g}_{ll'} = \widehat{i}_{ll'} - \sum_{j=1}^{\kappa} \widehat{\mathbf{M}}_{lj} \widehat{\mathbf{M}}_{l'j} \hat{A}_j^{-1},$$

and

$$\widehat{\mathbf{M}}_{lj} = \frac{1}{n} \sum_{i:z_i \in \mathbf{I}_j} \rho_i \frac{\partial}{\partial \underline{\mathbf{P}}} \ln \left[\lambda_{i,\widehat{\underline{\mathbf{P}}}}(z_i) \right].$$

The utilization of Bg.N statistics in censored simulation analyses constitutes a vital technique for the assessment and comparison of diverse probability distributions in the context of censored datasets. These investigations may yield significant understandings regarding the efficacy of the Bg.N assessments across various forms and degrees of censoring, thereby guiding determinations regarding the appropriate distribution to employ for future analyses. The objective is to ensure that the generated sample ($N = 13000$) will experience censorship at a rate of 24% and that the degrees of freedom will equal 5. In order to evaluate the conformity of the sample with the null hypothesis of the TPLF model, interval grouping methodologies will be employed. In order to analyze different theoretical levels, we compute the mean of the non-rejection values of the null hypothesis. $(\varepsilon = 0.01, 0.02, 0.04, 0.09)$, where $Y^2 \leq \chi^2_{\varepsilon}(r-1)$. The comparative analysis of the theoretical and empirical levels is presented in Table 5.6, which illustrates the degree of association between the identified empirical level and the corresponding theoretical level. We deduce that the specialized assessment is optimally aligned with the TPLF model as a result.

Table 5.6: Evaluation of the Bg.N statistic based on censored data for $\varepsilon = 0.01; 0.02; 0.05; 0.1$ and $N = 13000$

$n \downarrow$ & $\varepsilon \longrightarrow$	$\varepsilon = 0.01$	$\varepsilon = 0.02$	$\varepsilon = 0.04$	$\varepsilon = 0.09$
$n = 26$	0.9924	0.9819	0.9631	0.9120
$n = 40$	0.9920	0.9816	0.9625	0.9114
$n = 140$	0.9915	0.9807	0.9619	0.9111
$n = 350$	0.9911	0.9805	0.9613	0.9108
$n = 600$	0.9906	0.9804	0.9608	0.9103
$n = 1200$	0.9904	0.9801	0.9604	0.9101

It can be inferred from these results that the empirical significance level of the Y_n^2 aligns with the theoretical level of the chi-square distribution concerning degrees of freedom, which corresponds to the statistical level at which significance is established. The data that are censored obtained from the TPLF distribution can, therefore, be adequately modeled utilizing the proposed test, based on this fact.

5.6 An application Based on Emergency Care Data

The emergency unit of the hospital affiliated with a public health organization provided authentic data that were gathered during the month of March 2023, and these data were employed in the present research. The objective of this investigation was to analyze, within a cohort of patients receiving medical treatment at the department, the relationship between diverse clinical attributes and outcomes in the emergency department. The necessary authorizations were assured, and ethical principles were upheld during the data gathering process. The dataset encompassed 30 distinct individuals, each serving as a singular observation. A total of six distinct variables were documented for each participant: chronological age (years), systolic and diastolic blood pressure (mmHg), blood glucose concentration (mg/dL), heart rate (BPM), and peripheral oxygen saturation (SaO 2%). In order to guarantee the precision and integrity of the gathered data, stringent measures were implemented throughout the data collection procedure. This required the precise recording of patient information, conformity with established measurement protocols, and periodic quality assessments to identify any absent data or irregularities. This dataset is useful for investigating the associations between clinical variables and emergency room outcomes due to the accurate data collection method and the variety of the patient population.

Through an extensive evaluation of the TPLF model distribution, we can determine its validity and applicability by assessing the goodness-of-fit and its efficacy in accurately representing the observed trends and variations within emergency care data. We present the point estimates corresponding to each fitted TPLF model utilizing Weibull, Gompertz,

and Pareto baseline hazard functions. Among these models, the most suitable one for the data is identified through the application of the modified chi-squared test, as delineated by Bagdonavicius and Nikulin (2011) [17].

5.6.1 Evaluation of Two-Parameter Lindley Frailty Model Based on Weibull Baseline Hazard Function

Taking into account that the dataset previously discussed conforms to the TPLF model alongside the WBLHF distribution, and utilizing the R statistical programming (particularly the BB package), the maximum likelihood estimations for the parameter vector $\underline{\mathbf{P}}$ are delineated as

$$\begin{aligned}
\widehat{\kappa} &= 0.84965, \widehat{\rho} = 0.831994, \widehat{\sigma^2} = 1.019259, \\
\widehat{\beta_1} &= 0064875, \widehat{\beta_2} = -0.21048, \widehat{\beta_3} = -0.61541, \\
\beta_4 &= 0.49358, \widehat{\beta_5} = -0245174, \widehat{\beta_6} = 0.92547.
\end{aligned}$$

In the context of censored data, we consider, for instance, 5 intervals ($r = 5$) corresponding to the number of classes, a suggestion introduced by Bagdonaviius and Nikulin (2011) [17]. The estimated Fisher information matrix $I\left(\widehat{\underline{\mathbf{P}}}\right)$ consists of the subsequent components

$$\mathrm{I}\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix}
1.09654 & 2.15048 & 0.514872 & -5.91254 & 2.00215 & 1.09857 & 0.61472 & 3.00021 & 0.74581 \\
 & 0.65842 & -6.21547 & 0.53251 & -1.00248 & 2.02188 & -7.15482 & 0.33615 & 1.24182 \\
 & & 0.19547 & 0.00240 & 1.02548 & -6.32514 & -0.06254 & 9.32541 & 0.61245 \\
 & & & 1.00245 & 0.37948 & 0.12548 & 3.21547 & -5.00218 & 0.00097 \\
 & & & & 0.32458 & -4.12572 & 1.02458 & 0.95774 & 0.84752 \\
 & & & & & 3.00218 & -6.21542 & 0.900014 & 7.00015 \\
 & & & & & & 0.08457 & 2.00978 & 3.02157 \\
 & & & & & & & 0.85475 & -1.01021 \\
 & & & & & & & & 0.17548
\end{pmatrix}.$$

Following this, we determine the value of the test statistic to be $Y_n^2 = 9.120054$. The critical value is $\chi^2_{0.05}(4) = 9.488 > Y_n^2$. This dataset can be appropriately fitted by our proposed

TPLF model incorporating WBLHF.

5.6.2 Evaluation of Two-parameter Lindley Frailty Model Based on Exponential Baseline Hazard Function

Presuming that the dataset adheres to the TPLF model under an EBLHF specification, the maximum likelihood estimations of the parameter vector $\underline{\mathbf{P}}$ are derived utilizing the R statistical programming (BB package) and are subsequently detailed as follows

$$\begin{aligned}
\widehat{\lambda} &= 0.6254, \widehat{\sigma^2} = 1.02548, \\
\widehat{\beta_1} &= 3.0254, \widehat{\beta_2} = -9.03251, \widehat{\beta_3} = -0.61587, \\
\beta_4 &= 1.0254, \widehat{\beta_5} = 2.95014, \widehat{\beta_6} = -2.03215.
\end{aligned}$$

Following Bagdonavicius and Nikulin (2011) [17], the censored data are grouped into five classes ($r = 5$). Based on this classification, the elements of the estimated Fisher information matrix $I\left(\widehat{\underline{\mathbf{P}}}\right)$ are presented below

$$\mathrm{I}\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix}
1.63254 & -6.32517 & 2.61502 & -8.3265 & 1.02541 & 0.32514 & 0.96584 & 1.06025 \\
 & 2.95312 & 5.00002 & 1.02543 & 1.03268 & 1.92354 & -7.0095 & -10.7658 \\
 & & 0.88321 & 2.6157 & 3.00002 & 1.20451 & 2.03251 & 2.61547 \\
 & & & 1.54875 & 6.32514 & 3.26514 & 1.02547 & 0.00214 \\
 & & & & 2.00004 & -5.3268 & -4.7474 & 2.00315 \\
 & & & & & 0.96584 & 2.30142 & 1.02547 \\
 & & & & & & 1.02547 & 4.32510 \\
 & & & & & & & 0.32014
\end{pmatrix}$$

Subsequently, we determine the value of the test statistic, denoted as $Y_n^2 = 8.80451$. The critical value is represented by $\chi^2_{0.05}(4) = 9.488 > Y_n^2$. This dataset can be effectively accommodated by the proposed TPLF model employing EBLHF in a suitable manner.

5.6.3 Evaluation of the Two-parameter Lindley Frailty Model Based on Gompertz Baseline Hazard Function

Assuming that the dataset adheres to the TPLF model under a GBLHF specification, the maximum likelihood estimations of the parameter vector $\underline{\mathbf{P}}$ are derived utilizing the R statistical programming (BB package) and are subsequently detailed as follows

$$\begin{aligned}
\widehat{\gamma} &= 1.023014, \widehat{\varphi} = 0.98351, \widehat{\sigma^2} = 1.120034, \\
\widehat{\beta_1} &= 0.963514, \widehat{\beta_2} = 0.845271, \widehat{\beta_3} = -2.61847, \\
\widehat{\beta_4} &= 0.530018, \widehat{\beta_5} = -1.02947, \widehat{\beta_6} = 0.125437.
\end{aligned}$$

We use the approximated Fisher matrix with intervals of $r = 5$, which is denoted as

$$\mathrm{I}\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix}
0.93254 & 2.02154 & 0.21547 & 1.09587 & -2.00347 & 2.15427 & -8.06254 & -5.00214 & 0.61542 \\
 & 1.09651 & 1.09658 & -4.21571 & 0.61547 & 0.12548 & 1.23565 & 0.19574 & 1.02659 \\
 & & 0.85247 & 0.02154 & 1.03254 & -4.03251 & -6.21541 & 0.21547 & 1.02558 \\
 & & & 1.63254 & 0.91547 & 1.02548 & -4.02158 & 2.02154 & -7.02154 \\
 & & & & 2.00314 & -5.03268 & 4.02158 & -8.5479 & -4.01924 \\
 & & & & & 0.86592 & 0.21547 & 0.61547 & 0.2158 \\
 & & & & & & 0.46215 & 1.09754 & 0.21547 \\
 & & & & & & & 1.00985 & -1.00025 \\
 & & & & & & & & 0.36251
\end{pmatrix},$$

Subsequently, we proceed to compute the value of the statistic proposed by Bagdonavius and Nikulin (2011) [17], $Y_n^2 = 8.123048$. For various specified significance levels: $\alpha = 5\%$ and $\alpha = 10\%$, we ascertain $Y^2 < \chi^2_{0.05}(4) = 9.488$ and $Y_n^2 < \chi^2_{0.1}(5-1)$ respectively. Consequently, we deduce that the emergency medical data aligns with our suggested TPLF model incorporating GBLHF.

5.6.4 Evaluation of Two-parameter Lindley Frailty Model Based on Pareto Baseline Hazard Function

Provided that the dataset conforms to the TPLF model within a GBLHF specification, the maximum likelihood estimations of the parameter vector $\underline{\mathbf{P}}$ are computed employing the R statistical programming (BB package) and are subsequently articulated as follows.

$$\begin{aligned}
\widehat{\eta} &= 0.09584, \widehat{\alpha} = 1.24051, \widehat{\sigma^2} = 0.89574, \\
\widehat{\beta_1} &= 0.32658, \widehat{\beta_2} = 0.19487, \widehat{\beta_3} = -0.613548, \\
\beta_4 &= -2.164957, \widehat{\beta_5} = -1.94378, \widehat{\beta_6} = 0.379138.
\end{aligned}$$

Using five intervals $r = 5$, the estimated Fisher information matrix is expressed as follows

$$\mathrm{I}\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix}
0.93518 & 1.02547 & 0.321874 & -2.06587 & 1.02368 & 1.09557 & -2.45871 & 0.19385 & -3.21574 \\
 & 0.64875 & 1.026589 & 1.02547 & -8.02157 & -9.12547 & -4.02154 & 3.12574 & -6.21542 \\
 & & 0.61847 & 0.84753 & 2.03254 & -12.0214 & 0.12547 & 0.95847 & 4.02157 \\
 & & & 0.46158 & -3.16587 & -4.1257 & -4.91578 & 0.12548 & 1.9658 \\
 & & & & 1.02458 & 1.84576 & 1.3258 & 0.002157 & 1.21547 \\
 & & & & & 0.93784 & -2.12547 & -1.29568 & -2.02145 \\
 & & & & & & 1.19325 & -3.21547 & 0.23515 \\
 & & & & & & & 0.84571 & 1.55524 \\
 & & & & & & & & 1.90542
\end{pmatrix}$$

We then proceed to calculate the value of the Bagdonavicius and Nikulin (2011) [17] statistic, $Y_n^2 = 7.23197$. For varying significance levels: $\alpha = 5\%$ and $\alpha = 10\%$, we determine $Y_n^2 < \chi^2_{0.05}(5-1) = 9.488$ and $Y_n^2 < \chi^2_{0.1}(4) = 7.779$, respectively. Consequently, we conclude that the emergency care data aligns effectively with our suggested TPLF model incorporating PBLHF.

5.7 A Heart Attack Dataset Application

This dataset is categorized as multivariate, which denotes the analysis of multivariate numerical data that incorporates or offers a variety of unique mathematical or statistical variables. The collected data comprises 76 covariates; in our investigation, we have employed 5 specific covariates, which include: age, resting blood pressure, serum cholesterol levels, maximum heart rate attained, and oldpeak: ST segment depression elicited by exercise in comparison to rest. A principal objective of this dataset is to determine, by employing the characteristics provided by the patient, the probability of the individual receiving a diagnosis of heart disease. An additional experimental investigation encompasses the identification of the individual and the extraction of diverse insights from the dataset, which may contribute to a more profound comprehension of the subject. The dataset was developed by the Hungarian Institute of Cardiology, as referenced at https://doi.org/10.24432/C52P4X. Through the assessment of the goodness-of-fit related to the TPLF model distribution, we are able to determine the validity and applicability of this distribution by investigating its potential to precisely illustrate the observed trends and variability in heart attack data. Point estimates are reported for each fitted model, including the TPLF model with Weibull, exponential, Gompertz, and Pareto baseline hazard functions. To identify the most appropriate model for the data, the modified chi-squared test is applied, following the approach of Bagdonavicius and Nikulin (2011) [17].

5.7.1 Evaluation of Two Parameter Lindley Frailty Model under Weibull Baseline hazard function

The maximum likelihood estimations of the parameter vector $\underline{\mathbf{P}}$ are derived utilizing R programming (BB package), based on the assumption that the data aligns with the TPLF model incorporating WBLHF, and are presented as follows

$$\widehat{\kappa} = 1.00245, \widehat{\rho} = 0.61472,$$

$$\begin{aligned}\widehat{\sigma^2} &= 1.00214, \widehat{\beta_1} = -2.02154,\\ \widehat{\beta_2} &= 3.00215, \widehat{\beta_3} = -4.02875,\\ \beta_4 &= 1.023548, \widehat{\beta_5} = -1.23487.\end{aligned}$$

In accordance with Bagdonavicius and Nikulin (2011) [17], we categorize the censored data into $r = 8$ distinct classes. The components of the resultant estimated Fisher information matrix $I\left(\widehat{\underline{\mathbf{P}}}\right)$ are delineated below:

$$\mathrm{I}\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix} 2.003254 & 4.15785 & 0.15478 & -3.62501 & 0.00217 & -8.12544 & 6.00985 & 1.21545 \\ & 1.96542 & -6.30214 & 1.02458 & 3.11241 & 2.13547 & 1.75845 & -8.00002 \\ & & 0.96584 & 3.20145 & -2.15347 & 0.95135 & 1.54863 & -0.96584 \\ & & & 2.965847 & 8.215478 & -4.00215 & 12.3518 & 6.00214 \\ & & & & 3.021457 & 2.15475 & -5.88547 & -9.32514 \\ & & & & & 1.92547 & 0.35748 & 12.2514 \\ & & & & & & 2.11045 & -17.2152 \\ & & & & & & & 0.65847 \end{pmatrix},$$

The value of the test statistic is subsequently computed to be $Y_n^2 = 13.58497$. A critical observation is that $\chi^2_{0.05}(7) = 14.0689 > Y_n^2$. Therefore, the dataset can be accurately modeled using our proposed TPLF model with WBLHF.

5.7.2 Evaluation of Two parameter Lindley Frailty Model under Exponential Baseline Hazard Function

The maximum likelihood estimations of the parameter vector $\underline{\mathbf{P}}$ are obtained through R programming (BB package), predicated on the premise that the dataset conforms to the TPLF model that includes EBLHF, and are delineated as follows

$$\begin{aligned}\widehat{\lambda} &= 2.12501, \widehat{\sigma^2} = 1.02102,\\ \widehat{\beta_1} &= 0.2154, \widehat{\beta_2} = -9.3258, \widehat{\beta_3} = 1.95201,\\ \beta_4 &= -10.8124, \widehat{\beta_5} = 2.61024.\end{aligned}$$

Consistent with the findings of Bagdonavicius and Nikulin (2011) [17], we classify the censored data into $r = 8$ separate classes. The elements of the subsequently derived estimated Fisher information matrix $I\left(\widehat{\underline{\mathbf{P}}}\right)$ are specified below:

$$\mathrm{I}\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix} 1.32054 & -7.92518 & -9.3251 & 2.95147 & 1.92358 & 1.32547 & 3.00214 \\ & 0.26531 & -7.1658 & 1.32547 & 0.21574 & 0.96325 & 4.9513 \\ & & 1.92543 & -10.3201 & 5.0001 & 1.42152 & 3.0302 \\ & & & 2.30154 & 1.11124 & 0.85647 & 1.95487 \\ & & & & 3.00002 & 1.44475 & -7.9514 \\ & & & & & 1.20541 & 2.0215 \\ & & & & & & 2.15482 \end{pmatrix},$$

The value of the test statistic is subsequently determined to be $Y_n^2 = 12.95682$. A crucial observation is that $\chi^2_{0.05}(7) = 14.0689 > Y_n^2$. This dataset can be adequately modeled using our proposed TPLF model in conjunction with EBLHF.

5.7.3 Evaluation of Two-parameter Lindley Frailty Model under Gompertz Baseline Hazard Function

The estimations of the parameter vector $\underline{\mathbf{P}}$ utilizing the maximum likelihood method are derived via R programming (BB package), based on the assumption that the dataset adheres to the TPLF model, which encompasses GBLHF, and are articulated as follows

$$\begin{aligned} \widehat{\gamma} &= 1.36001, \widehat{\varphi} = 1.24801, \widehat{\sigma^2} = 1.03452, \\ \widehat{\beta_1} &= -3.51204, \widehat{\beta_2} = -4.671305, \widehat{\beta_3} = 1.30265, \\ \widehat{\beta_4} &= 2.101036, \widehat{\beta_5} = 1.063254. \end{aligned}$$

We utilize the approximated Fisher matrix corresponding to $r = 8$ intervals, which is represented as

$$I\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix} 0.965847 & -5.02147 & -9.21578 & 1.95847 & 3.21542 & -21.31544 & 0.93548 & 2.00031 \\ & 3.26150 & 2.301245 & -8.02547 & 1.362548 & 0.965842 & 1.20154 & 1.99658 \\ & & 2.61354 & 1.25698 & 0.15472 & -7.95382 & -12.4512 & 2.00001 \\ & & & 2.165847 & 1.952014 & 1.02154 & -9.32518 & 1.02458 \\ & & & & 1.92546 & 2.035214 & -12.32548 & 3.1102 \\ & & & & & 0.84571 & 4.03528 & -8.3219 \\ & & & & & & 4.00621 & 1.02326 \\ & & & & & & & 0.88827 \end{pmatrix},$$

Subsequently, we calculate the statistic as presented in Bagdonaviius and Nikulin (2011) [17]: $Y_n^2 = 11.684002$. For several critical levels at $\alpha = 5\%$, we find that $Y^2 < \chi^2_{0.05}(7) = 14.0689$. Consequently, we deduce that the data pertaining to emergency care aligns with our proposed TPLF model utilizing the GBLHF.

5.7.4 Evaluation of Two parameter Lindley Frailty Model under Pareto Baseline Hazard Function

The parameter vector $\underline{\mathbf{P}}$ estimations, obtained through the maximum likelihood approach, are computed using R programming (BB package). This computation is predicated on the premise that the dataset conforms to the TPLF model, which includes PBLHF, and is delineated as follows.

$$\begin{aligned} \widehat{\eta} &= 1.00254, \widehat{\alpha} = 1.36254, \widehat{\sigma^2} = 0.88695, \\ \widehat{\beta_1} &= 0.96584, \widehat{\beta_2} = -0.93747, \widehat{\beta_3} = -0.63625, \\ \beta_4 &= 1.02547, \widehat{\beta_5} = -9.002547. \end{aligned}$$

The elements of the estimated Fisher information matrix, derived through the utilization of $r = 8$ intervals, are delineated as follows:

$$I\left(\widehat{\underline{\mathbf{P}}}\right) = \begin{pmatrix} 0.236584 & -9.32514 & -11.2548 & -13.6258 & 2.15471 & 3.26587 & 2.15473 & 1.99685 \\ & 2.05487 & -12.3254 & 1.95324 & 4.00215 & 3.02130 & 1.95684 & -7.3184 \\ & & 1.92564 & 0.902457 & -4.9515 & 1.95684 & 2.00514 & 2.15043 \\ & & & 0.32678 & 2.2223 & 4.00125 & 3.12022 & 1.84502 \\ & & & & 3.12045 & 3.01254 & 1.90547 & -9.8124 \\ & & & & & 1.39584 & -8.8965 & 1.35486 \\ & & & & & & 1.32547 & 0.96584 \\ & & & & & & & 2.91573 \end{pmatrix}$$

Thereafter, we calculate the Bagdonaviius and Nikulin (2011) [17] statistic: $Y_n^2 = 11.57482$. For alternative significance levels of $\alpha = 5\%$, we observe that $Y_n^2 < \chi^2_{0.05}(7) = 14.0689$. Consequently, we deduce that the emergency care data aligns with our proposed TPLF model incorporating PBLHF.

Conclusions and perspectives

In this thesis, we introduced a new frailty model designed to address the challenges posed by unobserved heterogeneity in survival data. Building upon the two-parameter Lindley (TPL) distribution, we constructed the two-parameter Lindley frailty (TPLF) model and derived its principal analytical components, including its Laplace transform, marginal survival function, and hazard function. These developments offer a mathematically tractable framework for modeling heterogeneity while retaining the flexibility necessary for practical applications. Several baseline hazard functions namely the Gompertz, Weibull, exponential, and Pareto models were incorporated to assess the versatility of the TPLF model across a range of survival structures. A comprehensive set of simulation studies was carried out to investigate the finite-sample behavior of the maximum likelihood estimators associated with the TPLF model. The results confirmed that the estimators exhibit favorable convergence properties under varying degrees of censoring and across different sample sizes. These findings provide strong evidence of the robustness and stability of the proposed model in both small and large samples. Furthermore, to evaluate the model's adequacy, we developed a modified chi-squared goodness-of-fit test by integrating the statistics of Nikulin Rao Robson and the approach proposed by Bagdonavicius and Nikulin. The modified test was shown to be effective for both complete and censored survival datasets, offering a reliable tool for assessing the fit of frailty-based models. The performance of the proposed statistical test and the TPLF model was further evaluated through empirical analysis using newly collected emergency care data from an Algerian hospital and a heart attack dataset. These real-world applications demonstrated that the TPLF model, combined with Gompertz, Weibull, exponential, and Pareto baseline hazard functions, achieved an excellent fit to the observed survival patterns. According to the criteria of Bagdonavicius and Nikulin, the proposed goodness-of-fit test successfully validated the TPLF model under censoring. These results confirm the model's ability to

capture the latent heterogeneity inherent in clinical survival data and highlight its advantages over classical frailty models. Overall, the theoretical developments, simulation experiments, and empirical findings presented in this thesis collectively affirm the value of the TPLF model as a flexible and efficient alternative for frailty-based survival analysis. The model accommodates diverse censoring mechanisms, provides stable parameter estimation, and offers improved adaptability to real-world datasets. Nonetheless, several avenues for future research remain open. Extensions of the TPLF distribution to incorporate covariate-dependent frailty, exploration of Bayesian estimation techniques, or the study of the model's behavior under more complex censoring schemes would all enrich its applicability. In summary, the TPLF model constitutes a significant contribution to the statistical analysis of survival data, offering a powerful tool for capturing and interpreting unobserved heterogeneity across a wide range of applied domains. As a natural extension of this work, future research will focus on developing shared frailty versions of the TPLF model to accommodate clustered or correlated survival data, such as family-based or hospital-level structures. Investigating the behavior of the model under mixed censoring schemes combining right, left, and interval censoring will also broaden its applicability to more complex real-world scenarios. Such extensions would enhance the flexibility of the proposed framework and strengthen its relevance in biomedical and reliability studies. These perspectives constitute promising directions for advancing frailty modeling and improving the interpretation of heterogeneous survival processes. []

Bibliography

[1] Aalen, O. (1978). Nonparametric inference for a family of counting processes. The Annals of Statistics, 701-726.

[2] Aalen, O. O. (1987). Two examples of modelling heterogeneity in survival analysis. Scandinavian journal of statistics, 19-25.

[3] Aalen, O .O. (1988). Heterogeneity in Survival Analysis. Statistics in Medicine ,7, 1121-1137.

[4] Aalen, O.O. (1992). Modelling Heterogeneity in Survival Analysis by the Compound Poisson Distribution. Annals of Applied Probability, 4 (2), 951-972.

[5] Aalen, O.O, Tretli, S. (1999). Analysing incidence of testis cancer by means of a frailty model. Cancer Causes and Control ,10, 285-292.

[6] Abadir, K. M. (2005). The mean-median-mode inequality: counterexamples. Econometric Theory, 21(2), 477-482.

[7] Abbring, J. H., and Van Den Berg, G. J. (2007). The unobserved heterogeneity distribution in duration analysis. Biometrika, 94(1), 87-99.

[8] Aitchison, J., and Brown, J. A. (1957). The lognormal distribution. Cambridge University, 980.

[9] Al-Essa, L. A., Eliwa, M. S., El-Morshedy, M., Alqifari, H., and Yousof, H. M. (2023). Flexible extension of the lomax distribution for asymmetric data under different failure rate profiles: Characteristics with applications for failure modeling and service times for aircraft windshields. Processes, 11(7), 2197.

[10] Al-Zahrani, B., and Gindwan, M. (2015). Estimating the parameter of the Lindley distribution under progressive type-II censored data. Electronic Journal of Applied Statistical Analysis, 8(1).

[11] Alizadeh, M., Afshari, M., Ranjbar, V., Merovci, F. and Yousof, H. M. (2023). A novel XGamma extension: applications and actuarial risk analysis under the reinsurance data. São Paulo Journal of Mathematical Sciences, 1-31.

[12] Almeida, M. P., Paixao, R. S., Ramos, P. L., Tomazella, V., Louzada, F., and Ehlers, R. S. (2020). Bayesian non-parametric frailty model for dependent competing risks in a repairable systems framework. Reliability Engineering System Safety, 204, 107145.

[13] Andersen, P. K., Keiding, N. (2012). Interpretability and importance of functionals in competing risks and multistate models. Statistics in medicine, 31(11-12), 1074-1088.

[14] Anderson, J. E., Louis, T. A. (1995). Survival analysis using a scale change random effects model. Journal of the American Statistical Association, 90(430), 669-679.

[15] Bagdonavicius, V., Nikulin, M. (1997). Accelerated life testing when a process of production is unstable. Statistics probability letters, 35(3), 269-275.

[16] Bagdonavicius, V., Nikulin, M. (2001). Accelerated life models: modeling and statistical analysis. Chapman and Hall/CRC.

[17] Bagdonavicius, V., and Nikulin, M. (2011). Chi-squared goodness-of-fit test for right censored data.International Journal of Applied Mathematics and Statistics, 24, 30-50

[18] Bagdonavicius, V., Levuliene, R., J., and Nikulin, M. (2013). Chi-squared goodness-of-fit tests for parametric accelerated failure time models. Communications in Statistics-Theory and Methods, 42(15), 2768-2785.

[19] Bakouch, H. S., Al-Zahrani, B. M., Al-Shomrani, A. A., Marchi, V. A., Louzada, F. (2012). An extended Lindley distribution. Journal of the Korean statistical society, 41(1), 75-85.

[20] Balakrishnan, N.Peng., Y. (2006). Generalized gamma frailty model. Stat Med, 25(16), 2797-2816.

[21] Bandyopadhyay, D., Basu, A. P. (1990). On a generalization of a model by Lindley and Singpurwalla. Advances in Applied Probability, 22(2), 498-500.

[22] Barlow, R. B. (1975). Statistical theory of reliability and life testing. Holt.

[23] Bennett, S. (1983). Log-logistic regression models for survival data. Journal of the Royal Statistical Society Series C: Applied Statistics, 32(2), 165-171.

[24] Breslow, N. (1974). Covariance analysis of censored survival data. Biometrics, 89-99.

[25] Bretagnolle, J. Huber-Carol, C. (1988). Effects of omitting covariates in Cox's model for survival data. Scandinavian journal of statistics, 125-138.

[26] Cai, J., Prentice, R. L. (1995). Estimating equations for hazard ratio parameters based on correlated failure time data. Biometrika, 82(1), 151-164.

[27] Chang, S. H. (2004). Estimating marginal effects in accelerated failure time models for serial sojourn times among repeated events. Lifetime Data Analysis, 10(2), 175-190.

[28] Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65(1), 141-151.

[29] Clayton, D., Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. Journal of the Royal Statistical Society: Series A (General), 148(2), 82-108.

[30] Collett, D. (1994). Modelling survival data. In Modelling survival data in medical research (pp. 53-106). Springer US.

[31] Collett, D. (2015). Modelling Survival Data in Medical Research. Chapman and Hall/CRC.

[32] Congdon, P. (1995). Modelling frailty in area mortality. Statistics in medicine, 14(17), 1859-1874.

[33] Cook, R. D., Johnson, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. Journal of the Royal Statistical Society Series B: Statistical Methodology, 43(2), 210-218.

[34] Cox DR. (1972). Regression models and life-tables. J Roy Stat Soc: Ser B (Methodol), 34(2), 187-202.

[35] Cox, D. R. (1975). Partial likelihood. Biometrika, 62(2), 269-276.

[36] Cox, D.R., Oakes, D. (1984) Analysis of survival data. Chapman Hall, London.

[37] Cui, S., Sun, Y. (2004). Checking for the gamma frailty distribution under the marginal proportional hazards frailty model. Statistica Sinica, 249-267.

[38] Dos Santos, D. M., Davies, R. B., Francis, B. (1995). Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. Journal of statistical planning and inference, 47(1-2), 111-127.

[39] Duchateau, L., Janssen, P., Lindsey, P., Legrand, C., Nguti, R., Sylvester, R. (2002). The shared frailty model and the power for heterogeneity tests in multicenter trials. Computational Statistics Data Analysis, 40(3), 603-620.

[40] Duchateau, L., Janssen, P., Kezic, I., Fortpied, C. (2003). Evolution of recurrent asthma event rate over time in frailty models. Journal of the Royal Statistical Society Series C: Applied Statistics, 52(3), 355-363

[41] Duchateau, L., Janssen, P. (2004). Penalized partial likelihood for frailties and smoothing splines in time to first insemination models for dairy cows. Biometrics, 60(3), 608-614.

[42] Duchateau L, Janssen P (2007). The frailty model. Springer Science Business Media, Berlin.

[43] El-Morshedy, M., Eliwa, M. S., Al-Bossly, A., Yousof, H. M. (2022). A New Probability Heavy-Tail Model for Stochastic Modeling under Engineering Data. Journal of Mathematics, 2022(1), 1910909.

[44] Elbatal, I., Diab, L. S., Ghorbal, A. B., Yousof, H. M., Elgarhy, M. and Ali, E. I. (2024). A new losses (revenues) probability model with entropy analysis, applications and case studies for value-at-risk modeling and mean of order-P analysis. AIMS Mathematics, 9(3), 7169-7211.

[45] Elbers, C., Ridder, G. (1982). True and spurious duration dependence: The identifiability of the proportional hazard model. The Review of Economic Studies, 49(3), 403-409.

[46] Farewell, V. T. (1977). A model for a binary variable with time-censored observations. Biometrika, 64(1), 43-46.

[47] Feller, W. (1971). An introduction to probability theory and its applications (Vol. 963). New York: Wiley.

[48] Fine, J. P., Glidden, D. V., Lee, K. E. (2003). A simple estimator for a shared frailty regression model. Journal of the Royal Statistical Society Series B: Statistical Methodology, 65(1), 317-329.

[49] Fleming, T. R., Harrington, D. P. (2013). Counting processes and survival analysis. John Wiley Sons.

[50] Flinn, C., Heckman, J. (1982). New methods for analyzing structural models of labor force dynamics. Journal of econometrics, 18(1), 115-168.

[51] Genest, C., MacKay, J. (1986). The joy of copulas: Bivariate distributions with uniform marginals. The American Statistician, 40(4), 280-283.

[52] Ghitany, M. E., Atieh, B., Nadarajah, S. (2008). Lindley distribution and its application. Mathematics and computers in simulation, 78(4), 493-506.

[53] Glidden, D. V. (1999). Checking the adequacy of the gamma frailty model for multivariate failure times. Biometrika, 86(2), 381-393.

[54] Gompertz, B. (1825). XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS c. Philosophical transactions of the Royal Society of London, (115), 513-583.

[55] Goual, H., Yousof, H. M., Ali, M. M. (2019). Validation of the odd Lindley exponentiated exponential by a modified goodness of fit test with applications to censored and complete data. Pakistan Journal of Statistics and Operation Research, 15(3), 745-771.

[56] Goual, H., Yousof, H. M. and Ali, M. M. (2020a). Lomax inverse Weibull model: properties, applications, and a modified Chi-squared goodness-of-fit test for validation. Journal of Nonlinear Sciences Applications (JNSA), 13(6), 330-353.

[57] Goual, H., and Yousof, H. M. (2020b). Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test. Journal of Applied Statistics, 47(3), 393-423.

[58] Grandell, J. (1997). Mixed poisson processes (Vol. 77). CRC Press.

[59] Greenwood, M., Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the

occurrence of multiple attacks of disease or of repeated accidents. Journal of the Royal statistical society, 83(2), 255-279.

[60] Guo, G., Rodriguez, G. (1992). Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala. Journal of the American Statistical Association, 87(420), 969-976.

[61] Guo, G. (1993). Use of sibling data to estimate family mortality effects in Guatemala. Demography, 15-32.

[62] Gupta, P. L., Gupta, R. D. (1990). A bivariate random environmental stress model. Advances in Applied Probability, 22(2), 501-503.

[63] Gustafson, P. (1997). Large hierarchical Bayesian analysis of multivariate survival data. Biometrics, 230-242.

[64] Hamed, M. S., Cordeiro, G. M. and Yousof, H. M. (2022). A New Compound Lomax Model: Properties, Copulas, Modeling and Risk Analysis Utilizing the Negatively Skewed Insurance Claims Data. Pakistan Journal of Statistics and Operation Research, 18(3), 601-631. https://doi.org/10.18187/pjsor.v18i3.3652

[65] Hanagal, D. D. (2011). Modeling survival data using frailty models. Boca Raton: Chapman Hall/CRC.

[66] Hashem, A. F., Abdelkawy, M. A., Muse, A. H., Yousof, H. M. (2024). A novel generalized Weibull Poisson G class of continuous probabilistic distributions with some copulas, properties and applications to real-life datasets. Scientific Reports, 14(1), 1741.

[67] Hashempour, M., Alizadeh, M. and Yousof, H. M. (2023). A New Lindley Extension: Estimation, Risk Assessment and Analysis Under Bimodal Right Skewed Precipitation Data. Annals of Data Science, 1-40.

[68] Haukka, J., Suvisaari, J., Lönnqvist, J. (2003). Increasing age does not decrease risk of schizophrenia up to age 40. Schizophrenia research, 61(1), 105-110.

[69] Heckman, J., Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica: Journal of the Econometric Society, 271-320.

[70] Heckman, J. J., Walker, J. R. (1990). Estimating fecundability from data on waiting times to first conception. Journal of the American Statistical Association, 85(410), 283-294.

[71] Henderson, R., Oman, P. (1999). Effect of frailty on marginal regression estimates in survival analysis. Journal of the Royal Statistical Society Series B: Statistical Methodology, 61(2), 367-379.

[72] Holford, N. H., Sheiner, L. B. (1981). Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. Clinical pharmacokinetics, 6(6), 429-453.

[73] Holford, N. H., Sheiner, L. B. (1982). Kinetics of pharmacologic response. Pharmacology therapeutics, 16(2), 143-166.

[74] Holgate, P. (1970). The modality of some compound Poisson distributions. Biometrika, 57(3), 666-667.

[75] Horowitz, J. L. (1999). Semiparametric estimation of a proportional hazard model with unobserved heterogeneity. Econometrica, 67(5), 1001-1028.

[76] Hougaard, P. (1984). Life table methods for heterogeneous populations: distributions describing the heterogeneity. Biometrika, 71(1), 75-83.

[77] Hougaard, P. (1986a). A class of multivariate failure time distributions. Biometrika, 73(3), 671-678.

[78] Hougaard, P. (1986b). Survival models for heterogeneous populations derived from stable distributions. Biometrika, 73(2), 387-396.

[79] Hougaard, P. (1987). Modelling multivariate survival. Scandinavian Journal of Statistics, 291-304.

[80] Hougaard, P. (1991). Modelling heterogeneity in survival data. Journal of Applied Probability, 28(3), 695-701.

[81] Hougaard, P., Harvald, B., Holm, N. V. (1992). Measuring the similarities between the lifetimes of adult Danish twins born between 1881–1930. Journal of the American Statistical Association, 87(417), 17-24.

[82] Hougaard, P., Myglegaard, P., Borch-Johnsen, K. (1994). Heterogeneity models of disease susceptibility, with application to diabetic nephropathy. Biometrics, 1178-1188.

[83] Hougaard, P. (2000). Analysis of multivariate survival data (Vol. 564). New York: Springer.

[84] Hougaard, P. (2012). Analysis of multivariate survival data. Springer Science Business Media, Berlin.

[85] Huang, X., Wolfe, R. A. (2002). A frailty model for informative censoring. Biometrics, 58(3), 510-520.

[86] Ibrahim, J., Chen, M., Sinha, D. (2001). Bayesian survival analysis springer series in statistics. Springer, New York, 978-981.

[87] Ibrahim, M., Yadav, A. S., Yousof, H. M., Goual, H., and Hamedani, G. G. (2019). A new extension of Lindley distribution: modified validation test, characterizations and different methods of estimation. Communications for Statistical Applications and Methods, 26(5), 473-495.

[88] Ibrahim. M., Aidi, K., Ali, M. M. and Yousof, H. M. (2021). The Exponential Generalized Log-Logistic Model: Bagdonavičius-Nikulin test for Validation and Non-Bayesian Estimation Methods. Communications for Statistical Applications and Methods, 29(1), 681-705.

[89] Janosi, A., Steinbrunn, W., Pfisterer, M., and Detrano, R., (1988). Heart Disease. UCI Machine Learning Repository. https://doi.org/10.24432/C52P4X

[90] Joe, H. (1993). Parametric families of multivariate distributions with given margins. Journal of multivariate analysis, 46(2), 262-282.

[91] Kalbfleisch, J. D., Prentice, R. L. (2002). The statistical analysis of failure time data. John Wiley Sons.

[92] Keiding, N., Andersen, P. K., Klein, J. P. (1997). The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. Statistics in medicine, 16(2), 215-224.

[93] Khalil, M. G., Aidi, K., Ali, M. M., Butt, N. S., Ibrahim, M., Yousof, H. M. (2024). Modified Bagdonavicius-Nikulin Goodness-of-fit Test Statistic for the Compound Topp Leone Burr XII Model with Various Censored Applications. Statistics, Optimization Information Computing, 12(4), 851-868.

[94] Klein, J. P., Moeschberger, M., Li, Y. H., Wang, S. T., Flournoy, N. (1992). Estimating random effects in the Framingham heart study. In Survival Analysis: State of the Art (pp. 99-120). Dordrecht: Springer Netherlands.

[95] Klein, J. P., Pelz, C., Zhang, M. J. (1999). Modeling random effects for censored data by a multivariate normal regression model. Biometrics, 55(2), 497-506.

[96] Kleinbaum, D. G., Klein, M. (1996). Survival analysis a self-learning text. Springer.

[97] Kleinbaum, D. G., Klein, M. (2012). Survival Analysis. In Statistics for Biology and Health. Springer New York.

[98] Kosorok, M. R. (2008). Introduction to empirical processes and semiparametric inference. New York, NY: Springer New York.

[99] Kuk, A. Y., Chen, C. H. (1992). A mixture model combining logistic regression with proportional hazards regression. Biometrika, 79(3), 531-541.

[100] Lam, K. F., Lee, Y. W., Leung, T. L. (2002). Modeling multivariate survival data by a semiparametric random effects proportional odds model. Biometrics, 58(2), 316-323.

[101] Lam, K. F., Lee, Y. W. (2004). Merits of modelling multivariate survival data using random effects proportional odds model. Biometrical Journal: Journal of Mathematical Methods in Biosciences, 46(3), 331-342.

[102] Lam, K. F., Fong, D. Y., Tang, O. Y. (2005). Estimating the proportion of cured patients in a censored sample. Statistics in medicine, 24(12), 1865-1879.

[103] Lancaster, T. (1990). The econometric analysis of transition data (No. 17). Cambridge university press.

[104] Lawless, J. F. (2003). Statistical models and methods for lifetime data. John Wiley Sons.

[105] Lee, E. W., Wei, L. J., Amato, D. A., Leurgans, S. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In Survival analysis: state of the art (pp. 237-247). Dordrecht: Springer Netherlands.

[106] Lehmann, EL., Casella, G. (2006) .Theory of point estimation. Springer Science & Business Media, Berlin.

[107] Liang, K. Y., Self, S. G., Bandeen-Roche, K. J., Zeger, S. L. (1995). Some recent developments for regression analysis of multivariate failure time data. Lifetime data analysis, 1(4), 403-415.

[108] Lillard, L. A. (1993). Simultaneous equations for hazards: Marriage duration and fertility timing. Journal of econometrics, 56(1-2), 189-217.

[109] Lillard, L. A., Brien, M. J., Waite, L. J. (1995). Premarital cohabitation and subsequent marital dissolution: A matter of self-selection?. Demography, 32(3), 437-457.

[110] Lin, D. Y. (1994). Cox regression analysis of multivariate failure time data: the marginal approach. Statistics in medicine, 13(21), 2233-2247.

[111] Lindeboom, M., Van den Berg, G. J. (1994). Heterogeneity in models for bivariate survival: the importance of the mixing distribution. Journal of the Royal Statistical Society Series B: Statistical Methodology, 56(1), 49-60.

[112] Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. Journal of the Royal Statistical Society. Series B (Methodological), 102-107.

[113] Longini Jr, I. M., Halloran, M. E. (1996). A frailty mixture model for estimating vaccine efficacy. Journal of the Royal Statistical Society Series C: Applied Statistics, 45(2), 165-173.

[114] Mahé, C., Chevret, S. (1999). Estimating regression parameters and degree of dependence for multivariate failure time data. Biometrics, 55(4), 1078-1084.

[115] Manatunga, A. K., Oakes, D. (1999). Parametric analysis for matched pair survival data. Lifetime Data Analysis, 5(4), 371-387.

[116] Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. Journal of the American Statistical Association, 58(303), 690-700.

[117] Manton, K. G., Stallard, E., Vaupel, J. W. (1986). Alternative models for the heterogeneity of mortality risks among the aged. Journal of the American Statistical Association, 81(395), 635-644.

[118] Marshall, A. W., Olkin, I. (1988). Families of multivariate distributions. Journal of the American statistical association, 83(403), 834-841.

[119] Martinussen, T., Pipper, C. B. (2005). Estimation in the positive stable shared frailty Cox proportional hazards model. Lifetime data analysis, 11(1), 99-115.

[120] Mazucheli, J., Coelho-Barros, EA., Achcar, JA. (2016). An alternative reparametrization for the weighted lindley distribution. Pesquisa Operacional , 36(2), 345-353.

[121] McGilchrist, C.A., Aisbett, C.W. (1991). Regression with Frailty in Survival Analysis. Biometrics, 47, 461-466.

[122] McGilchrist, C. A. (1993). REML Estimation for Survival Models with Frailty. Biometrics, 49(1), 221.

[123] Minkah, R., de Wet, T., Ghosh, A., Yousof, H. M. (2023). Robust extreme quantile estimation for Pareto-type tails through an exponential regression model. Communications for Statistical Applications and Methods, 30(6), 531-550.

[124] Moger, T. A., Aalen, O. O., Halvorsen, T. O., Storm, H. H., Tretli, S. (2004). Frailty modelling of testicular cancer incidence using Scandinavian data. Biostatistics, 5(1), 1-14.

[125] Mohamed, H. S., Cordeiro, G. M., Minkah, R., Yousof, H. M., Ibrahim, M. (2024). A size-of-loss model for the negatively skewed insurance claims data: applications, risk analysis using different methods and statistical forecasting. Journal of Applied Statistics, 51(2), 348-369.

[126] Mota, A., Milani, E. A., Calsavara, V. F., Tomazella, V. L., Leao, J., Ramos, P. L., ... Louzada, F. (2021). Weighted Lindley frailty model: estimation and application to lung cancer data. Lifetime Data Analysis, 27(4), 561-587.

[127] Murphy, S. A. (1994). Consistency in a proportional hazards model incorporating a random effect. The Annals of Statistics, 22(2), 712-731.

[128] Murphy, S. A. (1995). Asymptotic theory for the frailty model. The annals of statistics, 182-198.

[129] Murphy, S. A., Rossini, A. J., van der Vaart, A. W. (1997). Maximum likelihood estimation in the proportional odds model. Journal of the American Statistical Association, 92(439), 968-976.

[130] Nelson, C. R., Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. Journal of monetary economics, 10(2), 139-162.

[131] Nielsen, G. G., Gill, R. D., Andersen, P. K., Sørensen, T. I. (1992). A counting process approach to maximum likelihood estimation in frailty models. Scandinavian journal of Statistics, 25-43.

[132] Nikulin, M.S, (1973a). Chi-squared test for continuous distributions with shift and scale parameters, Theory of Probability and its Applications. 18, 559-568.

[133] Nikulin, M.S. (1973b). Chi-squared test for normality. In proceedings of the International Vilnius Conference on Probability Theory and Mathematical Statistics, 2, 119-122.

[134] Nikulin, M.S, (1973c). On a Chi-squared test for continuous distributions, Theory of Probability and its Applications. 19, 638-639.

[135] Oakes, D. (1989). Bivariate survival models induced by frailties. Journal of the American Statistical Association, 84(406), 487-493.

[136] Ottestad, P. (1944). On certain compound frequency distributions. Scandinavian Actuarial Journal, 1944(1-2), 32-42.

[137] Parner, E. (1998). Asymptotic theory for the correlated gamma-frailty model. The Annals of Statistics, 26(1), 183-214.

[138] Pickles, A., Crouchley, R. (1995). A comparison of frailty models for multivariate survival data. Stat Med , 14(13), 1447-1461.

[139] Price, D. L., Manatunga, A. K. (2001). Modelling survival data with a cured fraction using frailty models. Statistics in medicine, 20(9-10), 1515-1527.

[140] Rao, K. C., Robson, D. S. (1974). A Chi-square statistic for goodness-of-fit tests within the exponential family. Communication in Statistics, 3, 1139-1153.

[141] Ripatti, S., Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. Biometrics, 56(4), 1016-1022.

[142] Ripatti, S., Larsen, K., Palmgren, J. (2002). Maximum likelihood inference for multivariate frailty models using an automated Monte Carlo EM algorithm. Lifetime Data Analysis, 8(4), 349-360.

[143] Robert, C., Casella, G.(2013). Monte Carlo statistical methods. Springer Science & Business Media, Berlin.

[144] Rocha, C. S. (1996). Survival models for heterogeneity using the non-central chi-squared distribution with zero degrees of freedom. In Lifetime Data: Models in Reliability and Survival Analysis (pp. 275-279). Boston, MA: Springer US.

[145] Salem, M., Emam, W., Tashkandy, Y., Ibrahim, M., Ali, M. M., Goual, H., Yousof, H. M. (2023). A new lomax extension: Properties, risk analysis, censored and complete goodness-of-fit validation testing under left-skewed insurance, reliability and medical data. Symmetry, 15(7), 1356.

[146] Sankaran, M. (1970). 275. note: The discrete poisson-lindley distribution. Biometrics, 145-149.

[147] Sastry, N. (1997). A nested frailty model for survival data, with an application to the study of child survival in northeast Brazil. Journal of the American Statistical Association, 92(438), 426-435.

[148] Schnier, C., Hielm, S., Saloniemi, H. S. (2004). Comparison of the breeding performance of cows in cold and warm loose-housing systems in Finland. Preventive veterinary medicine, 62(2), 135-151.

[149] Selvin, S. (2004). Statistical analysis of epidemiologic data (Vol. 35). Oxford University Press.

[150] Sen, S., Alizadeh, M., Aboraya, M., Ali, M. M., Yousof, H. M., Ibrahim, M. (2024). On truncated versions of xgamma distribution: Various estimation methods and statistical modelling. Statistics, Optimization Information Computing, 12(4), 943-961.

[151] Shanker, R., Mishra, A. (2013). A two-parameter Lindley distribution. Statistics in Transition new series, 14(1), 45-56.

[152] Shih, J. H., Louis, T. A. (1995). Assessing gamma frailty models for clustered failure time data. Lifetime Data Analysis, 1(2), 205-220.

[153] Spilerman, S. (1972). Extensions of the mover-stayer model. American Journal of Sociology, 78(3), 599-626.

[154] Struthers, C. A., Kalbfleisch, J. D. (1986). Misspecified proportional hazard models. Biometrika, 73(2), 363-369.

[155] SÁ CARVALHO, M. A. R. I. L. I. A., Henderson, R., Shimakura, S., Sousa, I. P. S. C. (2003). Survival of hemodialysis patients: modeling differences in risk of dialysis centers. International Journal for Quality in Health Care, 15(3), 189-196.

[156] Therneau, T. M., Grambsch, P. M. (2000). Frailty models. In Modeling survival data: Extending the Cox model (pp. 231-260). New York, NY: Springer New York.

[157] Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. Journal of the Royal Statistical Society: Series B (Methodological), 38(3), 290-295.

[158] Tweedy, M. (1984). An index which distinguishes between some important exponential families. Statistics: applications and new directions. In Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (pp. 579-604).

[159] Varadhan, R., Gilbert, P. (2010). BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function. Journal of statistical software, 32, 1-26.

[160] Vaupel, J.W., Manton, K.G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16(3), 439-454.

[161] Wang, S. T., Klein, J. P., Moeschberger, M. L. (1995). Semi-parametric estimation of covariate effects using the positive stable frailty model. Applied stochastic models and data analysis, 11(2), 121-133.

[162] Wei, L. J., Lin, D. Y., Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. Journal of the American statistical association, 84(408), 1065-1073.

[163] Wei, L. A., Glidden, D. V. (1997). An overview of statistical methods for multiple failure time data in clinical trials. Statistics in medicine, 16(8), 833-839.

[164] Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of applied mechanics.

[165] Wienke, A. (2010). Frailty models in survival analysis. CRC Press, Boca Raton.

[166] Xue, X., Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. Lifetime Data Analysis, 2(3), 277-289.

[167] Yadav, A. S., Goual, H., Alotaibi, R. M., Rezk, H., Ali, M. M., and Yousof, H. M. (2020). Validation of the Topp-Leone-Lomax model via a modified Nikulin-Rao-Robson goodness-of-fit test with different methods of estimation. Symmetry, 12(1), 57.

[168] Yadav, A. S., Shukla, S., Goual, H., Saha, M. and Yousof, H. M. (2022). Validation of xgamma exponential model via Nikulin-Rao-Robson goodness-of- fit test under complete and censored sample with different methods of estimation. Statistics, Optimization Information Computing, 10(2), 457-483.

[169] Yashin, A. I., Vaupel, J. W., Iachine, I. A. (1995). Correlated individual frailty: an advantageous approach to survival analysis of bivariate data. Mathematical population studies, 5(2), 145-159.

[170] Yashin, A. I., Begun, A. Z., Iachine, I. A. (1999). Genetic factors in susceptibility to death: a comparative analysis of bivariate survival models. Journal of Epidemiology and Biostatistics, 4(1), 53-60.

[171] Yousof, H. M., Goual, H., Khaoula, M. K., Hamedani, G. G., Al-Aefaie, A. H., Ibrahim, M., ... and Salem, M. (2023a). A novel accelerated failure time model: Characterizations, validation testing, different estimation methods and applications in engineering and medicine. Pakistan Journal of Statistics and Operation Research, 19(4), 691-717.

[172] Yousof, H. M., Ali, M. M., Aidi, K., Ibrahim, M. (2023b). The modified Bagdonavičius-Nikulin goodness-of-fit test statistic for the right censored distributional validation with applications in medicine and reliability. Statistics in Transition new series, 24(4), 1-18.

[173] Zakerzadeh, H., Dolati, A. (2009). Generalized Lindley Distribution. Journal of Mathematical extension.