FRAILTY MODELS FOR MODELLING HETEROGENEITY

By

ULVIYA ABDULKARIMOVA, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

© Copyright by Ulviya Abdulkarimova, April 2013
MASTER OF SCIENCE (2013) McMaster University
(Statistics) Hamilton, Ontario

TITLE: FRAILTY MODELS FOR MODELLING HETEROGENEITY

AUTHOR: Ulviya Abdulkarimova, B.Sc. (McMaster University)
SUPERVISOR: Professor N. Balakrishnan
NUMBER OF PAGES: ix, 76
Abstract

In the analysis of lifetime data, heterogeneity may be present in many situations. This heterogeneity is also referred to as frailty. Analysis that ignore frailty when it is present leads to incorrect inferences. In lifetime analysis, Cox proportional hazards model is commonly used to measure the effects of covariates. The covariates may fail to fully account for the true differences in risk. This may be due to an existence of another variable that is ignored in the model but can be explained by random frailty. Including frailty in the model can avoid underestimation and overestimation of parameters and also correctly measure the effects of the covariates on the response variable.

This thesis presents an extension of Cox model to parametric frailty model in which the exponential and Weibull distributions are used as the distributions of baseline hazard, and the gamma and Weibull distributions are used as frailty distributions. We examine the maximum likelihood estimation procedures and propose the use of Monte Carlo integration method or quadrature method in complicated cases where explicit solutions to the likelihood functions can not be obtained.

The gamma distribution is one of the most commonly used distributions for frailty. It has a closed form likelihood function that can be readily maximized. In this thesis, we study the performance of the Weibull distribution as a frailty distribution and compare with the gamma frailty model. Through simulation studies, the performance of the parameter estimates are evaluated. The effects of increasing the sample size and cluster size separately are also studied through Monte Carlo simulations. The Akaike Information Criteria (AIC) is used to compare the performance of the gamma and Weibull frailty models. The developed methods are then illustrated with numerical examples.

KEY WORDS: frailty, heterogeneity, Monte Carlo simulations, Monte Carlo integration, gamma distribution, Weibull distribution, Delta method, Numerical integration, AIC, BIC.
Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. N. Balakrishnan for his supervision throughout my research, continual support, encouragement, and invaluable advice.

I would also like to thank my thesis committee: Dr. Roman Viveros-Aguilera and Dr. Aaron Childs for their insightful comments.

I would like to thank my parents, siblings and husband, whose support and encouragement allowed me to finish this journey and my daughter, who joined us when I was preparing for my thesis defence, for giving me unlimited happiness.
Contents

Abstract iii

Acknowledgements iv

1 Introduction 1

1.1 Preliminary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
  1.1.1 Consequence of Ignoring Frailty . . . . . . . . . . . . . . . . . . . . . 2
1.2 Modelling Frailty . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
  1.2.1 Cox Proportional Hazards Model . . . . . . . . . . . . . . . . . . . . . 5
  1.2.2 Univariate Frailty Models . . . . . . . . . . . . . . . . . . . . . . . . . 6
  1.2.3 Multivariate Frailty Models . . . . . . . . . . . . . . . . . . . . . . . . 7
1.3 Scope of Thesis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9

2 Statistical Distributions for Frailty 11

2.1 Gamma Distribution . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
2.2 Weibull Distribution . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14

3 Likelihood Approach to Frailty Models 18

3.1 Survival Likelihood . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19
3.2 Gamma Frailty Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20
  3.2.1 Asymptotic Variance-Covariance Matrix . . . . . . . . . . . . . . . . . 24
3.3 Weibull Frailty Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
C.2 Kidney .......................................................... 69
C.2.1 Gamma Frailty ............................................. 69
C.2.2 Weibull Frailty ............................................ 70

Bibliography ..................................................... 73
## List of Tables

4.1 Shape and scale parameters of the Weibull distribution ........................................ 32

4.2 MSEs and biases (x100) of estimated treatment effect $\beta$ and the frailty variance $\theta$ of frailty models based on 1000 simulated data generated from frailty models with gamma distribution ................................................................. 34

4.3 MSEs and biases (x100) of estimated treatment effect $\beta$ and the frailty variance $\theta$ of frailty models based on 1000 simulated data generated from frailty models with Weibull distribution ................................................................. 35

4.4 Proportion of times the model is selected by AIC ($n=200, m_i=2$) .......................... 37

4.5 Proportion of times the model is selected by AIC ($n=400, m_i=2$) .......................... 38

5.1 The estimates of the regression coefficients, variance component and their standard errors for Infected Kidney Patients Data Set using Weibull Baseline Hazard ................................................................. 40

5.2 The estimates of the regression coefficients, variance component and their standard errors for Litter Data Set using Weibull Baseline Hazard .......................... 42

A.1 Litter-matched study of the tumorigenesis of a drug ................................................. 47

A.2 Recurrence times to kidney infection ................................................................. 48
List of Figures

1.1 Hazard plots illustrating the problem of ignoring heterogeneity . . . . . . . . . 3

2.1 Probability density and hazard functions of different gamma distributions with
mean 1 and variances 1, 0.75, 0.5, and 0.25 . . . . . . . . . . . . . . . . . . . . . . 13

2.2 Probability density and hazard functions of different Weibull distributions
with mean 1 and variances 1, 0.75, 0.5, and 0.25 . . . . . . . . . . . . . . . . . . . . 16

4.1 Gamma and Weibull density with mean 1 and variance 0.5 . . . . . . . . . . . 38
Chapter 1

Introduction

1.1 Preliminary

The statistical analysis of survival data is an important topic in many areas, including medicine, epidemiology, biology, demography, economics, engineering and other fields. A variety of techniques have been developed to analyse survival data. A common approach to the analysis of survival data is based on the assumption that the study population is homogeneous. That is, conditional on the covariates, every individual has the same risk of experiencing an event such as death or disease recurrence. The event times of individuals in the population, conditional on the observed covariates, are assumed to be independent. However, this can not be assumed in all applications as many applications require heterogeneous sample, i.e. individuals with different risks and hazards. In practice, there may be an association between the event times of some subgroups of the population since the individuals of these groups share a common trait that can not be observed. For example, there may be an association in the times to events of cancer or cardiovascular diseases between siblings or married couples, even occurrence of nonlethal diseases within the same individual. Though individuals may look identical in some aspects, they may differ in unmeasured ways. In applications of survival analysis, usually only a few covariates such as age, sex, severity
of disease or laboratory data are known. It is known that there are many other factors that can influence survival, including health status, lifestyle, smoking, occupation and genetic risk factors. These factors are unknown and can not be included in the analysis. When these factors are ignored, it is assumed that a population analysis describes an average person, not an individual.

Beard (1959), Vaupel et al. (1979), and Lancaster (1979) suggested a random effects model in order to account for the unobserved heterogeneity due to unobserved covariates. Beard (1959) used the term *longevity factor* to improve the fit of mortality models in populations. Vaupel et al. (1979) introduced the term *frailty* in order to account for unobserved heterogeneity, random effects, and association in univariate survival models. He introduced this concept of frailty to biostatistics by applying it on population mortality data. Lancaster (1979) introduced the model to the literature of economics and the model is called the mixed proportional hazards model. The concept, however, goes back to work of Greenwood and Yule on “accident proneness” in 1920. Clayton (1978) discussed the applications of the model to multivariate survival data in his seminar paper on chronic disease incidence in families. Frailty models account for unobserved heterogeneity that occurs because some observations are more prone to failure, and therefore more frail than others in a data set. Therefore, the objective is to introduce an additional parameter to the hazard rate that accounts for the random frailties. These frailties can be specific to individuals or groups, and are referred to as *individual frailty* or *shared frailty*.

1.1.1 Consequence of Ignoring Frailty

It is very important to consider the effect of ignoring frailty where the existence of heterogeneity may be present. Let’s consider a study of survival times for a population that consists of two subgroups, low-risk and high-risk. We will assume that both subgroups are exponentially distributed with hazard rates 0.01 and 0.04, respectively. We may assume that if these two subgroups are combined, subjects in high-risk group will experience the event
and exit earlier at a higher rate and the remaining population will have low-risk subjects with lower population hazard rate. Subjects with higher hazard rate are more ‘frail’ than subjects with lower hazard rate. Therefore, even though subjects in each subgroup have the same hazard, the population hazard function would decline over time. Figure 1.1 illustrates this problem of ignoring the heterogeneity. The estimated hazard rate for the population is shown by the downward sloping line, where the heterogeneity in risk factors is ignored.

![Hazard plots illustrating the problem of ignoring heterogeneity](image)

**Figure 1.1:** Hazard plots illustrating the problem of ignoring heterogeneity

Ignoring the existence of heterogeneity will produce incorrect estimation of parameters and their standard errors in survival analysis. According to Keyfitz and Littman (1979), ignoring heterogeneity overestimates life expectancy based on their study on estimating life expectancy in a heterogeneous population. Lancaster (1990) showed that when heterogeneity is ignored, it caused underestimation of covariate effects in his study of unemployment rates. Henderson and Oman (1999) showed that ignoring frailty leads to regression coefficient estimates biased towards zero by an amount depending on the distribution and the variability of the frailty terms.
1.2 Modelling Frailty

Using parametric or semiparametric regression models is an important way to handle heterogeneity. Regression models take lifetime as the dependent variable and explanatory variables as regressors. Sometimes these models may not provide adequate fit to the data. One of the reasons is due to omission of important covariates. Several methods have been developed to model the frailty in survival data during recent years. The generalization of the Cox proportional hazards model (Cox, 1972) is the best and widely applied model that allows for the random effect by multiplicatively adjusting the baseline hazard function.

Frailty models extend Cox proportional hazards model by introducing unobserved “frailties” to the model. In this case, the hazard rate will not be just a function of covariates, but also a function of frailties. A frailty model is a random effects model which has a multiplicative effect on the hazard rates of all the members of the subgroups. In univariate survival models, it can be used to model the heterogeneity among individuals, which is the influence of unobserved risk factors in a proportional hazards model. In multivariate survival models, shared frailty model is used to model the dependence between the individuals in the group. In the multivariate case, unobserved frailty is common to a group of individuals.

Suppose we have a data set of \( n \) individuals from some population and \( i = 1, ..., G \) subgroups or clusters. Each subgroup consists of \( n_i \geq 1 \) individuals. Individuals in each subgroup have dependent event times due to unobserved frailty, \( u_i \). This frailty term may represent aggregate effect of common genes or shared environmental effect on survival of members of a given family, such as siblings, husband and wives, etc. When the size of a subgroup is 1, then an individual is affected by his/her own frailty, and this case is referred to as univariate frailty.
1.2.1 Cox Proportional Hazards Model

Frailty models are generalizations of Cox proportional hazards model, which is known as the most popular model in survival analysis. It is a regression model with event time as the dependent variable. Cox proportional hazards model is used to assess the covariate effects on the hazard function. The model is given as

\[ h(t|Z) = h_0(t)c(\beta'Z), \]  

(1.2.1)

where \( h_0(t) \) is an arbitrary baseline hazard rate, \( \beta \) is a parameter vector, and \( Z \) is a covariate vector. A common model for \( c(\beta'Z) \) is

\[ c(\beta'Z) = exp(\beta'Z). \]

The exponential form ensures that the hazard rate remains positive for all values of \( Z \). One of the assumptions of this model is that the baseline hazard is common to all the individuals in the study population. The covariates act multiplicatively on the baseline hazard, which adds additional risks based on each individual’s prognostic information. The cumulative hazard and survival functions based on Cox proportional hazards model are given by

\[ H(t, Z) = \int_0^t h_0(u)e^{\beta'u}du = H_0(t)e^{\beta'Z}, \]  

(1.2.2)

and

\[ S(t, Z) = exp(-H_0(t)e^{\beta'Z}), \]  

(1.2.3)

where \( H_0(t) \) is the baseline cumulative hazard function.

The main objective of the Cox model is to assess the effects of the covariates by estimating their coefficients. However, the covariates do not always fully account for the true differences in risk and this is due to frailty variable that is not included in the model. Therefore, including frailties in the model allows to correctly measure the covariate effects and avoid underestimation or overestimation of the parameters.
1.2.2 Univariate Frailty Models

In frailty models, the variability of survival times can be divided into two parts. One part is observed risk factors, known as covariates, and the other part is unobserved risk factors, known as frailty. The univariate frailty model presents the population as a mixture in which baseline hazard is common to all individuals but each individual has his/her own frailty. Suppose we have a sample of \( j \) observations in a study. Some of these observations fail earlier than others due to unobserved heterogeneity. The proportional hazards model assumes that conditional on the frailty, the hazard function for an individual at time \( t > 0 \) is

\[
h_j(t) = h_0(t)e^{\beta'z_j + \psi_j}, \quad j = 1, \ldots, n, \tag{1.2.4}
\]

where \( \psi_j \) is a frailty term from a probability distribution with a mean of 0 and variance of 1. If \( \psi_j \) could be measured and included in the model, then \( \psi_j \) would go to 0 and we would obtain the standard proportional hazards model. The hazard function conditional on both covariates and frailty can be rewritten as

\[
h_j(t) = h_0(t)u_je^{\beta'z_j}, \quad j = 1, \ldots, n, \tag{1.2.5}
\]

where \( u_j = e^{\psi_j} \). This shows that the hazard of an individual also depends on an unobservable random variable, \( u_j \), which acts multiplicatively on the hazard rate. If frailty is not taken into account, then \( u_j = 1 \).

In the univariate case, frailty models are used to make adjustments for overdispersion. When unobserved or unmeasured effects are ignored, the estimates of survival may be misleading. Therefore, corrections for this overdispersion is needed in order to allow for adjustments for those important frailties.
1.2.3 Multivariate Frailty Models

Many statistical models and methods proposed to model failure time data assume that the observations are statistically independent of each other. This, however, does not hold in many applications. Multivariate or shared frailty model is a conditional independence model in which frailty is common to all subjects in a cluster. The shared frailty model is responsible for creating dependence between event times. It is also known as a mixture model because the frailties in each cluster are assumed to be random. It assumes that, given the frailty, all event times in a cluster are independent. Shared frailty model was introduced by Clayton (1978) without using the notion frailty and extensively studied in Hougaard (2000), Therneau and Grambsch (2000), Duchateau et al. (2002, 2003), and Duchateau and Janssen (2004).

Multivariate frailty model is an extension of the univariate frailty model which allows the individuals in the same cluster to share the same frailty value. When frailty is shared, dependence between individuals who share frailties is generated. However, when conditioning on the frailty, the individuals are independent of each other.

Shared frailty models are very important in analysing multivariate or clustered survival data. Multivariate survival data arise when individuals experience recurrent events such as recurrence of tumor or infection or when there exists some grouping of individuals. In multicenter clinical trials, clustering is very common. Here, frailty describes center-to-center variations that are not explained by observed covariates. Shared frailty model assumes that individuals in a subgroup or pair share the same frailty $u$, but frailty from group to group may differ.

Shared frailty model is similar to the individual frailty model except the only difference is that frailty is now shared among the $nj_i$ observations in the $ith$ group. Suppose we have $j$ observations and $i$ subgroups. Each subgroup consists of $n_i$ observations and $\sum_{i=1}^{G} n_i = n$, where $n$ is the total sample size. The hazard rate for the $j$th individual in the $ith$ subgroup is given by
where $W_i$ are frailty terms for subgroups and their distribution is again assumed to be independent with a mean of 0 and a variance of 1. The hazard function conditional on covariates and frailties can be rewritten as

$$h_{ij}(t) = h_0(t)u_i e^{\beta Z_{ij} + W_i \phi}, i = 1, \ldots, G, j = 1, \ldots, n_i,$$

(1.2.6)

where $u_i = e^{W_i \psi}$.

As in the proportional hazards model, parametric or non-parametric forms of baseline hazard can be assumed in frailty models. If non-parametric form is assumed for $h_0(t)$, then semiparametric proportional hazards model is considered and the estimates are usually obtained by using Expectation-Maximization (EM) algorithm. If parametric form for $h_0(t)$ is assumed, then maximum likelihood estimates can be obtained by maximizing the likelihood function. In this thesis, we will only consider the parametric forms of baseline hazard for simplicity. Most applications assume Weibull baseline hazard, however, other parametric baseline hazards can be used. Using parametric baseline hazards not only makes the estimation easier, but it can also describe explicitly the effect of the frailty on hazard ratios over time.

The frailty model can be thought as a misspecified model. This misspecification can be taken into account, if assumptions are made about the probability distribution of the frailty. Any continuous distribution with positive support, mean of 1, and finite unknown variance $\theta$ can be used for the distribution of $u$. The mean of the frailty term is constrained to 1 in order to make the average hazard identifiable. The frailty term $u$ accounts for the extra variability from unobserved risk factors that are not accounted for by other predictors. If we assume that $u$ has some distribution, then we can estimate the frailty model. The goal is to estimate the frailty variance, $\theta$. The variance of the frailty distribution is used to determine the degree of heterogeneity in the study population. When the variance is small,
the values of frailty are around one. In this setting, there are no frailty effects and failures are independent within as well as across groups. When the variance is large, then the values of frailty are dispersed and hence induce greater heterogeneity in the individual hazards. Larger variance also indicates stronger association within groups. The association between group members are measured by Kendall’s $\tau$, which is given by

$$\tau = \frac{\theta}{(\theta+2)} \quad \text{with} \quad SE(\tau) = \frac{2SE(\theta)}{(2+\theta)^2}.$$ 

The estimation methods of the parameters $\theta$ and $\beta$ are discussed in detail in Chapter 4.

### 1.3 Scope of Thesis

This thesis examines the frailty issue in the Cox proportional hazards model by introducing gamma and Weibull frailty into the model and compare these two frailty models through simulation studies and illustrative examples. We assess the performance of parameter estimates and compare between the models. Throughout this thesis, parametric forms of baseline hazard functions and frailty distributions are considered.

In Chapter 2, we review some important concepts relating to gamma and Weibull distributions as a basis for the methods developed later.

In Chapter 3, we develop inferential methods for both gamma and Weibull frailty models. We show that frailties can be integrated out of the conditional likelihood function of the gamma frailty model and apply the classical maximum likelihood estimation techniques to estimate the parameters of interest and their standard errors. For the Weibull frailty model, we show that frailties can not be integrated out of the conditional likelihood function and thus maximum likelihood estimates can not be obtained explicitly. We propose the Monte Carlo integration or quadrature method to approximate the integral to obtain the estimate of the parameters of interest and their standard errors.

In Chapter 4, we conduct simulation studies to investigate the performance of the parameters. The gamma frailty model is the most commonly used model for frailty distribution.
The parameter estimates based on the Weibull frailty model will be compared to the results from the gamma frailty model in order to assess the performance of the Weibull distribution as a frailty model.

In Chapter 5, we use two real data sets to illustrate the methods of inference developed in Chapter 3. Finally, in Chapter 6, we give a brief overview of the work and suggest some possibilities for further research.
Chapter 2

Statistical Distributions for Frailty

The concept of frailty was developed to allow for heterogeneity. The choice of the frailty distribution is very important in the area of frailty models. There are various frailty models that have been developed and suggested in the literature. Any distribution with a positive random variable can be used to model frailty. Unlike standard random effects models, inferential methods have been less developed in frailty models because of censoring and truncation. One-parameter gamma distribution is the most widely used frailty distribution proposed by Clayton (1978), since it is very tractable. Hougaard (1986) suggested the gamma, the degenerate and the inverse gaussian distributions on the positive stable family of distributions for the frailty model. Oakes (1989) suggested the inverse Gaussian and log-normal models for the distribution of the frailty. We will focus on the gamma and Weibull models for the purpose of this thesis.

The shape of the distribution also plays an important role in frailty models. The tails of the distributions can determine the type of dependence a frailty model describes. Distributions with a large right tail such as positive stable distribution lead to strong early dependence, whereas distributions with a large left tail such as gamma and Weibull distributions lead to strong late dependence (Hougaard, 2000).
2.1 Gamma Distribution

Suppose a random variable $T > 0$ is gamma distributed with scale parameter $\lambda > 0$ and shape parameter $\alpha > 0$. The probability density function (pdf) of a random variable $T$ is

$$f_T(t) = \frac{\lambda^\alpha t^{\alpha-1}e^{-\lambda t}}{\Gamma(\alpha)}, t > 0,$$  \hspace{1cm} (2.1.1)

where

$$\Gamma(k) = \int_0^\infty s^{k-1}e^{-s}ds$$  \hspace{1cm} (2.1.2)

is the Gamma function.

The expected value and variance of the gamma distribution are as follows:

$$E(T) = \frac{\alpha}{\lambda}$$

and

$$Var(T) = \frac{\alpha}{\lambda^2},$$

respectively.

The survival and hazard functions of the gamma distribution are given by

$$S(t) = \frac{\Gamma(\alpha,\lambda t)}{\Gamma(\alpha)},$$

$$h(t) = \frac{\lambda^\alpha t^{\alpha-1}e^{-\lambda t}}{\Gamma(\alpha,\lambda t)}, t > 0,$$

respectively, where $\Gamma(\alpha,\lambda t)$ is the upper incomplete gamma function.

The gamma distribution $\text{Gamma}(\alpha, \lambda)$ takes variety of shapes as $\alpha$ varies. When $\alpha=1$, it is identical to the exponential distribution. Figure 2.1 shows the density and hazard functions of gamma-distributed random variables with expectation 1 and variances 1, 0.75, 0.5, and 0.25.

$$h(t) = \begin{cases} \text{decreasing} & \text{when } 0 < \alpha < 1 \\ \text{constant} & \text{when } \alpha = 1 \\ \text{increasing} & \text{when } \alpha > 1 \end{cases}$$
Figure 2.1: Probability density and hazard functions of different gamma distributions with mean 1 and variances 1, 0.75, 0.5, and 0.25

The gamma distribution is very-well known and has simple densities. It is the most common distribution used for describing frailty. Even though gamma models do not have closed form expressions for survival and hazard functions, from a computational view, it fits well to frailty data and it is easy to derive the closed form expressions for unconditional survival and hazard functions. For this reason, this distribution is used often in most applications. Frailties appearing in the conditional likelihood can be integrated out and hence give simple expressions for marginal likelihood. Thus, it is easy to obtain parameter estimates by maximizing the marginal likelihood. The likelihood construction is discussed in more details in Chapter 3.

There are many applications of the gamma frailty model. Lancaster (1979) suggested this model for the duration of unemployment. Aalen (1987) studied the expulsion of intra-uterine contraceptive devices. Ellermann et al. (1992) studied recidivism among criminals using gamma-Weibull model. Andersen et al. (1993) used the gamma frailty model to check the proportional hazards assumptions in his study of malignant melanoma. Vaupel et al. (1979) used the gamma distribution in their studies on population mortality data.
from Sweden. The gamma distribution has two advantages as a frailty distribution. The frailty distribution of the survivors at any given age is again a gamma distribution, with the same parameter and a different scale parameter. The second advantage is that the frailty distribution among the persons dying at any age is also a gamma distribution, with the same shape parameter plus one, and a scale parameter as a function of the age at death. However, there are no known biological reasons which make the gamma distribution preferable than other distributions (Hougaard, 1995).

In gamma frailty models, the restriction $\alpha = \lambda$ is used, which results in expectation of 1. The variance of the frailty variable is then $1 / \lambda$. Assume that the frailty term $u$ is distributed as gamma with $E(U) = 1$ and $Var(U) = \theta$. Then $\lambda = \alpha = 1 / \theta$. The distribution function of the frailty term $u$ is then one-parameter gamma distribution

$$g(u) = \frac{u^{1/\theta - 1} exp(-u/\theta)}{\Gamma(1/\theta) \theta^{1/\theta}}, \theta > 0. \quad (2.1.3)$$

$u > 1$ indicates that individuals in group $i$ are frail, whereas $u < 1$ indicates that individuals are strong and have lower risk.

### 2.2 Weibull Distribution

The Weibull distribution was introduced by Swedish physicist Walloidi Weibull in 1939. He used this distribution to model the distribution of the breaking strengths of materials. It has been widely used in the statistical analysis of lifetime data. Peto and Lee (1973) suggested that failure times should follow Weibull distribution in certain cases as it has a simple hazard function. The flexibility and simplicity of the Weibull distribution makes it very suitable for analysing multivariate lifetime data. The Weibull distribution has been used in the analysis of many types of data including health sciences. Lancaster (1985) used the Weibull distribution in a proportional hazards model to model heterogeneity in duration models. Berry (1975) used this model in designing and analysing carcinogenic experiments.
The survival and hazard functions of the Weibull distribution are in a closed form and hence make Weibull useful as the distribution of the baseline hazard in a frailty model. As a frailty model, Weibull is not as convenient as the gamma model since the frailties can not be integrated out in the conditional likelihood function leading to a nontractable integral which needs to be approximated.

Several methods were developed to avoid integrating out frailties. Wang et al. (2011) used hierarchical likelihood approach for the Weibull frailty model using nonparametric baseline hazard and compared the results to the lognormal and the gamma frailty models. Through simulation studies they concluded that the Weibull frailty model can provide a good fit for survival data besides the gamma and the lognormal frailty models. Boneng (2001) used maximum hierarchical likelihood and restricted maximum hierarchical likelihood methods to estimate the Weibull frailty model with Weibull baseline hazard and compared it to the lognormal frailty model. His simulation studies showed that both of these methods give similar estimates for the regression coefficients. Also, he concluded that the Weibull frailty model gives similar results to those of the lognormal frailty model.

In Chapter 3, we propose some methods to approximate the conditional likelihood function and hence solve for the parameters of interest. One advantage of the Weibull distribution as a frailty distribution is that frailty cannot be negative and Weibull distribution is commonly used to model positive variables.

A continuous lifetime variable $T > 0$ has a Weibull distribution if its survivor, hazard, and cumulative hazard functions are

\[
S(t) = exp\{-\left(\frac{t}{\alpha}\right)^{\eta}\},
\]

\[
h(t) = \frac{\eta t^{\eta-1}}{\alpha^\eta},
\]

and

\[
H(t) = \left(\frac{t}{\alpha}\right)^{\eta}, t > 0,
\]
respectively, where $\alpha > 0$ is the scale parameter and $\eta > 0$ is the shape parameter. The probability density function, mean and variance of $T$ are

$$f_T(t) = \exp \left[ - \left( \frac{t}{\alpha} \right)^\eta \right] \frac{\eta t^{\eta-1}}{\alpha^\eta},$$

(2.2.1)

$$E(T) = \alpha \Gamma \left( 1 + \frac{1}{\eta} \right),$$

and

$$Var(T) = \alpha^2 \left[ \Gamma \left( 1 + \frac{2}{\eta} \right) - \left[ \Gamma \left( 1 + \frac{1}{\eta} \right) \right]^2 \right], t > 0, \alpha > 0, \eta > 0,$$

where $\Gamma()$ is the Gamma function as given in Eq. (2.1.2).

Figure 2.2: Probability density and hazard functions of different Weibull distributions with mean 1 and variances 1, 0.75, 0.5, and 0.25

Weibull model is very flexible in describing lifetime data and its shape parameter allows for different shapes of the hazard function. When $\eta=1$, it is identical to the exponential distribution. Figure 2.2 shows the density and hazard functions of a Weibull distributed
random variable with mean 1 and variances 1, 0.75, 0.5, and 0.25.

\[ h(t) = \begin{cases} 
    \text{increasing} & \text{when } \eta > 1 \\
    \text{constant} & \text{when } \eta = 1 \\
    \text{decreasing} & \text{when } \eta < 1
\end{cases} \]

Weibull model has been used in a wide range of biostatistical applications. For example, time until occurrence of cancer in laboratory animals, time to death by lung cancer are modeled using Weibull distribution. Weibull distribution is considered inappropriate for unimodal or bathtubshaped hazard functions.

Assume that the frailty term \( u \) is distributed as Weibull with mean 1 and variance \( \theta \). Then, the expected value and variance are as follows

\[ E(U) = \alpha \Gamma(1 + 1/\eta) = 1 \]

and

\[ Var(U) = \alpha^2[\Gamma(1 + 2/\eta) - (\Gamma(1 + 1/\eta))^2] = \theta. \]

It is not easy to express the parameters \( \alpha \) and \( \eta \) in terms of \( \theta \) as in the gamma model. However, we can express \( \alpha \) and \( \theta \) as a function of \( \eta \).

\[ \alpha = \frac{1}{\Gamma(1+1/\eta)} \]

and

\[ \theta = \left(\frac{1}{\Gamma(1+1/\eta)}\right)^2[\Gamma(1 + 2/\eta) - (\Gamma(1 + 1/\eta))^2]. \]

Thus, the distribution of the frailty term \( u \) can be written as

\[ g(u) = \text{exp}\left[-\left(\frac{u}{\Gamma(1+1/\eta)}\right)^\eta \frac{\eta u^{\eta-1}}{(\Gamma(1+1/\eta))^\eta}\right] = \text{exp}\{-u\Gamma(1 + 1/\eta)^\eta\} \eta u^{\eta-1}(\Gamma(1 + 1/\eta))^\eta \]

(2.2.2)

where \( \eta = \phi^{-1}(\theta) \). In the next Chapter, we discuss the inferential methods for the gamma and Weibull frailty models.
Chapter 3

Likelihood Approach to Frailty Models

In the literature of statistics, several methods are used to estimate unknown parameters in a frailty model. These methods include maximum likelihood methods that are only applicable in parametric models, the EM algorithm, penalized partial likelihood, Markov Chain Monte Carlo and best linear unbiased predictor (BLUP) method. Maximum likelihood estimation is not straightforward for models that involve intractable integration. Many authors used EM algorithm together with Markov Chain Monte Carlo method to deal with intractable integrals and to obtain the parameter estimates of frailty models. In this thesis, we approximate the integral in the likelihood function directly. The maximum likelihood estimation in the gamma frailty model is straightforward as we can easily integrate the frailties out in the likelihood function and obtain the parameter estimates using classical maximum likelihood techniques. The likelihood function for Weibull frailty model however contains an intractable integral and thus it is not possible to use the classical maximum likelihood techniques for this model. Therefore, approximation techniques are needed. In this thesis, we propose to use the Monte Carlo integration or quadrature method to approximate the nontractable integrals.
3.1 Survival Likelihood

As we know, survival data consists of a combination of event times and censored observations under random right censoring. The likelihood function for survival data is given by

\[ L = \prod_{j=1}^{n} [(1 - G_j(t))f_j(t)]^{\delta_j} [(1 - F_j(t))g_j(t)]^{1-\delta_j}, \]

where \( \delta_j \) is the censoring indicator, \( g \) and \( G \) are the density function and the cumulative distribution function of the censoring time, \( f \) and \( F \) are the density function and the cumulative distribution function of the event time, respectively.

The distribution of censoring times in the likelihood function can be ignored because it does not depend on the parameters of interest related to the survival function. Therefore, assuming right censoring, the likelihood function for the \( j \)th subject is of the form

\[ L = \prod_{j=1}^{n} (f_j(t))^{\delta_j}(S_j(t))^{1-\delta_j}. \]

Following the idea above, the likelihood function for the \( j \)th subject in the \( i \)th subgroup is given by

\[ L_i = \prod_{j=1}^{n_i} (f_{ij}(t))^{\delta_{ij}}(S_{ij}(t))^{1-\delta_{ij}}. \]

Since \( h_{ij}(t) = \frac{f_{ij}(t)}{S_{ij}(t)} \), we can replace \( f_{ij}(t) \) in the likelihood function by \( h_{ij}(t)S_{ij}(t) \). We can rewrite the conditional likelihood function in the form of

\[ L_i = \prod_{j=1}^{n_i} (h_{ij}(t))^{\delta_{ij}}S_{ij}(t). \] (3.1.1)

Following these ideas, we can easily derive the forms of the conditional and marginal likelihood functions of the frailty models. As discussed in Chapter 1, Cox proportional hazards model for frailties is given by

\[ h_{ij}(t) = h_0(t)u_ie^{\beta^Tz_{ij}}, \] (3.1.2)
where \( u_i \)'s are independent and identically distributed random sample from a distribution with mean of 1 and some unknown variance of \( \theta \). Eq. (3.1.2) can be rewritten as

\[
\frac{f(t_{ij})}{S(t_{ij})} = h_0(t)u_ie^{\beta z_{ij}}.
\]

(3.1.3)

Integrating both sides of the Eq. (3.1.3), we can get the expression for the survival function.

\[
\int_0^\infty \frac{f(t_{ij})}{S(t_{ij})} dt = \int_0^\infty h_0(t)u_ie^{\beta z_{ij}} dt
\]

\[- \ln(S_{ij}(t)) = H_0(t)u_ie^{\beta z_{ij}}\]

Therefore,

\[
S_{ij}(t) = \exp(-H_0(t)u_ie^{\beta z_{ij}}).
\]

(3.1.4)

The conditional likelihood function for the \( i \)th subgroup is then given by

\[
L_i(\psi, \beta | u_i) = \prod_{j=1}^{n_i} (h_0(t)u_i e^{\beta z_{ij}})^{\delta_{ij}} e^{-H_0(t)u_i e^{\beta z_{ij}}}.
\]

(3.1.5)

where \( \psi \) is a vector of parameters of the baseline hazard. It follows that, the marginal likelihood function for the \( i \)th subgroup is

\[
L_i(\psi, \theta, \beta) = \prod_{j=1}^{n_i} \int_0^\infty (h_0(t)ue^{\beta z_{ij}})^{\delta_{ij}} e^{-H_0(t)ue^{\beta z_{ij}}} g(u) du,
\]

(3.1.6)

where \( g(u) \) is the probability distribution function of frailties \( u_1, \ldots, u_G \).

### 3.2 Gamma Frailty Model

To obtain the marginal loglikelihood for the gamma frailty model, we proceed as follows.

Let \( u_i \) be independent and identically distributed (iid) sample of gamma random variables with density function
where \( g(u) = \frac{u^{1/\theta - 1}e^{-u/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}}, u > 0, \theta > 0, \)

with \( E(U) = 1 \) and \( Var(U) = \theta \). Larger values of \( \theta \) indicate that there is a higher degree of heterogeneity among groups and strong association within groups.

First, we show that gamma frailties can be integrated out in the conditional survival likelihood. This would lead to explicit and simple marginal likelihood function which only contains the parameters of interest. The marginal likelihood function for the \( i \)th group is given by

\[
L_i(\psi, \theta, \beta) = \prod_{j=1}^{n_i} \int_0^\infty (h_0(t)u e^{\beta' z_{ij}})^{\delta_{ij}} e^{-H_0(t)ue^{\beta' z_{ij}}} \frac{u^{1/\theta - 1}e^{-u/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}} du, \quad (3.2.1)
\]

where \( \psi \) contains the baseline hazard parameters. \( \psi = (\lambda) \) for the exponential baseline hazard and \( \psi = (\eta, \alpha) \) for the Weibull baseline hazard.

Rearranging the terms in Eq. (3.2.1), we obtain the following expression

\[
L_i(\psi, \theta, \beta) = \prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}} e^{\beta' z_{ij}} \frac{\Gamma(1/\theta + d_i) \theta^{1/\theta + d_i}}{\Gamma(1/\theta) \theta^{1/\theta}} \int_0^\infty \frac{u^{1/\theta + d_i - 1}e^{-u/\theta} e^{-u(1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta' z_{ij}})}}{\Gamma(1/\theta + d_i) \theta^{1/\theta + d_i}} du,
\]

where \( d_i = \sum_{j=1}^{n_i} \delta_{ij} \).

To make our problem tractable, we integrate out the frailty term \( u \). The term under the integral is the moment generating function (mgf) of a gamma distribution with a pdf \( \Gamma(1/\theta + d_i, 1/\theta) \). Using this fact, we can derive the expression for marginal likelihood function as

\[
L_i(\psi, \theta, \beta) = \prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}} e^{\beta' z_{ij}} \frac{\Gamma(1/\theta + d_i) \theta^{1/\theta + d_i}}{\Gamma(1/\theta) \theta^{1/\theta + d_i} \theta^{1/\theta + d_i}} \int_0^\infty \frac{u^{1/\theta + d_i - 1}e^{-u(1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta' z_{ij}})}}{\Gamma(1/\theta + d_i)} du.
\]
\[ n \prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}} e^{\beta^T z_{ij} \delta_{ij}} \frac{\Gamma(1/\theta + d_i)}{\Gamma(1/\theta)^{1/\theta} (1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta^T z_{ij}})^{(1/\theta + d_i)}} \]

\[ \int_0^{\infty} u^{1/\theta + d_i - 1} e^{-u(1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta^T z_{ij}})[1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta^T z_{ij}}]} (1/\theta + d_i) \Gamma(1/\theta + d_i) \, du. \]

It is easy to see that the term under the integral is the pdf of \( \Gamma(1/\theta + d_i, 1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta^T z_{ij}}) \), which integrates to 1. Therefore, the obtained marginal likelihood function is

\[ L_i(\psi, \theta, \beta) = \frac{\Gamma(1/\theta + d_i) \prod_{j=1}^{n_i} h_0(t)^{\delta_{ij}} e^{\beta^T z_{ij} \delta_{ij}}}{(1/\theta + \sum_{j=1}^{n_i} H_0(t)e^{\beta^T z_{ij}})^{(1/\theta + d_i)} \Gamma(1/\theta)^{1/\theta}}. \] (3.2.2)

Taking the logarithm of this expression and summing over the \( G \) clusters, we obtain the marginal loglikelihood function, \( l(\psi, \theta, \beta) \).

\[ l(\psi, \theta, \beta) = \sum_{i=1}^{G} [d_i \log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d_i)) - (1/\theta + d_i) \log(1 + \theta \sum_{j=1}^{n_i} H_0(t)e^{\beta^T z_{ij}}) + \sum_{j=1}^{n_i} \delta_{ij}(\beta^T z_{ij} + \log(h_0(t))]. \]

By maximizing this loglikelihood function, we can obtain maximum likelihood estimators for \( \psi, \theta, \) and \( \beta \).

We consider parametric forms of baseline hazards so that the marginal likelihood is also parametric and we can use classical maximum likelihood techniques to estimate the parameters of interest. The hazard and cumulative hazard functions for the exponential distribution are given by

\[ h(t) = \frac{f(t)}{S(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} = \lambda \]

and

\[ H(t) = \lambda t, \]
respectively. Therefore, the marginal loglikelihood function for gamma frailty with exponential baseline hazard rate is
\[
l(\psi, \theta, \beta) = \sum_{i=1}^{G} \left[ d_i \log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d_i)) - (1/\theta + d_i) \log(1 + \theta \sum_{j=1}^{n_i} \lambda t e^{\beta^t Z_{ij}}) + \sum_{j=1}^{n_i} \delta_{ij} (\beta^t Z_{ij} + \log(\lambda)) \right].
\]

The hazard and cumulative hazard functions for the Weibull distribution are given by
\[
h(t) = \frac{\eta t^{\eta-1}}{\alpha^\eta}
\]
and
\[
H(t) = \left(\frac{t}{\alpha}\right)^\eta,
\]
respectively. The marginal loglikelihood function for gamma frailty with Weibull baseline hazard rate is
\[
l(\psi, \theta, \beta) = \sum_{i=1}^{G} \left[ d_i \log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d_i)) - (1/\theta + d_i) \log(1 + \theta \sum_{j=1}^{n_i} (\frac{t}{\alpha})^{\eta} e^{\beta^t Z_{ij}}) + \sum_{j=1}^{n_i} \delta_{ij} (\beta^t Z_{ij} + \log(\frac{\eta t^{\eta-1}}{\alpha^\eta})) \right].
\]

As an example, we derive the first derivatives for the gamma frailty model with Weibull baseline hazard and one covariate.
\[
\frac{\partial l(\eta, \alpha, \theta, \beta)}{\partial \alpha} = \sum_{i=1}^{G} \left[ \frac{(1 + \theta d_i) \eta}{\alpha^2} \sum_{j=1}^{n_i} (\frac{t}{\alpha})^{\eta} e^{\beta^t Z_{ij}} Z_{ij} \right] - \frac{d_i \eta}{\alpha} \tag{3.2.3}
\]
\[
\frac{\partial l(\eta, \alpha, \theta, \beta)}{\partial \eta} = \sum_{i=1}^{G} \left[ -\frac{(1 + \theta d_i) \eta}{\alpha^2} \sum_{j=1}^{n_i} (\frac{t}{\alpha})^{\eta} e^{\beta^t Z_{ij}} \log(t) \right] + \frac{\sum_{j=1}^{n_i} \delta_{ij} (1/ \eta + \log(t))}{(1 + \theta \sum_{j=1}^{n_i} (\frac{t}{\alpha})^{\eta} e^{\beta^t Z_{ij}})} \tag{3.2.4}
\]
\[
\frac{\partial l(\eta, \alpha, \theta, \beta)}{\partial \theta} = \sum_{i=1}^{G} \left[ \frac{-(\frac{1}{\theta} + d_i) \sum_{j=1}^{n_i} \left( \frac{1}{\alpha} \right)^{\eta} e^{\beta Z_{ij}}}{(1 + \theta \sum_{j=1}^{n_i} \left( \frac{1}{\alpha} \right)^{\eta} e^{\beta Z_{ij}})} + \frac{1}{\theta^2} \log(1 + \theta \sum_{j=1}^{n_i} \left( \frac{t}{\alpha} \right)^{\eta} e^{\beta Z_{ij}}) + \frac{d_i}{\theta} + \frac{\Gamma'(\frac{1}{\theta})}{\theta^2 \Gamma(\frac{1}{\theta})} \right]
\] (3.2.5)

\[
\frac{\partial l(\eta, \alpha, \theta, \beta)}{\partial \beta} = \sum_{i=1}^{G} \left[ \frac{-(1 + \theta d_i) \sum_{j=1}^{n_i} \left( \frac{t}{\alpha} \right)^{\eta} e^{\beta Z_{ij}} Z_{ij}}{(1 + \theta \sum_{j=1}^{n_i} \left( \frac{t}{\alpha} \right)^{\eta} e^{\beta Z_{ij}})} + \sum_{j=1}^{n_i} \delta_{ij} Z_{ij} \right]
\] (3.2.6)

The maximum likelihood estimates can be obtained by setting each of the first-order derivatives to 0 and solving for the parameters of interest.

### 3.2.1 Asymptotic Variance-Covariance Matrix

Asymptotic variance-covariance matrix of the MLEs can be derived from the loglikelihood expression. Let \( H \) denote the Hessian matrix of the second partial derivatives of the marginal loglikelihood function \( l(\psi, \theta, \beta) \). The negative of the expected value of the Hessian matrix is known as the Fisher information. Let \( I \) denote the Fisher information matrix.

\[
I(\psi, \theta, \beta) = -E(H(\psi, \theta, \beta)).
\]

The observed information matrix, \( I \), is then the negative of the Hessian matrix.

\[
I(\psi, \theta, \beta) = -H(\psi, \theta, \beta).
\]

Therefore, we can obtain the asymptotic variance-covariance matrix of the estimates by taking inverse of the Fisher information matrix, and the estimated variance-covariance matrix can be obtained by taking the inverse of the observed information matrix by evaluating it at the actual values of the maximum likelihood estimators.
3.3 Weibull Frailty Model

Though Weibull distribution is widely used in survival analysis, its marginal likelihood function is impossible to derive in an explicit form in frailty models. In order to perform parameter estimation, more sophisticated estimation strategies are required. For the gamma frailty model, we obtained the parameter estimates directly from the exact loglikelihood function using the classical maximum likelihood method. Unfortunately, we can not use the same method for the Weibull distribution due to the complexity of the likelihood function. Recall from Chapter 2 that the expectation and variance of a frailty term \( u \) following a Weibull distribution are

\[
E(U) = \alpha \Gamma(1 + \frac{1}{\eta}) = 1
\]

and

\[
Var(U) = \alpha^2 \left[ \Gamma(1 + \frac{2}{\eta}) - \left\{ \Gamma(1 + \frac{1}{\eta}) \right\}^2 \right] = \theta,
\]

respectively. It follows that \( \alpha = \frac{1}{\Gamma(1 + \frac{1}{\eta})} \) and \( \theta = \frac{\Gamma(1 + \frac{2}{\eta})}{(\Gamma(1 + \frac{1}{\eta}))^2} - 1 \), where \( \alpha \) and \( \theta \) are functions of \( \eta \).

If the frailty \( u \) follows a Weibull distribution, then the pdf can be written as

\[
g(u) = e^{-(u\Gamma(1+\frac{1}{\eta}))^\eta} u^{\eta-1} (\Gamma(1 + \frac{1}{\eta}))^\eta, \ u > 0, \eta > 0.
\]

The marginal likelihood function is then given by

\[
L_i(\psi, \eta, \theta, \beta) = \prod_{j=1}^{n_i} \int_{0}^{\infty} (h_0(t) u e^{\beta^T Z_{ij}})^{\delta_{ij}} e^{-H_0(t) u e^{\beta^T Z_{ij}}} e^{-(u\Gamma(1+\frac{1}{\eta}))^\eta} u^{\eta-1} (\Gamma(1 + \frac{1}{\eta}))^\eta du
\]

\[
= \prod_{j=1}^{n_i} (h_0(t) e^{\beta^T Z_{ij}})^{\delta_{ij}} (\Gamma(1 + \frac{1}{\eta}))^\eta \int_{0}^{\infty} u^{\delta_{ij}} e^{-(u\Gamma(1+\frac{1}{\eta}))^\eta} u^{\eta-1} e^{\sum_{j=1}^{n_i} -H_0(t) u e^{\beta^T Z_{ij}}} du,
\]  

(3.3.1)
where $\psi$ is the vector of parameters of the baseline hazard function. Since the expression under the integral is not tractable, we propose to use the Monte Carlo integration method to approximate the integral. The integration part of the likelihood function is

$$
\int_0^\infty u^{d_i} e^{-\left(u\Gamma\left(1+\frac{1}{\eta}\right)\right)^\eta} \eta u^{\eta-1} \sum_{j=1}^{n_i} -H_0(t) e^{\beta^T z_{ij}}
$$

$$
= \int_0^\infty u^{d_i+\eta-1} \eta e^{-\left(u\Gamma\left(1+\frac{1}{\eta}\right)\right)^\eta} \sum_{j=1}^{n_i} -H_0(t) e^{\beta^T z_{ij}}
$$

Using the transformation $v = \sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}$, where $dv = \sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}} du$, the likelihood function can be written as

$$
\int_0^\infty \left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{d_i+\eta-1} \eta exp \left[-\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{\eta}\right] e^{-v} \frac{dv}{\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}}
$$

$$
= \int_0^\infty \left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{d_i+\eta} \eta exp \left[-\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{\eta}\right] e^{-v} \frac{dv}{\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}}
$$

$$
= \int_0^\infty \frac{\Gamma(d_i+\eta)}{\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{d_i+\eta}} \eta exp \left[-\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{\eta}\right] \frac{1}{\Gamma(d_i+\eta)} e^{-v} \frac{dv}{\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}},
$$

where $\frac{1}{\Gamma(d_i+\eta)} e^{-v} \frac{dv}{\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}}$ is the pdf of $\Gamma(d_i+\eta, 1)$.

It follows that the integration that is required is the expectation of

$$
\frac{\Gamma(d_i+\eta)}{\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{d_i+\eta}} \eta exp \left[-\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{\eta}\right]. \tag{3.3.2}
$$

Evaluating this expectation is difficult. However, if we can draw sufficiently large $N$ samples from $\Gamma(d_i+\eta, 1)$, we can estimate the expectation by

$$
\frac{1}{N} \sum_{l=1}^{N} \frac{\Gamma(d_i+\eta)}{\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{d_i+\eta}} \eta exp \left[-\left(\sum_{j=1}^{n_i} H_0(t) e^{\beta^T z_{ij}}\right)^{\eta}\right]. \tag{3.3.3}
$$
This is referred to as Monte Carlo integration. By Strong Law of Large Numbers, as \( N \to \infty \) the estimated expectation will approach the true value.

The approximate marginal loglikelihood function is then given by

\[
\hat{l}(\psi, \eta, \theta, \beta) = \log(L(\eta, \alpha, \theta, \beta)) = \sum_{i=1}^{G} \left[ \sum_{j=1}^{n_i} (\delta_{ij} \log(h_0(t)) + \beta^t Z_{ij}) + \eta \log(\Gamma\left(1 + \frac{1}{\eta}\right)) \right] + \log \left( \frac{1}{N} \sum_{l=1}^{N} \frac{\Gamma(d_i + \eta)}{(\sum_{j=1}^{n_i} H_0(t)e^{\beta^t Z_{ij}})^{d_i + \eta}} \exp \left[ - \left( \frac{v_l \Gamma\left(1 + \frac{1}{\eta}\right)}{\sum_{j=1}^{n_i} H_0(t)e^{\beta^t Z_{ij}}} \right)^\eta \right] \right),
\]

where \( v_1, ..., v_N \) are \( N \) independent realizations of \( \Gamma(d_i + \eta, 1) \). In order to achieve accurate estimates, \( N \) is required to be greater than 10000 (Balakrishnan and Peng, 2006).

If computer time is an issue, we also propose to use quadrature method to approximate the integral in Eq. (3.3.1). This method is available in statistical packages such as R, SAS, and S-PLUS.

\( \psi = (\lambda) \) for exponential baseline hazard and \( \psi = (\eta^*, \alpha^*) \) for the Weibull baseline hazard. The maximum likelihood estimates of \( \beta \) and \( \eta \) can be obtained by maximizing the loglikelihood function. Once we obtain the estimate of \( \eta \), the frailty variance \( \theta \) can be estimated by

\[
\hat{\theta} = \frac{\Gamma\left(1 + \frac{2}{\eta}\right)}{(\Gamma\left(1 + \frac{1}{\eta}\right))^2} - 1.
\]  

(3.3.4)

The first-order derivatives can be approximated with simulations in order to obtain the maximum likelihood estimates. \( \eta^* \) and \( \alpha^* \) are not of interest and hence we will treat them as fixed values in our simulation study. Note that the baseline hazard parameters for the Weibull distribution should be distinguished from the lifetime frailty parameters, \( \eta \) and \( \alpha \).

### 3.3.1 Asymptotic Variance-Covariance matrix

The second-order derivatives with respect to the parameters of interest can be approximated through simulations. The standard errors of the estimates can be approximated from
the inverse of the Hessian matrix. Since we cannot obtain the standard error of the frailty variance directly from the inverse of the Hessian matrix, we will use the Delta method. According to the Delta method, the variance of $\theta$ can be approximated by

$$Var(\hat{\theta}) = (g'(\hat{\eta}))^2 Var(\hat{\eta})$$

(3.3.5)

where

$$g'(\hat{\eta}) = \frac{\Gamma'(1 + \frac{2}{\eta})\Gamma(1 + \frac{1}{\eta})(\frac{-2}{\eta^2}) + \Gamma'(1 + \frac{1}{\eta})\Gamma(1 + \frac{2}{\eta})(\frac{2}{\eta^2})}{(\Gamma(1 + \frac{1}{\eta}))^3}$$

(3.3.6)

and $Var(\hat{\eta})$ can be obtained from the variance-covariance matrix of the MLEs $\eta$ and $\beta$. 

28
Chapter 4

Simulation Studies

In a medical study, patients that are treated in the same hospital may share similar facilities, physicians, nurses, and other care. They may also share some similarities that are not measured as covariates and vary by hospital, such as patient characteristics and medical practice patterns. These hospital effects may create dependence between the study outcomes at each hospital and lead the failure times to correlate. Inferences that ignore these effects can be seriously misleading. In our simulation study, we consider this situation as an example. We treat the clusters as hospitals and failure times as from patients from each hospital. Patients from each hospital are randomly assigned to a control group or treatment group.

We define the following notations to describe the frailty model. Suppose that we observe censored time-to-event data from a medical study with $G$ clusters and $n_i$ subjects per cluster ($i = 1, ..., G$), so that $n = \sum_{i=1}^{G} n_i$ is the total sample size. Let $T_{ij}$ be the failure time and $C_{ij}$ be the censoring time for subject $j$ in cluster $i$. We also observe a vector of covariates $Z_{ij}$. In our simulation example, we consider only one covariate, which is the treatment effect. We observe $\min(T_{ij}; C_{ij})$ and $\delta_{ij} = I\{T_{ij} \leq C_{ij}\}$, where $\delta_{ij}$ is the event indicator ($\delta_{ij} = 1$ if the event has occurred and $\delta_{ij} = 0$ if the lifetime is right-censored).

We carried out simulations to compare the performance of the gamma and Weibull frailty
models with respect to bias and mean squared error (MSE). Settings varied with respect to number of clusters \((n=100, 200, 400)\), number of subjects per cluster \((m_i=2, 4)\), and the frailty distribution. We consider the model with \(n=200\) clusters and \(m_i=2\) subjects in each cluster as the base model. The models with \(n=100\) clusters and \(m_i=4\) subjects in each cluster, and \(n=400\) clusters and \(m_i=2\) subjects in each cluster are compared to the base model to study the effects of increasing the cluster size and the number of clusters separately on the parameter estimates.

4.1 Simulation Algorithm

In order to perform the simulations, we first need to generate the data. The following algorithm is used in order to generate failure times from the frailty distributions. The failure times are assumed to be independent and follow a proportional hazards model, given the frailty. As shown in Chapter 3, the survival function for this model is given by

\[
S_{ij}(t) = \exp(-H_0(t)u_i e^{\beta'z_{ij}}).
\]

Then, the cumulative distribution function of the proportional hazards model is

\[
F_{ij}(t) = 1 - \exp(-H_0(t)u_i e^{\beta'z_{ij}}).
\]

The distribution function follows a uniform distribution on the interval from 0 to 1, denoted by \(U\). If \(U \sim U(0,1)\), then \(1-U \sim U(0,1)\). Therefore, the survival function follows a uniform distribution on the interval from 0 to 1.

\[
U = \exp(-H_0(t)u_i e^{\beta'z_{ij}}) \sim U(0,1).
\]

The failure time \(t\) can be solved by inverting \(H_0\)

\[
t = H_0^{-1}\left(\frac{-\log(U)}{u_i e^{\beta'z_{ij}}}\right),
\] (4.1.1)
where $H_0$ is the cumulative baseline hazard function. When exponential baseline hazard function is considered, $H_0^{-1}(t) = \frac{t}{\lambda}$ and the failure time is given by

$$ t = \frac{-\log(U)}{\frac{1}{u_i e^{\beta Z_{ij}}}}. \quad (4.1.2) $$

If the baseline hazard distribution is Weibull, then $H_0^{-1}(t) = \alpha t^{1/\eta}$ and the failure time is

$$ t = \alpha \left( -\log(U) \right)^{1/\eta} \left( \frac{1}{u_i e^{\beta Z_{ij}}} \right). \quad (4.1.3) $$

Using the expression in 4.1.1, we can easily generate the failure times. The frailty $u_i$ is generated from gamma or Weibull distributions and $U$ is generated randomly from $U(0,1)$. We consider one covariate model, where the covariate is the treatment effect, generated from the binomial distribution.

The censoring times are chosen to follow a uniform distribution on the interval from 0 to 4.5. This allows for censoring rates of approximately 39 per cent when the frailty distribution is gamma and 38 per cent when the frailty distribution is Weibull. Half of the subjects are assigned to treatment group and the other half are assigned to control group. The treatment effect parameter $\beta$ was set to $-\ln(2)=-0.693$ so that the hazard rate of subjects in the control group is twice the hazard rate in the treatment group.

We generated 1000 data sets under each model and applied the gamma and Weibull frailty models for each data set. We estimated the parameters for the variance of the frailty distribution, $\theta$ and fixed treatment effect, $\beta$. We also computed the biases and the MSEs of the estimates. The following formulas are used to find the bias and MSEs.

$$ Bias(\theta) = \frac{\sum_{r=1}^{1000} \theta_r}{1000} - 0.5, \quad Bias(\beta) = \frac{\sum_{r=1}^{1000} \beta_r}{1000} + \ln(2) $$

and

$$ MSE(\theta) = Bias^2(\theta) + \frac{\sum_{r=1}^{1000} Var(\theta_r)}{1000}, \quad MSE(\beta) = Bias^2(\beta) + \frac{\sum_{r=1}^{1000} Var(\beta_r)}{1000}, $$
where $\theta_r$ and $\beta_r$ are the estimates of interest within each of the 1000 simulations. The true value of $\theta$ is assumed to be 0.5 and for $\beta$ it is $-\ln(2)$.

We first assume that the frailty term follows a gamma distribution with the shape parameter 2 and scale parameter 0.5. The mean and variance of the gamma frailty distribution are 1 and 0.5, respectively. For the Weibull frailty model, we chose our shape and scale parameters so that the mean is 1 and variance is 0.5. The shape and scale parameters for the Weibull distribution can be found by solving the following equations:

\[
\alpha = \frac{1}{\Gamma(1 + \frac{2}{\eta})}
\]

and

\[
\left( \frac{\Gamma(1 + \frac{2}{\eta})}{\Gamma(1 + \frac{2}{\eta})^2} - 1 \right) = 0.5,
\]

with mean of 1 and variance of 0.5. Using the \texttt{uniroot} function in R, we found the parameters to be approximately $\eta = 1.435523$ and $\alpha = 1.101321$.

We considered two distributions as the baseline hazard distribution: the standard exponential distribution and the Weibull distribution. For simplicity, we fixed the parameters of each baseline hazard distributions. For the exponential distribution, we assumed $\lambda=1$. For the Weibull distribution, we fixed the mean and variance at 1 and 0.5 respectively and found the corresponding shape and scale parameters. For both frailty models with Weibull baseline hazard, we also study the impact of increasing or decreasing the variance of the Weibull distribution while keeping the mean at 1. We considered the following three cases: $E(T)=1$ and $Var(T)=0.25$, $E(T)=1$ and $Var(T)=0.5$, $E(T)=1$ and $Var(T)=0.75$. The corresponding shape and scale parameters are summarized in Table 4.1.

<table>
<thead>
<tr>
<th>$E(T)$</th>
<th>$Var(T)$</th>
<th>$\eta$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>2.101377</td>
<td>1.129063</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>1.435523</td>
<td>1.101321</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>1.157975</td>
<td>1.052847</td>
</tr>
</tbody>
</table>

Table 4.1: Shape and scale parameters of the Weibull distribution
The simulations are performed using R programming language. For the gamma frailty model, the nonlinear maximisation R function, “nlm”, is used to maximise the likelihood function as a function of the parameters.

In our simulation studies, we applied the quadrature method for the Weibull frailty model as the Monte Carlo integration method requires a lot of computer time. We used the “integrate” function to approximate the integral in the likelihood function and “numDeriv” package to find the Hessian matrix. One problem with the numerical methods is the choice of initial values that need to be considered carefully. A poor choice of initial values may result in non-convergence of the iterative methods.

4.2 Simulation Results

The results of the simulations are summarized in Tables 4.2 and 4.3. The results show that both models performed well in estimating the treatment effect. This indicates that the treatment effect may not be affected by the choice of frailty distribution. Also, using exponential baseline distribution gives very close MSE values in the estimates of the treatment effect to the MSE values when Weibull baseline distribution is used. Balakrishnan and Peng (2006) used the piecewise constant hazard baseline and found that using this distribution tends to give slightly larger MSEs in the estimates of the treatment effect and this could be due to more parameters to be estimated in a frailty model when piecewise constant baseline hazard is used.

We carried out two additional simulations by increasing the sample size of the groups and also subjects within a group. First, we increased the number of subjects in each group from 2 to 4, keeping the total number of subjects at 400. The results of the simulation compared to the original simulation with 2 subjects in each of the 200 clusters, showed that increasing the cluster size and decreasing the number of groups do not affect the MSEs of both estimates of the treatment effect and the frailty variance. The results also show that the
Table 4.2: MSEs and biases (×100) of estimated treatment effect $\beta$ and the frailty variance $\theta$ of frailty models based on 1000 simulated data generated from frailty models with gamma distribution $n = 200, m_i = 2$, $n = 100, m_i = 4$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$\theta$</th>
<th>$\beta$</th>
<th>$\theta$</th>
<th>$\beta$</th>
<th>$\theta$</th>
<th>$\beta$</th>
<th>$\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Hazard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E(T) = 1$</td>
<td>$Var(T) = 0.25$</td>
<td>$0.04$</td>
<td>$1.53$</td>
<td>$0.29$</td>
<td>$1.55$</td>
<td>$0.10$</td>
<td>$1.51$</td>
</tr>
<tr>
<td>$E(T) = 1$</td>
<td>$Var(T) = 0.5$</td>
<td>$0.15$</td>
<td>$1.50$</td>
<td>$1.28$</td>
<td>$1.48$</td>
<td>$0.22$</td>
<td>$1.34$</td>
</tr>
<tr>
<td>$E(T) = 1$</td>
<td>$Var(T) = 0.75$</td>
<td>$0.04$</td>
<td>$1.51$</td>
<td>$0.29$</td>
<td>$1.43$</td>
<td>$0.28$</td>
<td>$1.43$</td>
</tr>
</tbody>
</table>
Table 4.3: MSEs and biases (x100) of estimated treatment effect $\beta$ and the frailty variance $\theta$ of frailty models based on 1000 simulated data generated from frailty models with Weibull distribution

<table>
<thead>
<tr>
<th>Baseline Hazard</th>
<th>$n = 200, m_i = 2$</th>
<th>$n = 100, m_i = 4$</th>
<th>$n = 400, m_i = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>MSE</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Exponential</td>
<td>-0.09</td>
<td>1.48</td>
<td>1.10</td>
</tr>
<tr>
<td>Weibull $E(T) = 1, Var(T) = 0.25$</td>
<td>0.11</td>
<td>1.50</td>
<td>-0.10</td>
</tr>
<tr>
<td>$E(T) = 1, Var(T) = 0.5$</td>
<td>-0.84</td>
<td>1.52</td>
<td>-0.07</td>
</tr>
<tr>
<td>$E(T) = 1, Var(T) = 0.75$</td>
<td>-0.57</td>
<td>1.55</td>
<td>-0.11</td>
</tr>
</tbody>
</table>
biases in the estimates of the frailty variance is slightly larger when the frailty distribution is Weibull compared to the gamma frailty distribution, for both exponential and Weibull baseline hazard distributions. However, these differences are small enough that with an increase in the sample size, these differences become negligible.

We also examined the performance of clusters and their effects on the estimates by increasing the number of clusters from 200 to 400 keeping the number of subjects in each cluster at 2. The results of the simulation showed that MSEs slightly decreased for both estimates compared to the original simulation results, as expected.

We also examined the performance of the models using Weibull baseline hazard with fixed mean at 1 and allowed the variances to vary. The results showed that the MSEs of both estimates may not be affected by the choice of the variance of the baseline hazard distribution. The biases in the estimates of the treatment effect and the frailty variance increased for some cases and decreased for other. However, these changes are very small and may be negligible.

4.3 Model Selection

Several model-selection methods have been proposed in the literature of statistics. The most commonly used methods include information and likelihood based criteria. To compare the gamma and Weibull frailty models, we apply the information based criteria. The most commonly used model selection criteria are the Akaike information criterion (AIC) and Bayesian information criterion (BIC). AIC is given by the expression

$$AIC = -2\log(L) + 2k,$$

where $L$ is the maximized likelihood value and $k$ is the number of parameters in the model. BIC is given by the expression

$$BIC = -2\log(L) + k\ln(N),$$
where $N$ is the total sample size.

To compare the gamma and Weibull frailty models, we used the AIC, using exponential baseline hazard. The model with the smallest AIC value is considered a better fit. We ran 1000 simulations to count the number of times each model is chosen by AIC. Tables 4.4 and 4.5 summarize the results of the simulation. The main diagonals suggest that when fitting the model same as the true model, that model seems to fit the best. The probabilities in the main diagonals are higher than the probabilities in the off diagonals.

To test this hypothesis, we performed two sided proportion tests whether or not the true proportion is 0.5. We tested the proportion $p=0.518$ and with a $p$-value of 0.2549, we conclude that when the true model is gamma, fitting both Weibull and gamma models may give similar results. However, when the true frailty distribution is Weibull, the proportion test suggests that gamma distribution is not a good fit ($p$-value=0.002399).

Table 4.4: Proportion of times the model is selected by AIC (n=200, $m_i=2$)

<table>
<thead>
<tr>
<th>Fitted</th>
<th>True</th>
<th>Gamma</th>
<th>Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>0.518</td>
<td>0.482</td>
<td></td>
</tr>
<tr>
<td>Weibull</td>
<td>0.452</td>
<td>0.548</td>
<td></td>
</tr>
</tbody>
</table>

We performed the similar test by increasing the number of groups from 200 to 400 while keeping the cluster size at 2. We observe that when the sample size increases, the proportion of times the correct model is selected by AIC increases. The results are summarized in Table 4.5. Increasing the number of groups, the proportion test suggests that when the true frailty distribution is gamma, gamma model fits the best and Weibull model is not a good fit. Similarly, when the true frailty distribution is Weibull, the Weibull model fits the best and gamma model is not a good fit ($p$-values>0.05).

Based on the results using standard exponential distribution as the baseline distribution, we would expect the similar outcomes when the Weibull baseline distribution is used. The results of the AIC method show that when the fitted frailty distribution is the same as the true frailty distribution, the models seem to perform the best. This result also agrees with
Table 4.5: Proportion of times the model is selected by AIC (n=400, \( m_i=2 \))

<table>
<thead>
<tr>
<th>True</th>
<th>Fitted</th>
<th>Gamma</th>
<th>Weibull</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma</td>
<td>0.589</td>
<td>0.411</td>
<td></td>
</tr>
<tr>
<td>Weibull</td>
<td>0.425</td>
<td>0.575</td>
<td></td>
</tr>
</tbody>
</table>

the findings of Balakrishnan and Peng (2006).

We plotted the densities of both gamma and Weibull distributions with mean 1 and variance 0.5. As Figure 4.1 shows, when the frailty is greater than 1, gamma and Weibull densities almost overlap. When frailty value is around 0.5, the gamma distribution gives higher density than the Weibull distribution. This also indicates that, the models perform better when the fitted frailty distribution is the same as the true frailty distribution.

Figure 4.1: Gamma and Weibull density with mean 1 and variance 0.5
Chapter 5

Illustrative Examples

In this chapter, we provide two examples to illustrate the methods developed in the preceding sections. We consider McGilchrist and Aisbett’s recurrence times to kidney infection data and Mantel et al.’s litter-matched study of the tumorigenesis of a drug data. We fit the gamma and Weibull frailty models with Weibull baseline hazard and compare the fit of these models using the AIC and BIC methods. We also examine the effect of considering frailty on the coefficients of the treatment effect.

5.1 Recurrence times to kidney infection

In this example, we consider a survival data set introduced by McGilchrist and Aisbett (1991) on the recurrence times of infections of 38 kidney patients using a portable dialysis machine. A catheter is inserted into dialysis patients and time from insertion is recorded and is called infection time (in days). The catheters may be removed for other reasons, therefore censoring is present. The catheter is then inserted after some time and infection time is recorded again. The study recorded two times to recurrence of an infection ($T_1, T_2$) and event indicators ($\delta_1, \delta_2$) for each time. The study emphasized the effects of the risk variables of age, sex, and type of disease that are glomerulo nephritis (GN), acute nephritis (AN), and polycystic kidney disease (PKD) on the recurrence time.
We restrict our analysis to sex ($\beta_1$) and age ($\beta_2$) as covariates and treat the two failure times for each patient as independent. We fit Cox hazard model with and without the frailty. It is important to account for heterogeneity in the Cox hazard model because each patient is considered to have different risks and this is not accounted for by the covariate structure. We use the gamma and Weibull frailties, and Cox proportional hazards model with Weibull baseline hazards to study the effect of including frailty in our models. We model the hazard of infection as a function of the patient’s age and sex.

$$h_{ij}(t) = h_0(t)u_iexp(\beta_1 Sex + \beta_2 Age).$$

The summary of the models are given in Table 5.1.

Table 5.1: The estimates of the regression coefficients, variance component and their standard errors for Infected Kidney Patients Data Set using Weibull Baseline Hazard

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Gamma Frailty</th>
<th>Weibull Frailty</th>
<th>Without Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$ (SE)</td>
<td>0.510(0.255)</td>
<td>0.382(0.227)</td>
<td>-0.875(0.287)</td>
</tr>
<tr>
<td>$\beta_1$ (SE)</td>
<td>-1.912(0.539)</td>
<td>-1.939(0.559)</td>
<td>-0.007(0.012)</td>
</tr>
<tr>
<td>$\beta_2$ (SE)</td>
<td>0.007(0.012)</td>
<td>0.007(0.011)</td>
<td>0.004(0.02)</td>
</tr>
<tr>
<td>Kendall’s $\tau$ (SE)</td>
<td>0.20 (0.08)</td>
<td>0.16 (0.08)</td>
<td>0.004(0.02)</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-332.188</td>
<td>-331.982</td>
<td>-336.554</td>
</tr>
<tr>
<td>AIC</td>
<td>674.376</td>
<td>673.964</td>
<td>681.108</td>
</tr>
<tr>
<td>BIC</td>
<td>686.029</td>
<td>685.617</td>
<td>690.431</td>
</tr>
</tbody>
</table>

Comparing the regression coefficients, we notice that the effect of age and sex is biased downwards when frailty effects are not taken into account, as expected. The gamma and Weibull frailty models are able to account for this unobserved heterogeneity. Note that the regression estimates and their standard errors increase when frailty is introduced into the model.

We performed the Wald test to test the significance of the covariate effects. The Wald statistics for the covariate sex are -3.55 ($p$-value=0.0004) for the gamma frailty model, -3.47 ($p$-value=0.0005) for Weibull frailty model, and -3.05 ($p$-value=0.0023) for the model without frailty. The results indicate that sex has significant effect on the hazard of infection which indicates that males are more susceptible to infection than females. The Wald statistics for
the covariate age are 0.58 (p-value=0.56) for the gamma frailty model, 0.60 (p-value=0.52) for the Weibull frailty model, and 0.2 (p-value=0.84) for the model without frailty. The results indicate the hazard of infection is not affected by age.

The gamma frailty model gives slightly larger frailty variance compared to the Weibull frailty model. The ratio of the frailty variances is $0.510/0.382=1.34$ with a $p$-value of 0.21, suggesting that there is no significant difference in the variances of the frailty parameter for these models. Note that the standard deviations of the frailty variances for both models are large which suggests that there is a possibility of no observed heterogeneity.

To compare the gamma and Weibull frailty models, we use the AIC and BIC methods. Both frailty models have the same number of parameters ($k=5$). Comparing the two models, Weibull frailty model has the smallest AIC and BIC values, suggesting that it is a good fit for this data.

Kendall’s $\tau$ is 0.20 for gamma and 0.16 for Weibull, thus there is on average a positive correlation of 0.20 and 0.16, respectively, between the infection recurrence times.

5.2 Litter-matched study of the tumorigenesis of a drug

We consider the rat tumor data set that was first studied by Mantel et al. (1977). The results of a litter-matched study of the tumorigenesis of a drug was reported. The data set consists of fifty distinct litters with three female rats in each litter. One rat was randomly selected from each litter and was given the drug. The remaining two rats in each litter were selected as controls and given a placebo. The survival time is the time to the development of tumor and is measured in weeks. If death occurred before the appearance of tumor, the observation is right-censored. 40 rats developed a tumor, resulting in about censoring rate of 73 percent.

In this data set, we take the litter effects as clusters. We are interested in assessing whether there are any possible associations between the times to development of tumors
of the litter mates, possibly due to common genetic backgrounds shared by siblings. In particular, we are interested in assessing whether the time to tumor appearance for the control group is longer than in the treatment group, and if the times to tumor appearance are correlated in each litter.

For this data set, the litter mates share the same frailty and a single covariate, $\beta$, which indicates whether the rat was given the drug or placebo ($Z_{ij} = 1$ for treatment and $Z_{ij} = 0$ for control). We fit the gamma and Weibull frailty models with Weibull baseline hazard to the data. In addition, we fit the Cox proportional hazards regression model with Weibull baseline hazard ignoring possible dependence between litter mates. We model the hazard as a function of the treatment group.

$$h_{ij}(t) = h_0(t) u_i \exp(\beta treatment).$$

The summary of the models are given in Table 5.2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Gamma Frailty</th>
<th>Weibull Frailty</th>
<th>Without Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$ (SE)</td>
<td>0.489 (0.469)</td>
<td>0.449 (0.325)</td>
<td>0.831 (0.317)</td>
</tr>
<tr>
<td>$\beta$ (SE)</td>
<td>0.907 (0.322)</td>
<td>0.906 (0.323)</td>
<td></td>
</tr>
<tr>
<td>Kendall’s $\tau$ (SE)</td>
<td>0.20 (0.15)</td>
<td>0.18 (0.11)</td>
<td></td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-241.472</td>
<td>-241.446</td>
<td>-269.208</td>
</tr>
<tr>
<td>AIC</td>
<td>490.945</td>
<td>490.893</td>
<td>542.416</td>
</tr>
<tr>
<td>BIC</td>
<td>502.987</td>
<td>502.936</td>
<td>548.437</td>
</tr>
</tbody>
</table>

When the frailty is ignored, the estimate for $\beta$ and its estimated error is smaller compared to the gamma and Weibull frailty models. This is expected as the frailty models account for the extra variance associated with unmeasured risk factors. The treatment effect for all three models are found to be significant with the Wald statistics 2.62 ($p$-value=0.009) for the Cox proportional hazards model without frailty, 2.81 ($p$-value=0.005) for the Weibull frailty model and 2.82 ($p$-value=0.005) for the gamma frailty model.

Both models give similar estimates of the treatment effect and frailty variance. The estimated frailty variance $\theta$ is 0.489 for the gamma frailty model and 0.449 for the Weibull
frailty model. The ratio of these variances is $0.489/0.449=1.09$, which is close to 1, indicating that there is no difference in the frailty variances. As in the previous example, the standard deviations of the frailty variances for both models are large suggesting the possibility of no observed heterogeneity.

To compare these two frailty models, we used the AIC and BIC methods. Both frailty models have the same number of parameters ($k=4$). Both AIC and BIC values are the smallest when Weibull frailty model is used, suggesting that the Weibull frailty model is the best fit to this data set. In addition, Kendall’s $\tau$ is 0.20 for the gamma frailty model and 0.18 for the Weibull frailty model, indicating that there is on average a positive correlation of 0.20 and 0.18 between the times to tumor development, respectively.

Both examples show that the covariate effect estimates may not be affected by the distribution of frailty as we concluded from our simulation studies.
Chapter 6

Conclusions and Further Work

6.1 Conclusions

This thesis considered maximum likelihood estimation approaches to the frailty models that allow for extra random components in the linear predictors of the Cox proportional hazards model. For the gamma frailty model, we used the classical maximum likelihood techniques to estimate the model parameters. The integral in the likelihood function of the gamma frailty model is tractable and thus parameter estimation is straightforward. For the Weibull frailty model however, the integral in the likelihood function is not tractable. Therefore, it is not straightforward to use the classical maximum likelihood techniques. We proposed Monte Carlo integration method to approximate the integral in the likelihood function in Chapter 3. Though Monte Carlo integration is straightforward, it requires a lot of computer time. Therefore, we also proposed a quadrature method in R. In Chapter 4, we discussed the simulation algorithm and results.

This thesis considered parametric forms of the baseline hazard distributions in the frailty models. We considered exponential and Weibull baseline hazard distributions. The gamma and Weibull frailty models were considered and their performance were compared through simulation studies and illustrative examples.
The simulation study showed very small differences in the bias and MSEs of the estimate of the frailty variance between the gamma and Weibull frailty models. These differences however are very small and may be negligible. The simulation study also showed no noticeable differences in the estimate of the covariate effect in both the gamma and Weibull models. This indicates the the choice of the frailty distribution may not affect the estimation of the covariate effects.

We repeated our simulation study by increasing the cluster size to \( m_i = 4 \) and decreasing the number of groups to \( n = 100 \). The results showed that MSEs of the estimates of both frailty variance and treatment effect were not affected. However, increasing the number of groups to \( n = 400 \) and keeping the cluster size at \( m_i = 2 \) decreased the MSEs of both estimates for both models.

We finally used the AIC method to compare the fit of the Weibull and gamma frailty models. Through Monte Carlo simulations, we observed that the AIC chooses the correct model with higher proportion. Increasing the sample size also increases the proportion of times the correct model is chosen by AIC.

### 6.2 Further Work

For future work, the methods discussed in Chapter 3 may be extended into a model with nonparametric baseline hazard function. When nonparametric form is assumed for \( h_0 \), the semiparametric estimates can be obtained using an Expectation-Maximization (EM) algorithm. Then the loglikelihood function would be

\[
l = l(\theta) + l(\beta, H_0).
\]

The EM algorithm can be used to maximize complex likelihoods. If the likelihood function contains an integral which is not tractable, Markov Chain Monte Carlo (MCMC) method can be used to approximate the integral.
As mentioned in Chapter 4, Monte Carlo integration method requires a lot of computer time. Therefore, we used quadrature method to approximate the integral in the loglikelihood function in our simulation studies and illustrative examples. However, this method can have non-convergence problem if the initial values for the iteration were not chosen carefully. To avoid this problem, different methods can be considered in the future work. One can consider Bayesian approach using Markov Chain Monte Carlo algorithm. Bolstad and Manda (2001) applied a Bayesian approach to the gamma frailty model.

In addition, one can also focus on developing methods to test for heterogeneity. Testing this assumption is very important to decide whether or not to include frailties in a model. Various methods have been developed to test for heterogeneity in Cox proportional hazards model. Liang (1987) developed a score statistic for testing the heterogeneity. One can extend his approach to a case where parametric forms for both baseline hazard and frailty are assumed.
Appendix A

Data Sets

Table A.1: Litter-matched study of the tumorigenesis of a drug

<table>
<thead>
<tr>
<th>Group</th>
<th>Treated Rat</th>
<th>Control Rats</th>
<th>Group</th>
<th>Treated Rat</th>
<th>Control Rats</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>101⁺</td>
<td>104⁺, 49</td>
<td>26</td>
<td>89⁺</td>
<td>104⁺, 104⁺</td>
</tr>
<tr>
<td>2</td>
<td>104⁺</td>
<td>104⁺, 102⁺</td>
<td>27</td>
<td>78⁺</td>
<td>104⁺, 104⁺</td>
</tr>
<tr>
<td>3</td>
<td>104⁺</td>
<td>104⁺, 104⁺</td>
<td>28</td>
<td>104⁺</td>
<td>81, 64</td>
</tr>
<tr>
<td>4</td>
<td>77⁺</td>
<td>97⁺, 79⁺</td>
<td>29</td>
<td>86</td>
<td>94⁺, 55</td>
</tr>
<tr>
<td>5</td>
<td>89⁺</td>
<td>104⁺, 104⁺</td>
<td>30</td>
<td>54</td>
<td>104⁺, 54</td>
</tr>
<tr>
<td>6</td>
<td>88⁺</td>
<td>104⁺, 96</td>
<td>31</td>
<td>76⁺</td>
<td>87⁺, 74⁺</td>
</tr>
<tr>
<td>7</td>
<td>104⁺</td>
<td>94⁺, 77</td>
<td>32</td>
<td>103</td>
<td>84, 73</td>
</tr>
<tr>
<td>8</td>
<td>96⁺</td>
<td>104⁺, 104⁺</td>
<td>33</td>
<td>102</td>
<td>104⁺, 80⁺</td>
</tr>
<tr>
<td>9</td>
<td>82⁺</td>
<td>104⁺, 77⁺</td>
<td>34</td>
<td>80</td>
<td>104⁺, 73⁺</td>
</tr>
<tr>
<td>10</td>
<td>70⁺</td>
<td>104⁺, 77⁺</td>
<td>35</td>
<td>45</td>
<td>104⁺, 79⁺</td>
</tr>
<tr>
<td>11</td>
<td>89⁺</td>
<td>91⁺, 90⁺</td>
<td>36</td>
<td>94</td>
<td>104⁺, 104⁺</td>
</tr>
<tr>
<td>12</td>
<td>91⁺</td>
<td>92⁺, 70⁺</td>
<td>37</td>
<td>104⁺</td>
<td>104⁺, 104⁺</td>
</tr>
<tr>
<td>13</td>
<td>39⁺</td>
<td>50, 45⁺</td>
<td>38</td>
<td>104⁺</td>
<td>101, 94⁺</td>
</tr>
<tr>
<td>14</td>
<td>103⁺</td>
<td>91⁺, 69⁺</td>
<td>39</td>
<td>76⁺</td>
<td>84, 78</td>
</tr>
<tr>
<td>15</td>
<td>93⁺</td>
<td>104⁺, 103⁺</td>
<td>40</td>
<td>80</td>
<td>80, 76⁺</td>
</tr>
<tr>
<td>16</td>
<td>85⁺</td>
<td>104⁺, 72⁺</td>
<td>41</td>
<td>72</td>
<td>104⁺, 95⁺</td>
</tr>
<tr>
<td>17</td>
<td>104⁺</td>
<td>104⁺, 69⁺</td>
<td>42</td>
<td>73</td>
<td>104⁺, 66</td>
</tr>
<tr>
<td>18</td>
<td>104⁺</td>
<td>104⁺, 74⁺</td>
<td>43</td>
<td>92</td>
<td>104⁺, 102</td>
</tr>
<tr>
<td>19</td>
<td>81⁺</td>
<td>104⁺, 69⁺</td>
<td>44</td>
<td>104⁺</td>
<td>98⁺, 78⁺</td>
</tr>
<tr>
<td>20</td>
<td>67⁺</td>
<td>104⁺, 68</td>
<td>45</td>
<td>55⁺</td>
<td>104⁺, 104⁺</td>
</tr>
<tr>
<td>21</td>
<td>104⁺</td>
<td>104⁺, 104⁺</td>
<td>46</td>
<td>49⁺</td>
<td>83⁺, 77⁺</td>
</tr>
<tr>
<td>22</td>
<td>104⁺</td>
<td>104⁺, 104⁺</td>
<td>47</td>
<td>89</td>
<td>104⁺, 104⁺</td>
</tr>
<tr>
<td>23</td>
<td>104⁺</td>
<td>83⁺, 40</td>
<td>48</td>
<td>88⁺</td>
<td>99⁺, 79⁺</td>
</tr>
<tr>
<td>24</td>
<td>87⁺</td>
<td>104⁺, 104⁺</td>
<td>49</td>
<td>103</td>
<td>104⁺, 91⁺</td>
</tr>
<tr>
<td>25</td>
<td>104⁺</td>
<td>104⁺, 104⁺</td>
<td>50</td>
<td>104⁺</td>
<td>104⁺, 79⁺</td>
</tr>
</tbody>
</table>

⁺ Censored observation
Table A.2: Recurrence times to kidney infection

<table>
<thead>
<tr>
<th>Patient</th>
<th>First time</th>
<th>Second time</th>
<th>Censoring first time</th>
<th>Censoring second time</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>92</td>
<td>28</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>447</td>
<td>318</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>12</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>24</td>
<td>245</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>511</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>53</td>
<td>196</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>154</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>333</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>141</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>96</td>
<td>38</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>140</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>536</td>
<td>25</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>185</td>
<td>117</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>292</td>
<td>114</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>22</td>
<td>159</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>15</td>
<td>108</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>152</td>
<td>562</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>22</td>
<td>402</td>
<td>24</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>13</td>
<td>66</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>36</td>
<td>46</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>12</td>
<td>40</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>26</td>
<td>113</td>
<td>201</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>132</td>
<td>156</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>28</td>
<td>34</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>2</td>
<td>25</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>30</td>
<td>130</td>
<td>26</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>31</td>
<td>27</td>
<td>58</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>32</td>
<td>5</td>
<td>43</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>33</td>
<td>152</td>
<td>30</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>34</td>
<td>196</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>35</td>
<td>119</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>36</td>
<td>54</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>37</td>
<td>6</td>
<td>78</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>38</td>
<td>63</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

(Censoring (0); infection occurrence (1); male (1); female (2))
Appendix B

R codes for Chapter 4

Codes in this section are repeated for \( n=100, m_i=4 \) and \( n=400, m_i = 2 \).

B.1 Code for Gamma Frailty

B.1.1 Using exponential baseline hazard

\[
\begin{align*}
\text{cumhaz} &= \sum_{j=1}^{n_i} \lambda t e^{\beta^t Z_{ij}} \\
\lnhaz &= \sum_{j=1}^{n_i} \delta_{ij} \log(\lambda \exp(\beta^t Z_{ij}))
\end{align*}
\]

# Generating the data set
\( i \leftarrow \text{rep}(1:200, \text{rep}(2, 200)) \)
\( j \leftarrow \text{rep}(1:400) \)
\( r \leftarrow 1000 \)

# Covariate
\( Z_i \leftarrow \text{matrix(\text{rbinom(r*400, 1, 0.5), 400})} \)

# Generating frailties
\( u_{ii} \leftarrow \text{matrix(\text{rgamma(200*r, shape=2, scale=0.5), 200})} \)
\( u_i \leftarrow u_{ii} \text{[\text{rep}(1:200, \text{rep}(2, 200)), \text{]} } \)
\( u \leftarrow \text{matrix(\text{runif(r*400, 0, 1), 400})} \)
\( C \leftarrow \text{matrix(\text{runif(r*400, 0, 4.5), 400})} \)
\[ b \leftarrow -\log(2) \]

\[ \text{#Equation 4.1.2} \]

\[ T \leftarrow -\log(u)/(u_i \exp(bZ_i)) \]

\[ \text{time} \leftarrow \text{pmin}(C, T) \]

\[ \text{#Indicator variable} \]

\[ \text{del} \leftarrow \text{as.matrix(time==T)} \]

\[ \text{delta} \leftarrow \text{del}+0 \]

\[ \text{#Number of subjects in each cluster} \]

\[ n \leftarrow \text{length(levels(as.factor(i)))} \]

\[ \text{#Number of failures in each cluster} \]

\[ \text{Di} \leftarrow \text{aggregate(delta, by=list(i), FUN=sum)} \]

\[ \text{#Total number of failures} \]

\[ e \leftarrow \text{colSums(Di[2:(r+1)])} \]

\[ \text{di} \leftarrow \text{Di[2:(r+1)]} \]

\[ \text{output1} \leftarrow \text{vector("numeric", length(1:r))} \]

\[ \text{output2} \leftarrow \text{vector("numeric", length(1:r))} \]

\[ \text{output3} \leftarrow \text{vector("numeric", length(1:r))} \]

\[ \text{output4} \leftarrow \text{vector("numeric", length(1:r))} \]

\[ \text{#Loglikelihood function} \]


\[ \text{for (m in 1:r)} \]

\[ \text{likelihood.exponential} \leftarrow \text{function(p)} \]

\[ p[1] \leftarrow 0 \]

\[ \text{cumhaz} \leftarrow \text{(time}\exp(Zi*p[3]))\exp(p[1]) \]

\[ \text{cumhaz} \leftarrow \text{as.numeric(sapply(split(cumhaz[,m], i), sum))} \]

\[ \text{lnhaz} \leftarrow \text{delta*log(exp(Zi*p[3])}\text{exp(p[1])}) \]

\[ \text{lnhaz} \leftarrow \text{as.numeric(sapply(split(lnhaz[,m], i), sum))} \]

\[ \text{lik} \leftarrow \text{e[m]*log(exp(p[2]))-} \]

\[ \text{sum(((di[,m]+1/exp(p[2]))}\log(1+\text{cumhaz}\text{exp(p[2])})+\text{sum(lnhaz)}-n}\log(\text{gamma}(1/exp(p[2])))\text{+sum(log(\text{gamma}(di[,m]+1/exp(p[2]))))))} \]

\[ \text{lik} \]
# Obtaining the MLEs

```r
initial <- c(0, 0.5, log(0.5))
t <- nlm(likelihood.exponential, initial)
lambda <- exp(t$estimate[1])
theta <- exp(t$estimate[2])
beta <- t$estimate[3]
output1[m] <- theta
output2[m] <- beta
```

# Obtaining the variance of the MLEs

```r
likelihood.exponential2 <- function(p) {
  p[1] <- 1
  cumhaz <- -(time * exp(Zi * p[3])) * p[1]
  cumhaz <- as.numeric(sapply(split(cumhaz[, m], i), sum))
  lnhaz <- delta * log(exp(Zi * p[3]) * p[1])
  lnhaz <- as.numeric(sapply(split(lnhaz[, m], i), sum))
  lik <- e[m] * log(p[2]) -
  sum((di[, m] + 1/p[2]) * log(1 + cumhaz * p[2])) + sum(lnhaz) -
  n * log(gamma(1/p[2])) + sum(log(gamma(di[, m] + 1/p[2])))
  -lik
}
initial1 <- c(lambda, theta, beta)
th <- nlm(likelihood.exponential2, initial1, hessian=T)
hess <- th$hessian[2:3, 2:3]
cov <- solve(hess)
vartheta <- cov[1, 1]
varbeta <- cov[2, 2]
output3[m] <- vartheta
output4[m] <- varbeta
```

```r
theta <- as.vector(output1)
beta <- as.vector(output2)
vartheta <- as.vector(output3)
```
B.1.2 Using Weibull baseline hazard

\[ \text{cumhaz} = \sum_{j=1}^{n_i} \left( \frac{t}{\alpha} \right)^{\eta e^{\beta Z_{ij}}} \]

\[ \ln haz = \sum_{j=1}^{n_i} \delta_{ij} \log(\eta^{t-1} \exp(\beta^t Z_{ij})/\alpha^\eta) \]

# Generating the data set

\text{i} <- \text{rep}(1:200, \text{rep}(2, 200))
\text{j} <- \text{rep}(1:400)
\text{r} <- 1000

# Specifying the shape and scale parameters of Weibull distribution.

\text{alpha} <- 1.052847
\text{eta} <- 1.435523

# Covariate

\text{Zi} <- \text{matrix(\text{rbinom(r*400, 1, 0.5), 400})}

# Generating frailties

\text{uii} <- \text{matrix(rgamma(200*r, shape=2, scale=0.5), 200)}
\text{ui} <- \text{uii[rep(1:200, rep(2, 200)), ]}
\text{u} <- \text{matrix(\text{runif(r*400, 0, 1), 400})}
\text{C} <- \text{matrix(\text{runif(r*400, 0, 4.5), 400})}
\text{b} <- -\log(2)

# Equation 4.1.3

\text{T} <- \text{alpha*(-log(u)/(ui*exp(b*Zi)))}^{(1/\text{eta})}
\text{time} <- \text{pmin(C, T)}

# Indicator variable

\text{del} <- \text{as.matrix(time==T)}
\text{delta} <- \text{del+0}

# Number of subjects in each cluster

\text{n} <- \text{length(levels(as.factor(i)) for each cluster

# Number of failures in each cluster

52
Di<-aggregate(delta,by=list(i),FUN=sum)
#Total number of failures
e<-colSums(Di[2:(r+1)])
di<-Di[2:(r+1)]

output1<-vector("numeric",length(1:r))
output2<-vector("numeric",length(1:r))
output3<-vector("numeric",length(1:r))
output4<-vector("numeric",length(1:r))

#Loglikelihood function
for (m in 1:r){
likelihood.weibull<-function(p){
    p[1]<-log(alpha)
    cumhaz<-exp(Zi*p[3])*(time^(exp(p[4])))exp(p[1])
    cumhaz<-as.numeric(sapply(split(cumhaz[,i],sum)))
    lnhaz<-delta*(Zi*p[3]+log((exp(p[4])*time^(exp(p[4])-1))+exp(p[1])))
    lnhaz<-as.numeric(sapply(split(lnhaz[,i],sum)))
    lik<-e[m]*log(exp(p[2])-sum((di[,m]+1/exp(p[2]))*log(1+cumhaz*exp(p[2])))+sum(lnhaz)-n*log(gamma(1/exp(p[2]))))+sum(log(gamma(di[,m]+1/exp(p[2]))))
    -lik}
}

#Obtaining the MLEs
initial<-c(log(alpha),-log(2),2,log(eta))
t<-nlm(likelihood.weibull,initial)
theta<-exp(t$estimate[2])
beta<-t$estimate[3]
eta<-exp(t$estimate[4])
alpha<-exp(t$estimate[1])
output1[m]<-theta
output2[m]<-beta
# Obtaining the variance of the MLEs


likelihood.weibull2 <- function(p) {
  p[1] <- alpha
  p[4] <- eta
  cumhaz <- exp(Zi * p[3]) * (time^(p[4])) * p[1]
  cumhaz <- as.numeric(sapply(sapply(cumhaz[, m, i], sum)))
  lnhaz <- as.numeric(sapply(sapply(lnhaz[, m, i], sum)))
  lik <- exp(-c[m] * log(p[2]) - sum(di[, m] + 1/p[2]) * log(1 + cumhaz * p[2]) + sum(lnhaz) - n * log(gamma(1/p[2])) + sum(log(gamma(di[, m] + 1/p[2]))))
  -lik
}

initial1 <- c(alpha, theta, beta, eta)

th <- nlm(likelihood.weibull2, initial1, hessian = T)

hess <- th$hessian[2:3, 2:3]

cov <- solve(hess)

vartheta <- cov[1, 1]

varbeta <- cov[2, 2]

output3[m] <- vartheta

output4[m] <- varbeta

theta <- as.vector(output1)

beta <- as.vector(output2)

vartheta <- as.vector(output3)

varbeta <- as.vector(output4)
B.2 Code for Weibull Frailty

B.2.1 Using exponential baseline hazard

\[ cumhaz = \sum_{j=1}^{n_i} \delta_{ij} [\log(\lambda) + \beta^tZ_{ij}] + \eta \log(\Gamma(1 + \frac{1}{\eta})) + \log(\eta) \]

\[ haz2 = \sum_{j=1}^{n_i} -\lambda t e^{\beta^tZ_{ij}} \]

# Generating the data set
i <- rep(1:200, rep(2,200))
j <- rep(1:400)
r <- 1000
# Covariate
Zi <- matrix(rbinom(r*400, 1, 0.5), 400)
# Generating frailties
uii <- matrix(rweibull(r*200, scale = 1.101321, shape = 1.435523), 200)
ui <- uii[, rep(1:200, rep(2,200))]
u <- matrix(runif(r*400, 0, 1), 400)
C <- matrix(runif(r*400, 0, 4.5), 400)
b <- log(2)
# Equation 4.1.2
T <- log(u)/(ui*exp(b*gender))
time <- pmin(C, T)
# Indicator variable
del <- as.matrix(time == T)
delta <- del + 0
# Number of subjects in each cluster
n <- length(levels(as.factor(i)))
# Number of failures in each cluster
Di <- aggregate(delta, by=list(i), FUN=sum)

# Total number of failures
e <- colSums(Di[2:(r+1)])
di <- Di[2:(r+1)]

output1 <- vector("numeric", length(1:r))
output2 <- vector("numeric", length(1:r))
output3 <- vector("numeric", length(1:r))
output4 <- vector("numeric", length(1:r))
output5 <- vector("numeric", length(1:r))

# Log likelihood function
integ <- matrix(c(0), ncol=r, nrow=200)
for (m in 1:r) {
  ll <- function(p) {
    p[1] <- 0
    cumhaz <- delta*(log(exp(p[1]))+Zi*p[3])
    cumhaz <- as.numeric(sapply(split(cumhaz[,m], i), sum))+exp(p[2])*log(gamma(1+1/exp(p[2])))
    haz2 <- -(exp(p[1])*time*exp(Zi*p[3]))
    haz2 <- as.numeric(sapply(split(haz2[,m], i), sum))
    for (s in 1:200) {
      integrand <- function(x) {x^di[s,m]*exp(-x*gamma(1+1/exp(p[2])))^exp(p[2])*x^(exp(p[2])-1)*exp(-x*haz2[s])}
      integ[s,m] <- integrate(integrant, 0, Inf)$value
    }
    int <- as.vector(integ[,m])
    lik <- sum(log(int)+cumhaz)
    -lik
  }

  # Obtaining the MLEs
  initial <- c(0, 1.5, log(0.5))
t <- nlminb(initial, ll)
\[ \lambda < \exp(t \cdot \text{par}[1]) \]
\[ \eta < \exp(t \cdot \text{par}[2]) \]
\[ \beta < -t \cdot \text{par}[3] \]

# Obtaining the variance of the MLEs

\[ \text{ll2} <- \text{function(p)} \{ \]
\[ \text{p[1]} <- 1 \]
\[ \text{cumhaz} <- \text{delta} \ast (\log(p[1]) + Zi \ast p[3]) \]
\[ \text{cumhaz} <- \text{as.numeric(sapply(split(cumhaz[,m], i), sum))} + p[2] \ast \log(\gamma(1+1/p[2])) + \log(p[2]) \]
\[ \text{haz2} <- (p[1] \ast \text{time} \ast \exp(Zi \ast p[3])) \]
\[ \text{haz2} <- \text{as.numeric(sapply(split(haz2[,m], i), sum))} \]

for (s in 1:200) { 
\[ \text{integrand} <- \text{function(x)} \{ x \ast \text{di[s,m]} \ast \exp((-x \ast \gamma(1+1/p[2])) \ast p[2]) \ast x^{(p[2]-1)} \ast \exp(-x \ast \text{haz2[s]}) \} \]
\[ \text{integ[s,m]} <- \text{integrate(integrand,0,Inf)$value} \]
\}
\[ \text{int} <- \text{as.vector(integ[,m])} \]
\[ \text{lik} <- \text{sum(log(int)+cumhaz)} \]
\[ -\text{lik} \]
\}
\[ \text{initial1} <- \text{c(\lambda, \eta, \beta)} \]
\[ \text{t2} <- \text{nlminb(initial1, ll2)} \]

# Using numDeriv package to obtain the Hessian matrix

library(numDeriv)
\[ \text{hess} <- \text{hessian(ll2, t2$par)[2:3,2:3]} \]
\[ \text{cov} <- (\text{diag(solve(hess))}) \]
\[ \text{vareta} <- \text{cov[1]} \]
\[ \text{varbeta} <- \text{cov[2]} \]
\[ \text{output1}[m] <- \lambda \]
\[ \text{output2}[m] <- \eta \]
\[ \text{output3}[m] <- \beta \]
\[ \text{output4}[m] <- \text{vareta} \]
output5 [m]<−varbeta
}
lambda<−as.vector(output1)
etta<−as.vector(output2)
beta<−as.vector(output3)
vareta<−as.vector(output4)
varbeta<−as.vector(output5)

#Equation 3.2.4
theta<−gamma(1+2/eta)/(gamma(1+1/eta))ˆ2−1

#Equation 3.2.6
a<−(digamma(1+2/eta)*(-2/etaˆ2)*gamma(1+1/eta)+(2/etaˆ2)*gamma(1+2/eta)*
digamma(1+1/eta))/(gamma(1+1/eta))ˆ3

#Equation 3.2.5
vartheta<−a^2*vareta

B.2.2 Using Weibull baseline hazard

cumhaz = \sum_{j=1}^{n_i} \delta_{ij} [log(\eta^\eta^{t/\alpha}) + \beta^T Z_{ij}] + \eta log(\Gamma(1 + \frac{1}{\eta})) + log(\eta)

haz2 = \sum_{j=1}^{n_i} -(t/\alpha)^\eta^\eta^{t/\alpha} e^{\beta^T Z_{ij}}

The code is repeated for \alpha=1.101321, \eta^* = 1.435523 and \alpha=1.052847 , \eta^* = 1.157975.

#Generating the data set
i<−rep(1:200, rep(2,200))
j<−rep(1:400)
r<−1000
#Covariate
Zi<−matrix(rbinom(r*400,1,0.5),400)
#Shape and scale parameter values for the baseline hazard
alpha<-1.129063
etaa<-2.101377
u1<-matrix(rweibull(r*200,scale=alpha,shape=etaa),200)
u<-u1[rep(1:200,rep(2,200)),]
u<-matrix(runif(r*400,0,1),400)
C<-matrix(runif(r*400,0,4.5),400)
b<-log(2)

# Equation 4.1.3
T<-alpha*(-log(u)/(ui*exp(b*gender)))^(1/etaa)
time<-pmin(C,T)

# Indicator variable
del<-as.matrix(time==T)

delta<-del+0

# Number of subjects in each cluster
n<-length(levels(as.factor(i)))

# Number of failures in each cluster
Di<-aggregate(delta,by=list(i),FUN=sum)

# Total number of failures
e<-colSums(Di[2:(r+1)])
di<-Di[2:(r+1)]

output1<-vector("numeric",length(1:r))
output2<-vector("numeric",length(1:r))
output3<-vector("numeric",length(1:r))
output4<-vector("numeric",length(1:r))

# Loglikelihood function

integ<-matrix(c(0),ncol=r,nrow=200)
for (m in 1:r){
  ll<-function(p){
    cumhaz<-del*Zi*p[2]+log(((etaa*time^((etaa-1))/alpha^etaa))
    cumhaz<-as.numeric(sapply(split(cumhaz[,i],i),sum))exp(p[1])*log(gamma
      (1+1/exp(p[1])))+log(exp(p[1]))
    haz2<-exp(Zi*p[2])*((time/alpha)^(etaa))
\begin{verbatim}

haz2<-.numeric(sapply(split(haz2[,m], i), sum))
for (s in 1:200){
  integrand<-function(x) {x^di[s,m]*exp(-x*gamma(1+1/exp(p[1]))^exp(p[1]))*
  x^((exp(p[1])-1)*exp(-x*haz2[s]))
  integ[s,m]<-integrate(integrand,0,Inf)$value
}
int<-as.vector(integ[,m])
lik<-sum(log(int)+cumhaz)-lik
}
initial<-c(1.5,log(0.5))
t<-nlminb(initial, ll)
etac<-exp(t$par[1])
beta<-t$par[2]

# Obtaining the variance of the MLEs
ll2<-function(p){
  cumhaz<-delta*(Zi*p[2]+log((etaa*time^(etaa-1))/alpha^etaa))
  cumhaz<-as.numeric(sapply(split(cumhaz[,m], i), sum))+p[1]*log(gamma(1+1/p
       [1]))+log(p[1])
  haz2<-exp(Zi*p[2])*((time/alpha)^(etaa))
  haz2<-as.numeric(sapply(split(haz2[,m], i), sum))
  for (s in 1:200){
    integrand<-function(x) {x^di[s,m]*exp(-x*gamma(1+1/p[1])^p[1])*x^(p[1]-1)
       *exp(-x*haz2[s])
    integ[s,m]<-integrate(integrand,0,Inf)$value
  }
  int<-as.vector(integ[,m])
  lik<-sum(log(int)+cumhaz)-lik
}
initial1<-c(1.5,log(0.5))
t2<-nlminb(initial1, ll2)

\end{verbatim}
#Using numDeriv package to obtain the Hessian matrix

library(numDeriv)
hess <- hessian(ll2, t2$par)
cov <- (diag(solve(hess)))
vareta<-cov[1]
varbeta<-cov[2]
output1[m]<-eta
output2[m]<-beta
output3[m]<-vareta
output4[m]<-varbeta
}
eta<-as.vector(output1)
beta<-as.vector(output2)
vareta<-as.vector(output3)
varbeta<-as.vector(output4)

#Equation 3.2.4
theta<-(gamma(1+2/eta)/(gamma(1+1/eta)))^2-1

#Equation 3.2.6
ac<-(digamma(1+2/eta)*(-2/eta^2)*gamma(1+1/eta)+(2/eta^2)*gamma(1+2/eta)*
digamma(1+1/eta))/(gamma(1+1/eta))^3

#Equation 3.2.5
vartheta<-a^2*vareta

##B.3 Code for AIC selection

The code is repeated for n=400 and m=2 and also for the gamma frailty.
i<-rep(1:200,rep(2,200))
j<-rep(1:400)
r<-1000
Zi<-matrix(rbinom(r*400,1,0.5),400)

#Generating frailties from gamma distribution
#uui<-matrix(rgamma(200*r,shape=2,scale=0.5),200)
# Generating frailties from Weibull distribution

```r
uii <- matrix(rweibull(r*400, scale = 1.101321, shape = 1.435523), 400)
ui <- uii[rep(1:200, rep(2, 200)), ]
uc <- matrix(runif(r*400, 0, 1), 400)
Cc <- matrix(runif(r*400, 0, 4.5), 400)
b <- log(2)
T <- log(u)/uii*exp(b*gender)
time <- pmin(Cc, T)
del <- as.matrix(time == T)
delta <- del + 0
nc <- length(levels(as.factor(i)))
Di <- aggregate(delta, by=list(i), FUN=sum)
e <- colSums(Di[2:(r+1)])
di <- Di[2:(r+1)]
results <- numeric(r)
for (m in 1:r) {
  llg <- function(p) {
    p[1] <- 0
    cumhaz1 <- (time*exp(Zi*p[3])*exp(p[1]))
    cumhaz <- as.numeric(sapply(spli(cumhaz[, m], i), sum))
    lnha <- delta*log(exp(Zi*p[3])*exp(p[1]))
    lnha <- as.numeric(sapply(spli(lnha[, m], i), sum))
    lik <- e[m]*log(exp(p[2])) -
          sum((di[, m]+1/exp(p[2]))*log(1+cumhaz*exp(p[2])))+sum(lnha) -
          n*log(gamma(1/exp(p[2]))) + sum(log(gamma(di[, m]+1/exp(p[2]))))
    - as.numeric(lik)
  }
  integ <- matrix(c(0), ncol=r, nrow=200)
  llw <- function(p) {
    p[1] <- 0
    cumhaz <- delta*(log(exp(p[1]))+Zi*p[3])
    cumhaz <- as.numeric(sapply(spli(cumhaz[, m], i), sum)) + exp(p[2])*log(gamma
      (1+1/exp(p[2]))) + log(exp(p[2]))
```

62
haz2 <- (exp(p[1]) * time * exp(Zi * p[3]))  
haz2 <- as.numeric(sapply(split(haz2[, , i], sum), sum))  
for (s in 1:200) {
  integrand <- function(x) {x^dim [s, , m] * exp(-(x * gamma(1 + 1/exp(p[2]))) * exp(p[2]))) * 
    x^(exp(p[2]) - 1) * exp(-x * haz2[s])}
  integ [s, , m] <- integrate(integrand, 0, Inf)$value
}
int <- as.vector(integ [, , m])  
lik <- sum(log(int) + cumhaz) - lik

initial1 <- c(0, 0.5, log(0.5))  
initial2 <- c(0, 1.5, log(0.5))  
t1 <- nlm(llg, initial1)  
lambda1 <- exp(t1$estimate[1])  
theta1 <- exp(t1$estimate[2])  
beta1 <- t1$estimate[3]  
t2 <- nlminb(initial2, llw)  
lambda2 <- exp(t2$par[1])  
eta2 <- exp(t2$par[2])  
beta2 <- t2$par[3]
if (llg(c(log(lambda1), log(theta1), beta1)) <= llw(c(log(lambda2), log(eta2), 
  beta2))) {count <- -1}
if (llg(c(log(lambda1), log(theta1), beta1)) > llw(c(log(lambda2), log(eta2), 
  beta2))) {count <- 0}
counts <- as.numeric(count)
results [m] <- counts
}

countn <- as.vector(results)
countn[is.na(countn)] <- 0
sum(countn)
Appendix C

R codes for Chapter 5 examples

C.1 Rats

C.1.1 Gamma Frailty

#Data Set
library(survival)
data(rats)

#Renaming the variables
i <- rats$litter
delta <- rats$status
time <- rats$time
Z <- rats$rx

#Number of subjects in a cluster
n <- length(levels(as.factor(i)))

#Number of failures in each cluster
di <- aggregate(delta, by=list(i), FUN=sum)[,2]

#Total number of failures
e <- sum(di)

#Loglikelihood function
likelihood.weibull <- function(p)
{


cumhaz <- \exp(Zi \cdot p[3]) \cdot ((time/\exp(p[1])) \cdot (\exp(p[4])))
cumhaz <- aggregate(cumhaz, by=list(i), FUN=sum)[,2]

lnhaz <- delta \cdot (Zi \cdot p[3] + \log((\exp(p[4]) \cdot \text{time} \cdot (\exp(p[4]) - 1))/\exp(p[1]) \cdot \exp(p[4])))

lnhaz <- aggregate(lnhaz, by=list(i), FUN=sum)[,2]

lik <- \exp(-\log(\exp(p[2]))) -
sum((di+1/\exp(p[2])) \cdot \log(1+cumhaz \cdot \exp(p[2]))) + \text{sum(lnhaz)} -
\text{n} + \log(1/\exp(p[2])) + \text{sum(log(gamma(di+1/\exp(p[2]))))}

# Obtaining the MLEs
initial <- c(0.5, 0.5, 0.5, 0.5)
t <- nlm(likelihood.weibull, initial)
theta <- \exp(t$estimate[2])

beta <- t$estimate[3]
eta <- \exp(t$estimate[4])
alpha <- \exp(t$estimate[1])

# Obtaining the SE of the MLEs

ll <- function(p){
cumhaz <- \exp(Zi \cdot p[3]) \cdot ((time/p[1]) \cdot (\exp(p[4])))
cumhaz <- aggregate(cumhaz, by=list(i), FUN=sum)[,2]

lnhaz <- delta \cdot (Zi \cdot p[3] + \log((p[4] \cdot \text{time} \cdot (p[4]) - 1))/p[1] \cdot p[4]))

lnhaz <- aggregate(lnhaz, by=list(i), FUN=sum)[,2]
lik <- \exp(-\log(\exp(p[2]))) -
sum((di+1/\exp(p[2])) \cdot \log(1+cumhaz \cdot p[2])) + \text{sum(lnhaz)} -
\text{n} + \log(\gamma(1/\exp(p[2]))) + \text{sum(log(\gamma(di+1/\exp(p[2]))))}

lik}
initial1 <- c(alpha, theta, beta, eta)
th <- nlm(ll, initial1, hessian=T)

hess <- th$hessian
```r
var <- diag(solve(hess))
se <- sqrt(var)

# Loglikelihood Value
likelihood.weibull(c(log(alpha), log(theta), beta, log(eta)))

C.1.2 Weibull Frailty

# Loglikelihood Function
ll <- function(p) {
  integ <- vector("numeric", length(1:50))
  cumhaz <- delta*(Zi*p[3]+log((exp(p[4])*time^(exp(p[4]))-1))/exp(p[1])^(exp(p[4])))
  cumhaz <- aggregate(cumhaz, by=list(i), FUN=sum)[2] + exp(p[2]) * log(gamma(1+1/exp(p[2])))
  haz2 <- exp(Zi*p[3]) * (time^(exp(p[4]))) / exp(p[1])^(exp(p[4]))
  haz2 <- aggregate(haz2, by=list(i), FUN=sum)[2]
  for (s in 1:50) {
    integrand <- function(x) {x^di[s] * exp(-(x*gamma(1+1/exp(p[2])))^(exp(p[2])) * x
                         ^(exp(p[2])-1) * exp(-x*haz2[s])}
    integ[s] <- integrate(integrand, 0, Inf)$value
  }
  int <- as.vector(integ)
  lik <- sum(log(int) + cumhaz)
  lik
}

# Obtaining the MLEs
initial <- c(log(100), log(1.4), -0.8, log(1.2))
t <- optim(initial, ll)
lambda <- exp(t$par[1])
et <- exp(t$par[2])
eta <- exp(t$par[3])
et2 <- exp(t$par[4])
```

66
\[ \theta < \gamma (1 + 2/\eta ) / (\gamma (1 + 1/\eta ) )^2 - 1 \]

# Obtaining the variance of the MLEs


\[ l12 \leftarrow \text{function}(p) \{ \right. \]

\[ \text{integ} \leftarrow \text{vector} ("numeric", \text{length}(1:50)) \]

\[ \text{cumhaz} \leftarrow \text{delta} \ast (Zi \ast p[3] + \log ((p[4] \ast \text{time} ^ (p[4] - 1)) / (p[1]^p[4]))) \]

\[ \text{cumhaz} \leftarrow \text{aggregate} (\text{cumhaz}, \text{by}=\text{list}(i), \text{FUN}=\text{sum})[2] + p[2] \ast \log (\gamma (1 + 1/p[2])) + \log (p[2]) \]

\[ p2 \leftarrow \exp (Zi \ast p[3]) \ast (\text{time} ^ (p[4])) / (p[1] \ast p[4]) \]

\[ p2 \leftarrow \text{aggregate} (p2, \text{by}=\text{list}(i), \text{FUN}=\text{sum})[2] \]

\[ \text{for} (s \text{ in } 1:50) \{ \]

\[ \text{integrand} \leftarrow \text{function}(x) \{ x ^ { di[s]} \ast \exp (-x \ast \gamma (1 + 1/p[2])) \ast p[2] \ast x ^ (p[2] - 1) \ast \exp (-x \ast p2[s]) \} \]

\[ \text{integ} [s] \leftarrow \text{integrate} (\text{integrand}, 0, \text{Inf}) \}$value \}

\[ \text{int} \leftarrow \text{as.vector} (\text{integ}) \]

\[ \text{lik} \leftarrow \text{sum}(\text{log}(\text{int}) + \text{cumhaz}) \]

\[ -\text{lik} \}

\[ \text{initial1} \leftarrow c(140, 1.5, 0.9, 3.9) \]

\[ t2 \leftarrow \text{nlminb} (\text{initial1}, l12) \]

# Obtaining the Hessian matrix

\text{library} (\text{numDeriv})

\[ \text{hess} \leftarrow \text{hessian}(l12, t2$par) \]

\[ \text{var} \leftarrow (\text{diag} (\text{solve} (\text{hess}))) \]

\[ \text{se} \leftarrow \text{sqrt} (\text{var}) \]

# Obtaining the SE of the \( \tau \) using Equation 3.2.5

\[ a \leftarrow (-\text{digamma}(1 + 2/\eta) \ast (-2/\eta^2) \ast \text{gamma}(1 + 1/\eta) + (2/\eta^2) \ast \text{gamma}(1 + 2/\eta) \ast \text{digamma}(1 + 1/\eta)) / (\text{gamma}(1 + 1/\eta))^3 \]

\[ \text{vartheta} \leftarrow a^2 \ast \text{cov}[2] \]

\[ \text{settheta} \leftarrow \text{sqrt} (\text{vartheta}) \]

# Loglikelihood Value

\[ l1 (c(\text{log}(\lambda), \text{log}(\eta), \beta, \text{log}(\eta^2))) \]
# Model without frailty

```
fit <- parfm(formula = Surv(time, status) ~ rx,
             data = rats, dist = "exponential", frailty = "none")
```
C.2 Kidney

C.2.1 Gamma Frailty

```r
# Data set
library(survival)
data(kidney)

# Renaming the variables in the data set
kidney$sex <- kidney$sex - 1
i <- kidney$id
delta <- kidney$status
Zi <- kidney$sex
time <- kidney$time
age <- kidney$age

# Number of subjects in a cluster
n <- length(levels(as.factor(i)))

# Number of failures in each cluster
di <- aggregate(delta, by=list(i), FUN=sum) [, 2]

# Loglikelihood function
likelihood.gamma <- function(p){
cumhaz <- exp(Zi * p[3] + age * p[5]) * (time ^ (exp(p[4]))) * exp(p[1])
cumhaz <- aggregate(cumhaz, by=list(i), FUN=sum) [, 2]

lnhaz <- aggregate(lnhaz, by=list(i), FUN=sum) [, 2]
lik <- e * log(exp(p[2])) -
sum((di + 1/exp(p[2])) * log(1 + cumhaz * exp(p[2]))) + sum(lnhaz) -
n * log(gamma(1/exp(p[2]))) + sum(log(gamma(di + 1/exp(p[2])))) - lik
}

# Obtaining the MLEs
```
initial<-c(0.5,0.5,0.5,0.5,0.5)
t<-nlm(likelihood.gamma,initial)
theta<-exp(t$estimate[2])
beta1<-t$estimate[3]
beta2<-t$estimate[5]
eta<-exp(t$estimate[4])
alpha<-exp(t$estimate[1])

# Obtaining the SE of the MLEs
likelihood.gamma2<-function(p){
cumhaz<-aggregate(cumhaz,by=list(i),FUN=sum)[,2]
lnhaz<-aggregate(lnhaz,by=list(i),FUN=sum)[,2]
lik<-e*log(p[2])-sum((di+1/p[2])*log(1+cumhaz*p[2]))+sum(lnhaz)-
n*log(gamma(1/p[2]))+sum(log(gamma(di+1/p[2])))
-lik
}
initial1<-c(alpha, theta, beta1, eta, beta2)
th<--optim(initial1,likelihood.gamma2,hessian=T)
hess<--th$hessian
cov<--solve(hess)
vartheta<--cov[2,2]
varbeta1<--cov[3,3]
varbeta2<--cov[5,5]

# Loglikelihood Value
likelihood.gamma2(c(alpha,theta,beta1,eta,beta2))

C.2.2 Weibull Frailty

# Loglikelihood function

ll.weibull <- function(p) {
  integ <- vector("numeric", length(1:38))
  cumhaz <- aggregate(cumhaz, by=list(i), FUN=sum)[, 2] + exp(p[2]) * log(gamma(1 + 1 / exp(p[2])))
  haz2 <- exp(Zi * p[3] + age * p[4]) * (time ^ (exp(p[5]))) / exp(p[1]) ^ exp(p[5])
  haz2 <- aggregate(haz2, by=list(i), FUN=sum)[, 2]
  for (s in 1:38) {
    integrand <- function(x) {x ^ di[s] * exp(-x * gamma(1 + 1/exp(p[2]))) ^ exp(p[2]) ^ x
                                ^ (exp(p[2]) - 1) * exp(-x * haz2[s])}
    integ[s] <- integrate(integrand, 0, Inf)$value
  }
  int <- as.vector(integ)
  lik <- sum(log(int) + cumhaz)
  -lik
}

# Obtaining the MLEs
initial <- c(log(0.2), log(0.6), -0.5, 0.07, log(0.2))
t <- nlminb(initial, ll.weibull)
lambda <- exp(t$par[1])
eta <- exp(t$par[2])
beta1 <- t$par[3]
beta2 <- t$par[4]
eta2 <- exp(t$par[5])
theta <- gamma(1 + 2 / eta) / (gamma(1 + 1 / eta)) ^ 2 - 1

# Obtaining the SEs of the MLEs
ll2.weibull <- function(p) {
  integ <- vector("numeric", length(1:38))
  ...
cumhaz<-aggregate(cumhaz,by=list(i),FUN=sum)[,2]+p[2]*log(gamma(1+1/p[2]))+
log(p[2])
haz<-aggregate(haz2,by=list(i),FUN=sum)[,2]
for (s in 1:38){
  integrand<-function(x) {x^di[s]*exp(-(x*gamma(1+1/p[2]))^p[2])*x^(p[2]-1)*
    exp(-x*haz[s])}
  integ[s]<-integrate(integrand,0,Inf)$value
}
int<-as.vector(integ)
lik<-sum(log(int)+cumhaz)
lik
}
initial1<-c(0.9,0.6,-1.5,0.007,0.2)
t2<-nlminb(initial1,ll2,weibull)
library(numDeriv)
hess <- hessian(ll2, t2$par)
var <- (diag(solve(hess)))
se<-sqrt(var)
#Obtaining the SE of theta
a<-(digamma(1+2/eta)*(-2/eta^2)*gamma(1+1/eta)+(2/eta^2)*gamma(1+2/eta)*
    digamma(1+1/eta))/(gamma(1+1/eta))^3
vartheta<-a^2*var[2]
sqrt(vartheta)
#Loglikelihood Value
loglv<-ll(c(log(lambda),log(eta),beta1,beta2,log(eta2)))

#Model without frailty
fit <- parfm(formula = Surv(time, status) ~ sex+age,
data = kidney, dist = "weibull", frailty = "none")
Bibliography


