NON-PARAMETRIC COMPETING RISKS WITH MULTIVARIATE FRAILTY MODELS

By

Emad Masuadi

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT
OXFORD BROOKES UNIVERSITY
OXFORD, UNITED KINGDOM
JANUARY 2013

OXFORD BROOKES UNIVERSITY

DEPARTMENT OF

MECHANICAL ENGINEERING AND MATHEMATICAL SCIENCES

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "**Non-parametric Competing Risks with Multivariate Frailty Models**" by **Emad Masuadi** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dated: January 2013

OXFORD BROOKES UNIVERSITY

Date: **January 2013**

Author: **Emad Masuadi**

Title: **Non-parametric Competing Risks with Multivariate Frailty Models**

Department: **Mechanical Engineering and Mathematical Sciences**

Degree: **Ph.D.**

Signature of Author

To my parents, my wife and my children.

Table of Contents

List of Tables

List of Figures

Abstract

This research focuses on two theories: (i) competing risks and (ii) random effect (frailty) models. The theory of competing risks provides a structure for inference in problems where cases are subject to several types of failure. Random effects in competing risk models consist of two underlying distributions: the conditional distribution of the response variables, given the random effect, depending on the explanatory variables each with a failure type specific random effect; and the distribution of the random effect. In this situation, the distribution of interest is the unconditional distribution of the response variable, which may or may not have a tractable form. The parametric competing risk model, in which it is assumed that the failure times are coming from a known distribution, is widely used such as Weibull, Gamma and other distributions. The Gamma distribution has been widely used as a frailty distribution, perhaps due to its simplicity since it has a closed form expression of the unconditional hazard function. However, it is unrealistic to believe that a few parametric models are suitable for all types of failure time.

This research focuses on a distribution free of the multivariate frailty models. Another approach used to overcome this problem is using finite mixture of parametric frailty especially those who have a closed form of unconditional survival function. In addition, the advantages and disadvantages of a parametric competing risk models with multivariate parametric and/or non-parametric frailty (correlated random effects) are investigated. In this research, four main models are proposed: first, an application of a new computation and analysis of a multivariate frailty with competing risk model using Cholesky decomposition of the Log-normal frailty. Second, a correlated Inverse Gaussian frailty in the presence of competing risks model. Third, a non-parametric multivariate frailty with parametric competing risk model is proposed. Finally, a simulation study of finite mixture of Inverse Gaussian frailty showed the ability of this model to fit different frailty distribution. One main issue in multivariate analysis is the time it needs to fit the model. The proposed non-parametric model showed a significant time decrease in estimating the model parameters (about 80% less time compared the Log-Normal frailty with nested loops). A real data of recurrence of breast cancer is used as the applications of these models.

Acknowledgements

First of all, I would like to thank God, the Almighty, for having made everything possible by giving me strength and courage to write this dissertation.

I would like to thank my supervisor Dr. Reza Oskrochi, for his many suggestions, guidance and constant support during this research. I am also tremendously grateful to Prof. Kilani Ghodi, my co-advisor for his critical suggestions which enrich the entire thesis. My deepest thank to my second co-advisor Dr. Hooshang Izadi for his valuable suggestions and comments. I would like to thank the exam committee Dr.Hafiz.T.A.Khan and Dr.Robert Beale for their remarkable suggestions. Of course, I am grateful to my parents, my wife and my kids for their patience and *love* throughout the years of my study and without them this work would never have come into existence. Finally, I wish to thank the following: Dr. Omar Al Atari for the proof reading of this dissertation. Thanks to Prof. Yahia El-Bassiouni for his endless support at the academic as well as the professional level.

Oxford, UK Emad Masuadi

January 31, 2013

Symbols and abbreviations

AFT accelerated failure time

c.d.f cumulative distribution function

$\boldsymbol{E}[Z]$ the expected value of the random variable Z

h(t) hazard function

c.h.f cumulative hazard function

H(t) cumulative hazard function

HR hazard ratio

$L(\theta; t_i)$ likelihood fuction

$\mathcal{L}$ the Laplace transformation

LMMs Linear mixed models

p.d.f probability density function

p.m.f probability mass function

PH proportional hazards

S(t) survival function

$T \sim EXP(\lambda)$ Exponential distribution scale parameter λ

$T \sim \Gamma(\alpha, \lambda)$ Gamma distribution with shape parameter α and scale parameter λ

$T \sim LogL(\alpha, \lambda)$ Log-Logistic distribution with shape parameter α and scale parameter λ

$T \sim IG(\alpha, \lambda)$ Inverse Gaussian distribution with shape parameter α and scale parameter λ

$T \sim LogN(\mu, \sigma^2)$ Log-Normal distribution with location parameter μ and scale parameter σ

$T \sim Weib(\alpha, \lambda)$ Weibull distribution with shape parameter α and scale parameter λ

$\mathbf{Z}_{k\times 1} \sim MIG(\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ Multivariate Inverse Gaussian distribution with parameters $\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\Omega}$

$\boldsymbol{V}[Z]$ the variance of the random variable Z

Statisticians, like artists, have the bad habit of falling in love with their models.
(George Box)

Chapter 1

Introduction

The term survival analysis summarises statistical models and methods for analysing lifetime data or time-to-event data. These models are frequently employed in a variety of disciplines including Bio-statistics, Epidemiology, Engineering, Social Sciences and Economics. Survival analysis differs from other statistical procedures in many features; one of these features is the incompleteness of the survival times due to the censoring mechanism that gives a mixture of discrete and continuous data. Another difference is the shape of the distribution of the survival times that are non-negative random variables and usually skewed to right. The proportional hazards models (PH) by Cox (1972) which assume that covariates have a multiplicative effect on the hazard have dominated survival analysis. In addition, accelerated failure time models are also used for the analysis of survival data by modelling the survival time it-self and the covariates are assumed to act directly on it. During recent decades, these models have been extended to become suitable for handling more complex survival data so as to include frailty models and competing risks models. They provide a powerful tool to analyse models with repeated measures, clustered survival data and multiple types of failure.

One of the main assumptions in analysing survival data is that all subjects have the same risk of failure, which means that populations are homogeneous. However, this is usually not true as different subjects could have different hazards. In univariate survival data, frailty models are used to take into account the heterogeneity between subjects due to exclusion of some

important covariates in the model. In multivariate survival data, frailty models are used when there are repeated measures or clustering. Repeated data occur in case of longitudinal data or multiple recurrences of an event for the same individual. Competing risks models are another form of multivariate survival data where the censoring variable is decomposed into different variables. For each type of failure the subject experiences the event of that failure or is censored.

Competing risks with frailty models frequently arise in a number of substantive scientific research areas, particularly within the Social Sciences, Bio-statistics and Epidemiology. These models combine two theories: (i) competing risks and (ii) random effect (or frailty) models. The theory of competing risks provides a structure for reference to problems where cases are subject to several types of failure, i.e. multiple causes of failure. There are two approaches to analyse competing risks models in the literature. One places emphasis on cause specific hazard functions and sub-distribution functions, while the other uses the concept of latent failure times, where there is an inherent failure time for each type of failure, and only one such time, the smallest, is observable. Both approaches arrive at the same inference, using different notation (Kalbfleisch and Prentice, 2002 and Kundu, 2004). The concept of latent failure times is the one employed in this thesis.

Random effects in competing risk models consist of two underlying distributions: the conditional distribution of the response variables (i.e. failure types), given the random effect, depending on the explanatory variables each with a failure type specific random effect; and the distribution of the random effect in the population (i.e. frailty distribution). In this situation, the distribution of interest is the unconditional distribution of the response variables which may or may not have a tractable form.

Due to its simplicity, the parametric competing risk model where the failure times have a known distribution with monotonically increasing or decreasing baseline hazard and known distribution of random effect, is widely used in practice (Hougaard, 2000, Lambert et al., 2004 and Oskrochi and Crouchley, 2004), but it is unrealistic to believe that a few parametric models are suitable for all types of failure time. Unlike the parametric models, a distribution free model for the random effect in which only the baseline hazard follows a specific distribution will be less demanding in term of assumptions and would be more robust. In this situation the unconditional distribution of the competing risks does not have a tractable form and hence a more complex non-linear multivariate optimisation procedure is needed for parameter estimation (Oskrochi and Davies, 1997). One should differentiate between the ways frailty introduced into the model. In univariate failure time, frailty is included to accommodate heterogeneity between individuals. When individuals in the same group or cluster are assumed to share the same frailty then it accommodates the heterogeneity between clusters not individuals, the so-called shared frailty. Another way to include frailty is by assuming different frailties for different individuals or for different competing risks. In correlated frailty models, frailties are correlated through a covariance matrix and have the same set of marginal distributions but not coming from a multivariate distribution. In multivariate frailty models, frailties have a multivariate distribution with a general correlation structure between the frailties.

This study will investigate the methodology and the applications of competing risk models with multivariate frailty. In the first stage, the Chloeskey decomposition is applied in analysing competing risks model for censored survival data with multivariate Log-Normal frailty to a real data of breast cancer. In the second stage, a correlated Inverse Gaussian frailty as well as a multivariate Inverse Gaussian frailty is proposed. In the third stage, a methodology for analysing a competing risk model with non-parametric multivariate frailty

is proposed. In the last stage, a finite mixture of Inverse Gaussian frailty is proposed. The advantages and disadvantages of competing risk model with non-parametric and/or semi-parametric multivariate frailty is compared using simulations data as well as real data.

In Chapter 2, a general introduction to survival analysis and its main characteristics is summarized along with a description of a set of survival data on the recurrence of breast cancer in the UK, which is used throughout the thesis to demonstrate the development of the proposed models. In Chapter 3, a literature review of the univariate frailty models is conducted. Chapter 4 generalises these models to suite correlated and multivariate frailty models in the presence of competing risks in cases of parametric as well as non-parametric frailty. In chapter 5, a different approach is used to fit frailty models by decomposing the frailty distribution as a finite mixture model (semi-parametric model). Finally, chapter 6 concludes the main results and discusses the advantages and disadvantages of the proposed methodologies.

Chapter 2

Survival Analysis

The term "survival analysis" is used for describing data that measure the time to some event. The statistical analysis of survival data or 'time-to-event' has applications in disciplines as diverse as Medicine, Social Sciences, Engineering, Epidemiology, Economics, as well as many others. Time-to-event could mean the time until some electrical component fails, time of remission of a certain disease after treatment, or time from graduation until employment. These applications have ensured that survival analysis has expanded rapidly in the last three decades. In this study, the applications are within the biomedical framework where real data from medical fields are used and our subjects are individuals. Two features of survival data make them differ from the data used in classical methods (e.g. general linear models). First, there is a mixture of discrete and continuous variables. The time-to-event is the continuous part and the censoring is the discrete part. An individual is said to be censored if s/he does not experience the time of interest before the end of the study: for example, a patient with breast cancer may stay alive after the termination of the study. Second, in classical methods, the dependent variable is modelled through a link function with a linear combination of the explanatory variables. In survival analysis, the model is built either by the hazard function, which represents the failure rate at time t; for example, proportional hazard model (PH) Cox (1972), or by the survival function which represents the probability of surviving beyond time (t), for example, accelerated failure time (AFT) (Lawless, 1982). These two models coincide in the case of Weibull distribution.

2.1 Definitions

2.1.1 Definition. Let T be a non-negative random variable that represents the survival time (failure time, lifetime), of a subject, with probability density function (p.d.f) $f(t)$, and cumulative distribution function (c.d.f) $F(t) = \Pr(T \leq t)$.

If T is absolutely continuous then the probability density function is

$$f(t) = \lim_{\Delta t \to 0^+} \frac{P(\text{Failure occurs in}[t, t + \Delta t))}{\Delta t}.$$

2.1.2 Definition. The **survival function** $S(t) = \Pr(T \geq t)$, is the probability of an individual surviving beyond time t, or more generally, the probability that the event of interest has not occurred by duration t.

From the definition, if it is a continuous random variable then,

$$S(t) = \Pr(T \geq t) = \int_t^\infty f(u)du.$$

If T is discrete with mass points at t_j with probability mass function (p.m.f)$f_j = P(T = t_j)$ then, for $t_j \leq t < t_{j+1}$,

$$S(t) = \sum_{i \geq j} f(t_i).$$

2.1.3 Definition. If T is an absolutely continuous random variable then, the **hazard function** is given by

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \tag{2.1.1}$$

which represents the probability of failure at time t given that the individual survives up to

time t. If T is a discrete random variable, then the hazard function is given by

$$h(t) = P(T \geq t | T = t) = \frac{\Pr(T = t)}{P(T \geq t)} = \frac{f(t)}{S(t^-)},$$

where $S(t^-) = \lim_{x \to t^-} S(x)$ (t^-: from left). Moreover,

$$S(t) = \prod_{i=1}^{j} \frac{S(t_{i+1})}{S(t_i)} = \prod_{i=1}^{j} \frac{S(t_i) - f_i}{S(t_i)} = \prod_{i=1}^{j} \left(1 - \frac{f_i}{S(t_i)}\right)$$

and hence,

$$S(t) = \prod_{i=1}^{j} \left(1 - h(t_i)\right).$$

2.1.4 Definition. If T is an absolutely continuous random variable then, the cumulative hazard function c.h.f is

$$H(t) = \int_0^t h(u)du = -\log S(t). \tag{2.1.2}$$

Or equivalently,

$$S(t) = e^{-H(t)} = \exp\left[-\int_0^t h(u)du\right].$$

The p.d.f $f(t)$ can be written in terms of the hazard and the cumulative hazard function,

$$f(t) = h(t)exp[-H(t)].$$

If T is discrete,

$$H(t) = \sum_{i \leq j} h(t_i)$$

Sometime it is desirable to find the mean or the expected lifetime of subjects, for instance if T is a continuous random variable with p.d.f $f(t)$ then expected value of T is

$$\mu = \int_0^\infty t f(t) dt.$$

Another way to get this expected value is by integrating the survival function, assuming that the event of interest is bound to occur (i.e. $S(\infty) = 0$)

$$\mu = \int_0^\infty S(t) dt.$$

2.2 Censoring

As mentioned above, censoring is one of the reasons that survival analysis differs from standard statistical analysis, so censored data are those observations whose time-to-event is not observed before the end of the study. There are three different mechanisms of censoring: *type I*, *type II*, and *random censoring*. In *type I censoring*, a sample of n subjects is followed for a specific time T^* under the control of the researcher, so that the total duration of the study is fixed whilst the number of subjects who experience the event of interest is random, the actual failure time t_i cannot be observed if $t_i > T^*$. This type of censoring is usually used in medical applications. The opposite of this mechanism, *type II censoring,* occurs when a sample of n subjects is followed until the failure time of the first $r(r \leq n)$ of the subjects is observed. This type of censoring is usually used in industrial applications. Another possible mechanism of censoring is *random censoring*, where each subject has associated with it a potential censoring time C_i and a potential survival time T_i, which are usually assumed to

be independent of one another (the so-called independent censoring). This type of censoring will be the main censoring mechanism that is used within this thesis. Usually, the observed variables are $Y_i = \min(T_i, C_i)$, and the indicator variable δ_i.

$$Y_i = \begin{cases} T_i & if \quad T_i \leq C_i \\ C_i & if \quad T_i > C_i \end{cases}, \quad \text{and} \quad \delta_i = \begin{cases} 1 & if \quad T_i \leq C_i \\ 0 & if \quad T_i > C_i. \end{cases}$$

The observed data takes the form, $(y_1, \delta_1), ..., (y_n, \delta_n)$, and possibly some factors (independent variables). There are three different kinds of censoring, right-censoring, left-censoring, and interval-censoring:

Right-censoring: when subjects leave the study or the study ends before observing their survival (failure) time. It is only known that their survival time T_i lies in an interval (t, ∞). Throughout this thesis, right-censoring is assumed.

Left-censoring: when subjects experience the event (failure) before a certain duration. It is only known that their survival time T_i lies in an interval $[0, t)$.

Interval-censoring: when it is not clear when the event occurred. It is only known that the time-to-event occurred within some interval $[t_1, t_2)$. For more information see Lee and Wang (2003).

2.3 Non-parametric survival distribution

When it is difficult to determine the distribution of the survival time, or no assumption about the distribution is made, non-parametric or distribution-free survival time approaches represent viable alternatives. In this case, the empirical distribution function is used to estimate the survival function assuming that all subjects have experienced failure (i.e., no censored data). The empirical survival function is given by

$$\tilde{S}(t) = \frac{Number\ of\ subjects\ with\ survival\ times \geq t}{Number\ of\ subjects\ in\ the\ data\ set}.$$

The following subsections give a briefly review of the two most popular estimators of the survival function using both censored and uncensored data.

2.3.1 Life-table estimator

In the case of some subjects with censored time, the empirical function is not applicable any more. Life-table estimator divides the study time into, usually equal intervals. The interval width and number of intervals varies from one study to another depending on the length of the study and the number of observations. Suppose that $(t_1, ..., t_k)$ are the boundaries of these intervals and let d_i,c_i and n_i denote the number of failures, number of censored subjects and number of subjects who are at risk during the interval $[t_i, t_{i+1})$ respectively. Assuming that the censoring is uniformly distributed along the interval, the average number of subjects at risk during the interval $[t_i, t_{i+1})$ is $n_i^* = n_i - c_i/2$, and hence the probability of failure is d_i/n_i^*. For $t \in [t_j, t_{j+1})$, and $j = 1, ..., k$, the life-table estimator of the survival function is given by

$$\prod_{i=1}^{j} \left(1 - \frac{d_i}{n_i^*}\right).$$

2.3.2 Product-limit estimator

Kaplan and Meier (1958) provided a special case of the life-table estimator, where the interval boundaries are chosen so that there is at least one failure. Suppose that $(t_{(1)}, ..., t_{(m)})$are the ordered time points in which there is at least one failure so that, $t_{(1)} < t_{(2)} < \cdots < t_{(m)}$. For $t \in [t_{(j)}, t_{(j+1)})$, and $j = 1, ..., m$, the product-limit estimator of the survival function is given

by

$$\prod_{i=1}^{j}\left(1-\frac{d_i}{n_i}\right).$$

2.4 Parametric survival distribution

Although the non-parametric methods mentioned above are widely used in applications and do not require any specific assumptions they are unsuitable for handling complex data sets with explanatory variables, and if the distribution of the survival time is known, the inferences will be more accurate. This section reviews some of the most commonly used distributions for survival time. Actually, any non-negative random variable, whether discrete or continuous, can be used to describe the survival time, while in this thesis our focus will be on continuous distributions. Other random variables defined over the real line can be used say, $x \in (-\infty, \infty)$ such that $x = log(t)$ or equivalently $t = e^x$. There are some distributions that have been used frequently in the literature of survival analysis, such as the Exponential, Weibull, Gamma, Log-Normal and Log-Logistic distributions. For each distribution, the probability density function, survival function, hazard function, expected value and the variance of survival time are summarised in Table 2.1.

The Log-Normal and the Gamma distributions are generally less convenient computationally, but are still frequently applied, as well as non-parametric approaches, such as the product limit estimator suggested by Kaplan and Meier (1958) and related techniques. The advantages and disadvantages of different parametric, semi-parametric and non-parametric models as methodologies for statistical inference can be found in books such as Kalbfleisch and Prentice (2002), Miller (1981), Lawless (1982), Cox and Oakes (1984) and Klein and Moeschberger (1997).

2.4.1 Exponential distribution

The simplest distribution for survival time is the Exponential distribution ($T \sim EXP(\lambda)$), especially used in reliability analysis in engineering applications. The p.d.f, the mean and the variance of T are given in Table 2.1. It is used to model data with a constant failure rate (indicated by the hazard plot which is simply equal to a constant). The exponential distribution is a member of the exponential family. It has a unique property of "lack of memory", because of its constant hazard rate λ. The probability to failure within a particular time interval depends only on the length, not on the location of this interval. In real-world applications, the assumption of a constant rate is rarely satisfied.

2.4.2 Weibull distribution

The Weibull model is the most widely used parametric survival model. The Weibull distribution was introduced by Weibull (1939); it is an important generalisation of the exponential distribution with two positive parameters $T \sim Weib(\alpha, \lambda)$, where α is the shape parameter and λ is the scale parameter. Its first parameter allows different failure rates; if it is less than one this indicates a decreasing hazard function while a value more than one indicates an increasing hazard function, but if it is one then the distribution becomes the exponential distribution and the hazard function is constant. Figure 2.1 describes the density and the hazard curves of the Weibull distribution with a scale parameter that equals one with different values of the shape parameter . The Weibull model also has another property in the sense that if the plot of $log(-log(\hat{S}(t)))$ against $log(t)$ shows a linear trend, that would suggest Weibull model. $\hat{S}(t)$ is the empirical survival function which can be obtained the by the Kaplan-Meier estimate.

2.4.3 Gamma distribution

Another possible distribution of the survival time is the Gamma distribution $T \sim \Gamma(\alpha, \lambda)$ with two positive parameters: α the shape parameter and $\frac{1}{\lambda}$ the scale parameter. Like the Weibull distribution, it includes the Exponential distribution as a special case when $\alpha = 1$. The Gamma distribution is of limited use in survival analysis because the Gamma models do not have closed-form expressions neither for the survival nor the hazard function, if α is not an integer since both include the incomplete Gamma integral (If α is an integer then the distribution reduces to Erlang distribution). Its maximum likelihood estimation is difficult to get and involves incomplete Gamma integrals. This requires additional numerical calculations in parameter estimation. It can be shown that the limit of its hazard as time goes to infinity is equal to the shape parameter ($\lim_{t\to\infty} h(t) = \lambda$). The Gamma hazard increases monotonically if $\alpha > 1$, from a value of zero at the origin to a maximum of λ; and is constant if $\alpha = 1$; and decreases monotonically if $\alpha < 1$, from infinity at the origin to an asymptotic value of λ. Figure 2.2 describes the density and the hazard curves of the Gamma distribution with a scale parameter which equals one with different values of the shape parameter.

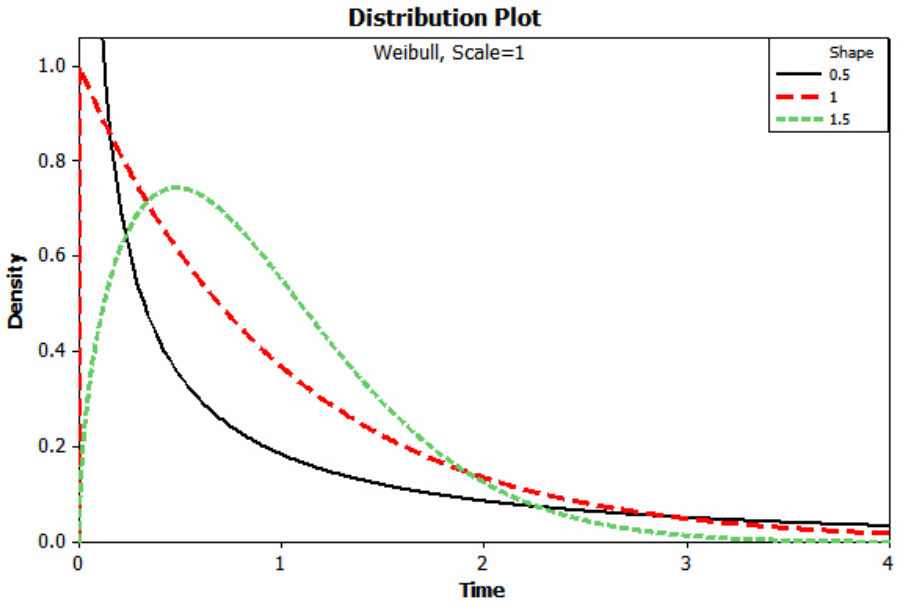
2.4.4 Log-Normal distribution

A random variable has a log-Normal distribution if the logarithm of the random variable is normally distributed $T \sim LogN(\mu, \sigma^2)$ if and only if $log(T) \sim N(\mu, \sigma^2)$. The log-Normal distribution is self-replicating under multiplication and division. That is, multiplying or dividing Log-Normal random variables will result in Log-Normal distributions. The hazard function of the log-Normal differs from the previous two distributions; it starts from zero at $t = 0$, increases to a maximum and then decreases and approaches zero as time goes to infinity. The decreasing form of the hazard function as time increases makes the distribution

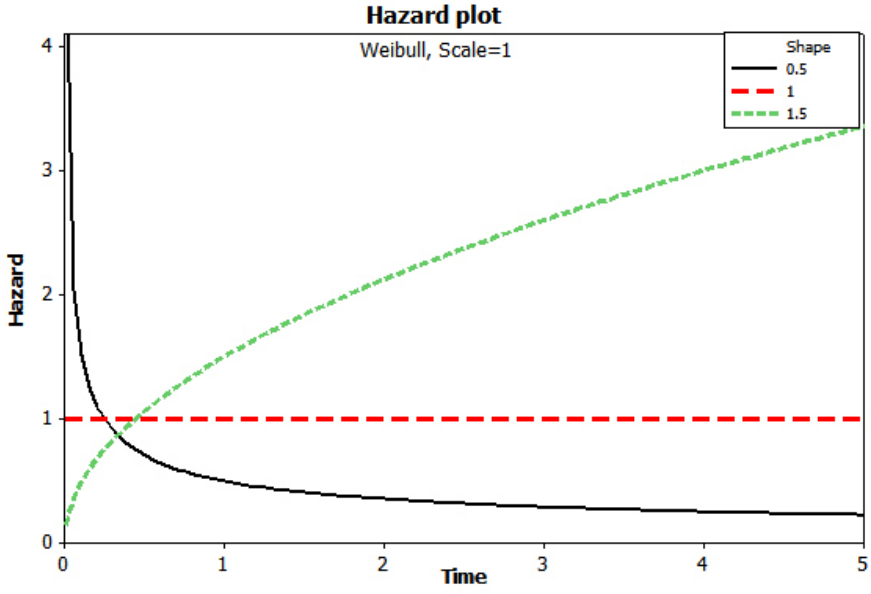
unsuitable to model lifetime data in most medical applications. However, the Log-Normal distribution can be very useful for representing lifetimes for situations with non-monotonic hazards such as the analysis of electrical insulation or time to occurrence of lung cancer among smokers. Figure 2.3 shows the density and the hazard curves of the Log-Normal distribution with a location parameter which equals one with different values of the shape parameter. It is similar to the Gamma distribution in the complexity of the hazard function since numerical integration needs to be used to fit the distribution.

2.4.5 Log-Logistic distribution

The Log-Logistic distribution is the probability distribution of a random variable whose logarithm has a logistic distribution $T \sim LogL(\alpha, \lambda)$. It is one of the parametric survival time models in which the hazard rate may be decreasing (if $\lambda \leq 1$), increasing, or hump-shaped (if $\lambda > 1$), that is, it initially increases and then decreases. Figure 2.4 describes the density and the hazard curve of the log-logistic distribution with scale parameter equals one with different values of the shape parameter. It has also another property that if the plot of the *logit* of the survival function $S(t)$, $(log(\frac{S(t)}{1-S(t)}))$ against $log(t)$ has a linear trend then it is an indication of log-logistic model.

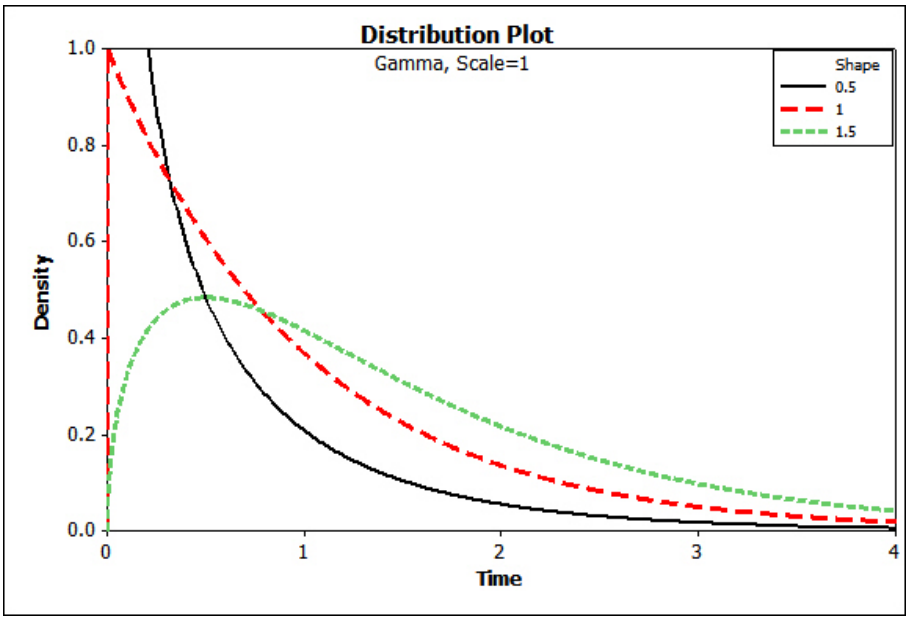


(a) Weibull densities

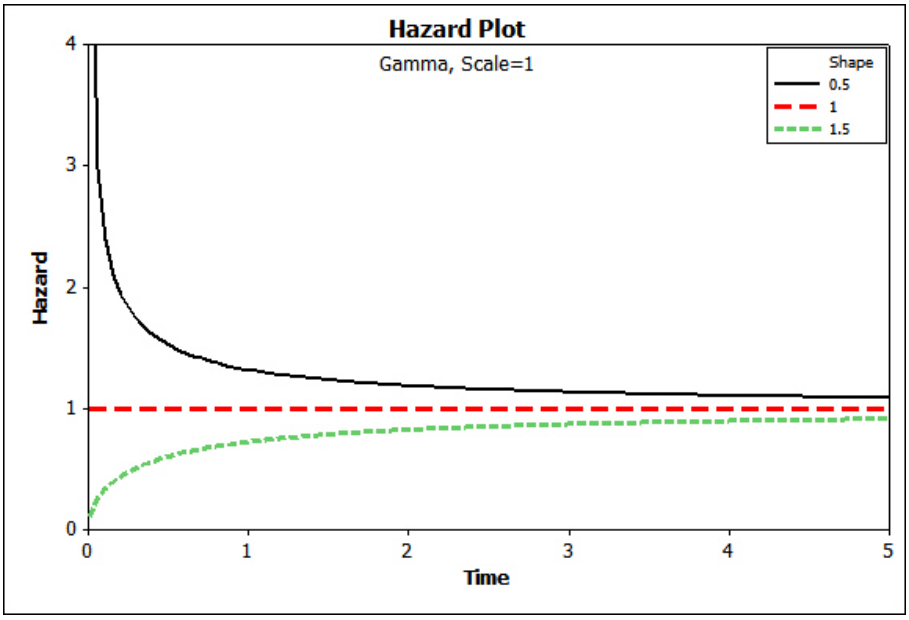


(b) Weibull hazards

Figure 2.1: Weibull densities and hazards with scale parameter $\lambda = 1$ and shape parameter $\alpha = (0.5, 1, 1.5)$



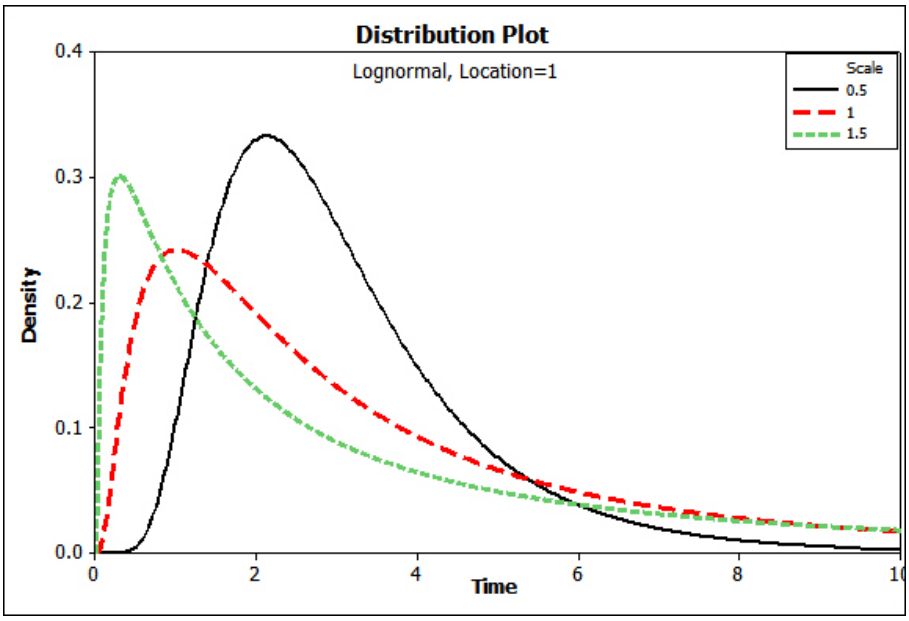
(a) Gamma densities



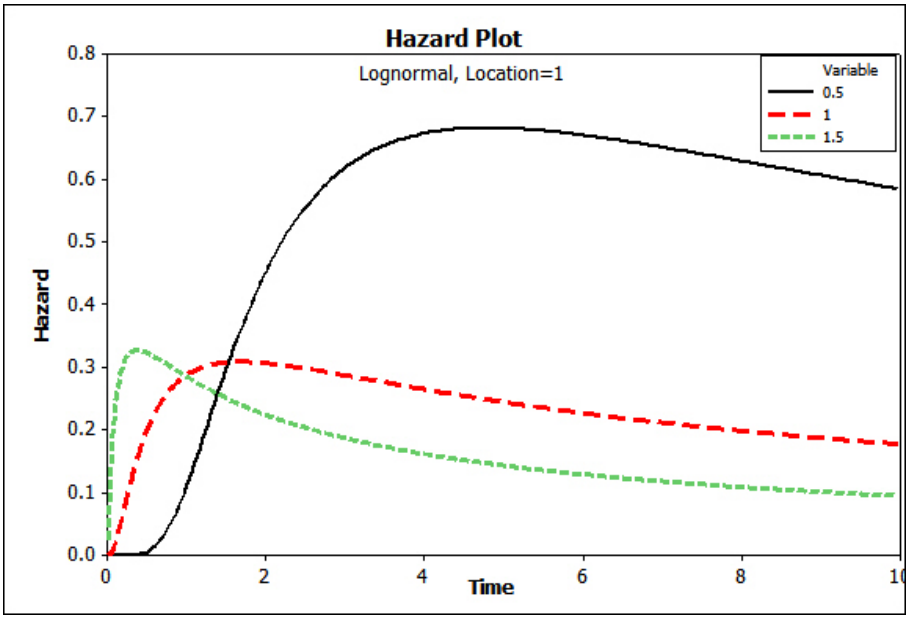
(b) Gamma hazards

Figure 2.2: Gamma densities and hazards with scale parameter $\lambda = 1$ and shape parameter $\alpha = (0.5, 1, 1.5)$

.



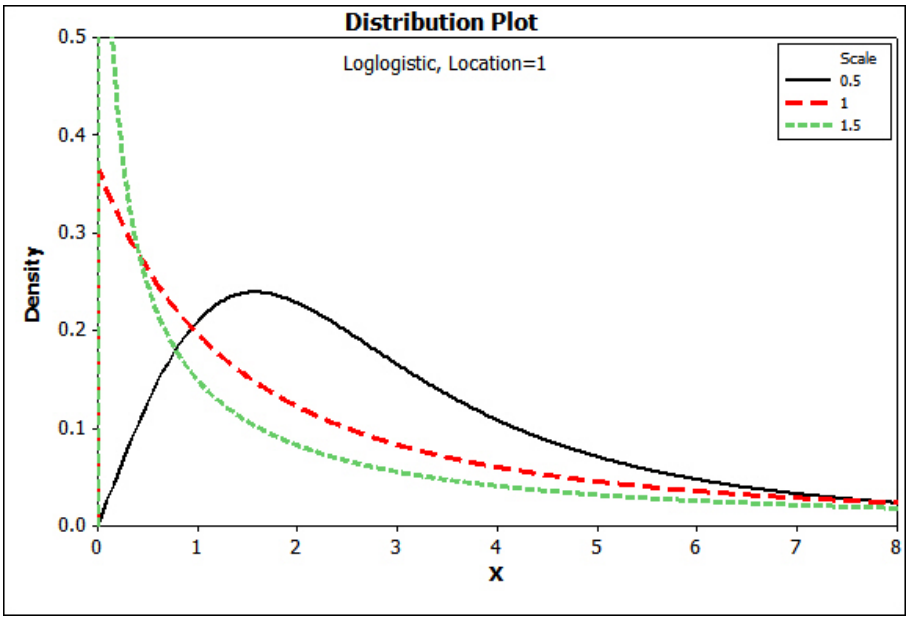
(a) Log-Normal densities



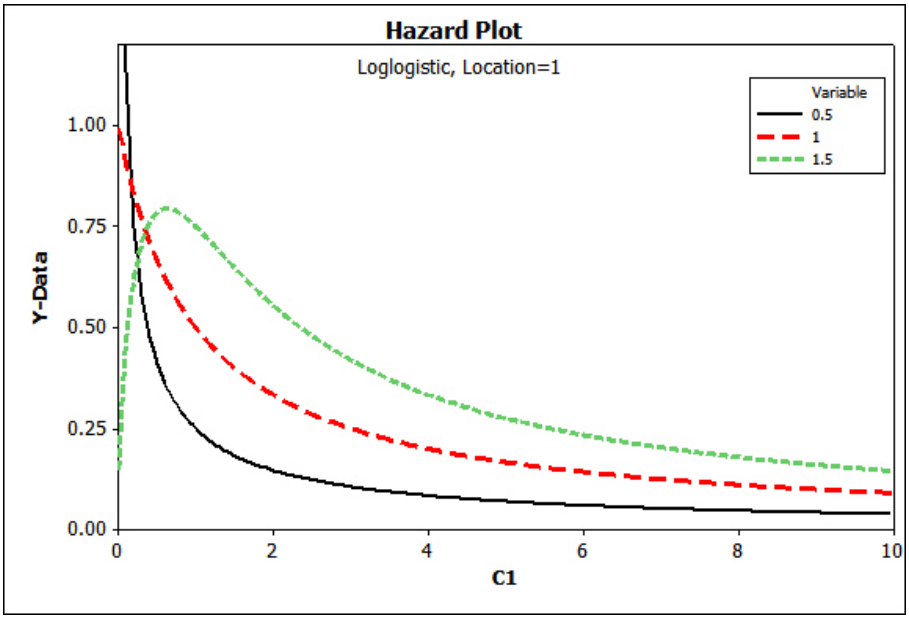
(b) Log-Normal hazards

Figure 2.3: Log-Normal densities and hazards with location parameter ($\mu = 1$) and shape parameter($\alpha = (0.5, 1, 1.5)$)

.



(a) Log-Logistic densities



(b) Log-Logistic hazards

Figure 2.4: Log-Logistic densities and hazards with scale parameter $\lambda = 1$ and shape parameter $\alpha = (0.5, 1, 1.5)$

.

Table 2.1: Some common parametric survival distribution along with their associated functions

Distribution	**Probability density function** $f(t)$	**Survival function** $S(t)$	**Hazard function** $h(t)$	**Cum. hazard function** $H(t)$	**Mean** $E(T)$	**Variance** $Var(T)$
Exponential	$\lambda exp(-\lambda t)$, $\lambda > 0, t \geq 0$	$exp(-\lambda t)$	λ	λt	$(1/\lambda)$	$\left(1/\lambda^2\right)$
Weibull	$\alpha\lambda t^{\alpha-1} exp(-\lambda t^{\alpha})$, $\alpha, \lambda > 0$, $t \geq 0$	$exp(\lambda t^{\alpha})$	$\alpha\lambda t^{\alpha-1}$	λt^{α}	$\frac{\Gamma(1+(1/\alpha))}{\lambda^{(1/\alpha)}}$	$\frac{\Gamma(1+2/\alpha)-\Gamma(1+1/\alpha)^2}{\lambda^{(2/\alpha)}}$
Gamma	$\frac{1}{\Gamma(\alpha)}\lambda^{\alpha} t^{\alpha-1} exp(-\lambda t)$, $\alpha, \lambda > 0, t \geq 0$	$1 - I_{\alpha}(\lambda t)^{*}$	$\frac{\lambda^{\alpha} t^{\alpha-1} exp(-\lambda t)}{(1-I_{\alpha}(\lambda t))\Gamma(\alpha)}$	—	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
Log-Normal	$\frac{exp\left(\frac{1}{2\sigma^2}(ln(t)-\mu)^2\right)}{t\sigma\sqrt{2\pi}}$, $\sigma > 0, t \geq 0$	$1 - \Phi\left(\frac{ln(t)-\mu}{\sigma}\right)^{**}$	$\frac{f(t)}{S(t)}$	—	$exp(\mu + \frac{\sigma^2}{2})$	$(e^{\sigma^2} - 1)e^{(\sigma^2+2\mu)}$
Log-Logistic	$\frac{\alpha\lambda t^{\alpha-1}}{(1+\lambda t^{\alpha})^2}$, $\alpha, \lambda > 0, t \geq 0$	$\frac{1}{1+\lambda t^{\alpha}}$	$\frac{\alpha\lambda t^{\alpha-1}}{1+\lambda t^{\alpha}}$	$\ln(1 + \lambda t^{\alpha})$	$\frac{\pi \csc(\pi/\alpha)}{\alpha\lambda^{(1/\alpha)}}$, if $\alpha > 1$	$\frac{2\alpha\pi \csc(\frac{2\pi}{\alpha}) - \pi^2 \csc^2(\frac{\pi}{\alpha})}{\alpha^2\lambda^{(2/\alpha)}}$, if $\alpha > 2$

$^{*}I_{\alpha}(t) = \int_0^t (1/\Gamma(\alpha)) x^{\alpha-1} e^{-x} dx$ is the incomplete Gamma function, $^{**}\Phi(\cdot)$ is the cumulative distribution function of standard Normal distribution.

2.5 Likelihood function

In order to estimate the parameters involved in the survival analysis models, the likelihood function is usually used. Since the observed data takes the form,$(t_1,\delta_1),...,(t_n,\delta_n)$, where (t_i,δ_i) are respectively the survival time and the censoring indicator for the i^{th} individual, the likelihood function $L(\theta;t_i)$ is given by

$$L(\theta;t_i)=\prod_{i=1}^{n}[f(t_i)]^{\delta_i}[S(t_i)]^{1-\delta_i}.$$

Using the relation in (2.1.2) the likelihood function becomes

$$L(\theta;t_i)=\prod_{i=1}^{n}[h(t_i)]^{\delta_i}S(t_i). \tag{2.5.1}$$

The maximum likelihood estimator of θ is the value in the parameter space that maximises the likelihood function or equivalently the log-likelihood function which is given by

$$\ell(\theta;t_i)=\sum_{i=1}^{n}\delta_i\log[h(t_i)]+\log[S(t_i)]. \tag{2.5.2}$$

Example 2.1 Assume the survival times follow the Weibull distribution $T\sim Weib(\alpha,\lambda)$, then from Table 2.1, the probability density function, the hazard, and the survival function are given by

$$f(t)=\alpha\lambda t^{\alpha-1}exp(\lambda t^{\alpha}), \qquad h(t)=\alpha\lambda t^{\alpha-1}, \qquad S(t)=exp(\lambda t^{\alpha}).$$

Applying formula (2.5.2), the log-likelihood is

$$\ell(\lambda, \alpha; t_i) = \sum_{i=1}^{n} \delta_i \log[\alpha\lambda t_i^{\alpha-1}] + \lambda t_i^{\alpha},$$

and consequently,

$$\ell(\lambda, \alpha; t_i) = r \log(\alpha\lambda) + (\alpha - 1) \sum_{i=1}^{n} \delta_i \log t_i - \lambda \sum_{i=1}^{n} t_i^{\alpha},$$

where $r = \sum_{i=1}^{n} \delta_i$ is the number of uncensored individuals. Differentiating the log-likelihood function and setting it to zero, the maximum likelihood estimators of λ and α are given by

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^{n} t_i^{\hat{\alpha}}}, \qquad \text{and} \qquad \frac{r}{\hat{\alpha}} + \sum_{i=1}^{n} \delta_i \log t_i - \hat{\lambda} \sum_{i=1}^{n} t_i^{\hat{\alpha}} \log t_i = 0.$$

The values of $\hat{\lambda}$ and $\hat{\alpha}$ can be found by an iterative numerical procedure such as the Newton-Raphson algorithm.

2.6 Proportional hazard models

Many studies focus on determining the risk factors affecting the survival times of individuals or subjects. Cox (1972) introduced the proportional hazard model (PH) in order to estimate the effects of such risk factors or covariates influencing survival time data. It assumes that the covariates have a multiplicative effect on the hazard. The proportional hazard model is the most popular model for survival data and has been used extensively in the literature. The proportional hazard model is given by

$$h(t_i, \boldsymbol{x_i}) = h_0(t_i) \exp(\boldsymbol{x_i'}\boldsymbol{\beta}). \tag{2.6.1}$$

where $h_0(t)$ is the baseline hazard function corresponding to the hazard function of a subject with covariate variables $\boldsymbol{x_i}$ equal to 0, and $\boldsymbol{\beta}$ is a $(p \times 1)$ vector of unknown parameters.

If the proportionality assumption is not satisfied, an alternative way to include the effect of covariates is using additive models. The hazard function then takes the form

$$h(t_i, \boldsymbol{x_i}) = h_0(t_i) + \boldsymbol{x_i'}\boldsymbol{\beta}. \tag{2.6.2}$$

For more details, see (Lin and Ying, 1994) and (Beamonte and Bermdez, 2003). A more general model that includes both types is called an additive-multiplicative model (Lin and Ying, 1995).

Example 2.2 Assume the survival times follow the Weibull distribution $T \sim Weib(\alpha, \lambda)$. Then $h_0(t_i) = \alpha t_i^{\alpha-1}$, and under the assumption of multiplicative hazard model, and setting $\lambda = \exp(\boldsymbol{x_i'}\boldsymbol{\beta})$ and $r = \sum_{i=1}^{n} \delta_i$,

$$h(t_i, \boldsymbol{x_i}) = \alpha t_i^{\alpha-1} \exp(\boldsymbol{x_i'}\boldsymbol{\beta}), \qquad \text{and} \qquad S(t_i, \boldsymbol{x_i}) = \exp(-t_i^{\alpha} \exp(\boldsymbol{x_i'}\boldsymbol{\beta})).$$

Consequently, the log-likelihood is given by

$$\ell(\beta, \alpha; t_i) = r(\log(\alpha) + \boldsymbol{x_i'}\boldsymbol{\beta}) + (\alpha - 1) \sum_{i=1}^{n} \delta_i \log t_i - \sum_{i=1}^{n} t_i^{\alpha} \exp(\boldsymbol{x_i'}\boldsymbol{\beta}). \tag{2.6.3}$$

2.7 Accelerated failure time models

Another way to estimate the effect of covariates on the survival time is through modelling the survival time by accelerated failure time (AFT). The covariates are assumed to act directly on survival times. The proportional hazard model is given by

$$S(t_i, \boldsymbol{x_i}) = S_0(t_i \exp(\boldsymbol{x_i'}\boldsymbol{\beta})), \tag{2.7.1}$$

where $S_0(t)$ is the baseline survival function corresponding to the survival function of a subject with covariate variables $\boldsymbol{x_i}$ equal to 0. Another way to represent (2.7.1) is using the survival time,

$$T = T_0 \exp(\boldsymbol{x_i'}\boldsymbol{\beta}),$$

where T_0 has survival function $S_0(t)$. The AFT model assumes the effect of covariates is multiplicative with respect to survival time. The AFT models are similar to the usual linear regression model

$$log(T_i) = \boldsymbol{x_i'}\boldsymbol{\beta} + \sigma\boldsymbol{\varepsilon}_i,$$

where $\boldsymbol{\beta}$ is the unknown regression coefficient and $\boldsymbol{x_i'}$ is the vector of observed covariates. The random errors ε_i is assumed to be independently and identically distributed with the mean zero and standard deviation one. If there are no censored data, the model can be readily estimated by ordinary least squares. One can simply generate a new variable, $Y = log(T)$, and use the linear regression model with Y as the dependent variable. If the error ε_i is normally distributed, the OLS estimates will also be maximum likelihood estimates of the model parameters. Survival data usually have at least some censored observations, and these are difficult to handle with OLS. Alternatively, one can use Maximum Likelihood Estimation (MLE) method with different distribution assumption. For each distribution of ε_i, there is a corresponding distribution for T. For instance, if the survival times follow the Weibull distribution $T \sim Weib(\alpha, \lambda)$, then $y = \log(T)$ has the extreme-value distribution $y \sim ext(a, b)$, where $a = \log\lambda$ and $b = 1/\alpha$, therefore,

$$\ell(\beta, b; y_i) = -r\log(b) + \sum_{i=1}^{n} \delta_i \left(\frac{y_i - \boldsymbol{x_i'}\boldsymbol{\beta}}{b}\right) - \sum_{i=1}^{n} \exp\left(\frac{y_i - \boldsymbol{x_i'}\boldsymbol{\beta}}{b}\right).$$

The following table lists some distributions of ε_i and their corresponding distributions of T.

Distribution of ε_i	Distribution of T
Extreme value (2 parameters)	Weibull
Extreme value (1 parameter)	Exponential
Log-gamma	Gamma
Logistic	Log-logistic
Normal	Log-normal

Table 2.2: Some distributions of ε_i and their corresponding distributions of T in modelling AFT.

2.8 Breast cancer recurrence data

This section describes a set of survival data on the recurrence of breast cancer in the UK which will be used throughout the thesis to demonstrate the development of the proposed models. The data used in this study was collected and provided by Research division of Christie Hospital in Manchester U.K. and includes more than 2850 women who were referred to the Christie Hospital, U.K. during 1985 and 1995, by their GPs with diagnosis of breast cancer. The data also includes the subsequent monitoring of these women up to 2001. This is an observational data set, hence, no randomisation or clinical trial was involved. Recurrence in this study is defined as clinical recurrence of breast cancer (i.e. after remission). The data were checked for values that were out of range, incorrect sequence of events or dates, and logical inconsistencies such as discrepancy between date of death and date of follow-up. As the result of this check a few individuals were excluded from the data set. The event of interest in this study is the first recurrence time of breast cancer in patients after initial treatment (surgery). There are three types of recurrence: Type one, local recurrence which cancerous tumour cells remain in the original site and grow over time. Type two, regional

recurrence of breast cancer is more serious because it usually indicates that the cancer has spread past the breast, into the axillary (underarm) lymph nodes and beyond. Type three, metastasis, where secondary cancer cells metastasise spread to other parts of the body and cause tumour. In this study, the response variable is the time from initial treatment to either local, regional recurrence or metastasis. In addition to these three observed recurrence times, other situations considered where the recurrence time was not observed because the patient was either symptomless at the end date of the study (independent right censoring) or the patient dropped out for some reason before the end of the study. The patients in this study can be classified into six categories:

1. patients who were alive when last seen with no disease and no recurrence (right censoring)
2. patients who experienced a local recurrence (LR) as the first recurrence (T_1)
3. patients who experienced a regional recurrence (RR) as the first recurrence (T_2)
4. patients who experienced metastasis (MT) as the first recurrence (T_3)
5. patients who died from breast cancer (DB) (i.e. drop-out due to the breast cancer) before the first recurrence (T_4) was observed
6. patients who died from other causes (DO) (i.e. drop-outs due to other causes) before the first recurrence (T_5) was observed

More details are in Appendix A. It is generally assumed that the right censoring mechanism is independent of the recurrence time (Kalbfleisch and Prentice, 2002). However, this assumption may not apply to both types of drop-outs. For instance, patients diagnosed with an advanced stage of breast cancer may die due to that cancer before any clinical recurrence of it. Similarly, patients with severe sickness tend to have shorter survival time and are more likely to die from other diseases due to general weakness. Ignoring such informative drop-outs

while employing the commonly used estimation procedures based on treating drop-outs as independent right censored observations tends to underestimate the parameters of interest. Hence, in this study drop-outs were not treated as right censored observations. Some of the variables which have been observed for these patients to act as potential covariates are: *age, stage of the disease at first diagnosis, type of surgery, histology, the cohort of initial surgery, chemotherapy, menopausal status, radiotherapy* and *side of the body.* More details about the data are given in chapter three.

2.9 Summary

This chapter summarised the main features of the survival data and the methods available in estimating the survival and hazard functions. Both parametric and non-parametric estimates of the survival function are described. Estimating the empirical survival function by Product-limit estimator can be used to judge the best fit for the survival function. Matching the graph of empirical survival function with those in figure 2.1 to figure 2.4 can be used to decide the parametric survival function. One of the limitations of these models is that they implicitly assume homogeneity of study populations which may not be true. Adding covariates to the model may relax this assumption. There are two ways to estimate the effects of covariates or risk factors influencing survival time data, proportional hazard models and accelerated failure time models. The proportional hazard models assume that the covariates have a multiplicative effect on the hazard. Whereas in accelerated failure time models, the covariates are assumed to act directly on survival times. An extension of these models is by including random effects which will be discussed in the next chapter.

Chapter 3

Frailty Models in the literature

3.1 Introduction

In this chapter, first, a literature review of the frailty models in survival analysis and the distributions used in modelling frailty is given. Second, the models discussed are applied to a simulation data and to the breast cancer data presented in the previous chapter. Standard methods in survival analysis implicitly assume homogeneity of study populations. That means that all subjects have the same degree of failure risk and that the survival times are independently and identically distributed. Models with covariates relax this assumption by introducing observed sources of heterogeneity. But it is not realistic to assume that all relevant risk factors or covariates were measured and included in the model. Either the relevant risk factors are unknown or they are known but it may be costly to measure them. Unmeasured covariates or omitted risk factors generate a between-subject variation usually referred to as frailty (unobserved heterogeneity). Frailty models can be viewed as an extension of Cox proportional hazard, (Cox, 1972). Considering these omitted risks as random variables with a probability, a joint distribution of failure time and the frailty could be generated and since the frailty is unobservable (i.e. no data on that) it has to be integrated out. In the context of survival analysis, a frailty model is a *mixed-effects* model where the frailty is the

random effect component which usually has a multiplicative effect on its hazard function. Vaupel et al. (1979) introduced the term frailty as a measure of susceptibility to all causes of death to describe mortality in non-homogeneous populations and used it in univariate survival models. Clayton (1978) applied the idea of the frailty model to the multivariate situation of chronic disease incidence in families, but he did not use the term "frailty".

3.2 Linear mixed models

Linear mixed models (LMMs) are statistical models for continuous responses in which the residuals are assumed to be normally distributed but may not be independent or have constant variance. Therefore, they provide the flexibility for modelling not only the means of the data, but the variances and covariances as well. In a linear mixed-effects model, responses from a subject are assumed to be the sum of *fixed* and *random* effects. A factor is considered to be *fixed* if all levels or categories that are of interest are included in the study. A factor is considered to be *random* if the levels or categories included in the study represent a random sample from a larger population of values. The random effects contribute only to the covariance structure of the data. It often introduces correlations between cases. Such correlations are usually encountered in studies where data are grouped in clusters or in longitudinal and repeated measures studies with multiple observations for the same subject. If only *fixed* effects are included in the model, then the dependent variable Y is modelled in relation to several explanatory variables by

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad \text{and} \quad \boldsymbol{\varepsilon} \sim \boldsymbol{N(0, \Sigma)}, \tag{3.2.1}$$

where $\mathbf{Y}$ is $(n \times 1)$ vector, $\boldsymbol{\varepsilon}$ is $(n \times 1)$ vector of residuals with variance-covariance matrix $\boldsymbol{\Sigma}$, $\boldsymbol{\beta}$ is $(p \times 1)$ vector of unknown parameters, and $\mathbf{X}$ is $(n \times p)$ design matrix, the matrix of explanatory variables. If the intercept is included in the model, then a vector of ones should be included in the design matrix. If the model contains only continuous explanatory variables, it is usually called a regression model while models containing only qualitative variables are called Analysis of Variance models (ANOVA). Both of these models are special cases of the general linear model, where both types of explanatory variables could be included in the model. In general linear models, the response variables are assumed to be independent and normally distributed with common variance, and link function $\mu_i = E(Y_i) = \boldsymbol{X_i\beta}$. In longitudinal and cluster data, it is more appropriate to include both *fixed* and *random* effects in the model, which is extension of the model given in (3.2.1). The the random effect is included as follows

$$\begin{aligned} &\boldsymbol{Y_i = X_i\beta + Z_ib_i + \varepsilon_i,} \\ &\boldsymbol{b_i \sim N(0, D)} \text{ and } \boldsymbol{\varepsilon_i \sim N(0, \Sigma_i)}. \end{aligned} \tag{3.2.2}$$

where $\boldsymbol{Y_i}$ is $(n_i \times 1)$ vector of responses of individual or cluster $i, (i = 1, ..., N)$, $\boldsymbol{\varepsilon}_i$ is $(n_i \times 1)$ vector of residuals with variance-covariance matrix $\boldsymbol{\Sigma}_i$, $\boldsymbol{\beta}$ is $(p \times 1)$ vector of *fixed* effect, $\boldsymbol{b_i}$ is $(q \times 1)$ vector of *random* effect, $\boldsymbol{X_i}$ and $\boldsymbol{Z_i}$ are $(n_i \times p)$ and $(n_i \times q)$ matrices of known covariates and $\boldsymbol{D}$ is $(q \times q)$ covariance matrix of the random effect. Maximum likelihood (ML) and restricted maximum likelihood (REML) estimation are the methods commonly used to estimate the model parameters. For more information about linear mixed models see Verbeke and Molenberghs (2000), McCulloch and Searle (2001) and Muller and Stewart (2006). For statistical software fitting Linear mixed models see West et al. (2007).

3.3 Model Identifiability

A statistical model should be identifiable to make a valid inference about its parameters. A model is considered to be identifiable if its parameter values uniquely determine the probability distribution of the data and the probability distribution of the data determines the parameter values uniquely. The identifiability as defined in Casella and Berger (2002)

3.3.1 Definition. A parameter θ for a family of distributions $\{f(x|\theta) : \theta \in \Theta\}$ is *identifiable* if distinct values of θ correspond to distinct p.d.fs. That is, if $\theta \neq \theta'$, then $f(x|\theta)$ is not the same function of x as $f(x|\theta')$.

One of the most common source of model non-identifiability is a poorly defined model. Over-parameterisation of the model usually creates such a problem.

3.4 Univariate frailty models

In the frailty framework, when there is only one time-point measure and no clustering of individuals, univariate frailty models are used to take into account the heterogeneity between individuals due to the exclusion of important covariates in the model. In LMMs, random effects are usually included when there is clustering or repeated measures while in survival analysis random effects (frailty) are included to account for unobserved heterogeneity between subjects. Vaupel et al. (1979) proposed a univariate frailty model to survival analysis assuming a Gamma distribution to account for unobserved heterogeneity, i.e, assuming that different subjects have different frailties so that subjects which are more frail tend to have shorter survival time than those which are less frail. Many authors have discussed the univariate frailty models. (See for example, Lancaster and Nickell (1980), Heckman and Singer (1984, 1985), Vaupel and Yashin (1985), Hougaard (1984, 1986a,b, 2000), Vaupel

(1990), Aalen (1988, 1992) and Richardson and Green (1997)).

There are different ways to include the *random effect* (frailty) in survival analysis. Under the assumption of proportional hazard, the multiplicative frailty effects model which is commonly used in the literature, the frailty acts multiplicatively on the underlying baseline hazard function. In this case the conditional hazard function on the random effect z takes the form

$$h(t_i, \boldsymbol{x_i}|z) = zh(t_i, \boldsymbol{x_i}) = zh_0(t_i)exp(\boldsymbol{x_i'}\boldsymbol{\beta}). \tag{3.4.1}$$

where $h_0(t_i)$ is the baseline hazard, $\boldsymbol{x_i}$ is the vector of covariates of the i^{th} subject, and $\boldsymbol{\beta}$ is the fixed effect vector. In (3.4.1), Z is assumed to have some density $g(z, \theta)$ with parameter vector θ, $E[Z] = 1$ and $V[Z] = \tau^2$. Any other value for this expectation could be used since it would be absorbed into the baseline hazard function. The conditional survival function is given by

$$\begin{aligned} S(t_i, \boldsymbol{x_i}|z) &= \exp\left(-\int_0^t h(s, \boldsymbol{x_i}|z)ds\right) \\ &= \exp\left(-z\int_0^t h(s, \boldsymbol{x_i})ds\right) \\ &= \exp\left(-zH_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}\right). \end{aligned} \tag{3.4.2}$$

The unconditional survival function is given

$$S(t_i, \boldsymbol{x_i}) = \int_0^\infty S(t_i, \boldsymbol{x_i}|z)g(z)dz = \int_0^\infty \exp\left(-zH_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}\right) g(z)dz,$$

and, hence,

$$S(t_i, \boldsymbol{x_i}) = \mathcal{L}_Z[H_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}]. \tag{3.4.3}$$

where $\mathcal{L}_Z[.]$ is the Laplace transformation with respect to the random variable Z and $H_0(t_i)$

is the cumulative baseline hazard function. This model is identifiable when the the expected value of Z is finite (Elbers and Ridder, 1982). An alternative way to write the above model is by setting $(u = \log z)$ in (3.4.1), the conditional hazard can be written as

$$h(t_i, \boldsymbol{x_i}|u) = h_0(t_i)exp(\boldsymbol{x_i'}\boldsymbol{\beta} + u). \tag{3.4.4}$$

This model is a special case of the LMMs (3.2.2) (intercept model) where the design matrix of the random effect contains only a column vector of ones. Frailty models differ from LMMs in several ways: First, they do not include residual components $\boldsymbol{\varepsilon}$. In the frailty models the residual variability is modelled through the survival distribution. Secondly, the expected survival time, given the random effects, is not equal to the linear combination of covariates as LMMs. Thirdly, the inferential methods of frailty models have been less developed than LMMs due to the incompleteness of data due to censoring and truncation especially in multivariate models with non-Gaussian frailty distribution. This thesis focuses on multiplicative frailty models. However, other types of frailty models exist such as additive frailty models, where the frailty acts additively on the baseline hazard function. The hazard function as defined by Cai and Zeng (2011)takes the form

$$h(t_i, \boldsymbol{x_i}|z) = h_0(t_i) + \boldsymbol{x_i'}\boldsymbol{\beta} + z.$$

For more details, see Lin and Ying (1994), Korsgaard and Andersen (1998), Peterson (1998), Li (2002), Zhong and Li (2004), Pipper and Martinussen (2004), Yin and Ibrahim (2005), Yin (2007) and Cai and Zeng (2011). Another way to include the frailty effect in the survival analysis is through accelerated failure time (AFT) models. Many authors considered AFT frailty models namely, Anderson and Louis (1995); Keiding et al. (1997), Klein et al. (1999), Pan (2001), Lambert and Collett (2002), Lambert et al. (2004), Chang (2004), Zhang and

Peng (2007) and Xu and Zhang (2009, 2010). In general, any distribution with positive range, mean one and finite variance is a suitable candidate to represent the frailty distribution. Gamma and Inverse Gaussian distributions are the mostly used distributions in the literature since they provide a closed form expression for the unconditional survival function.

3.4.1 Gamma frailty model

The Gamma distribution is a member of the exponential family and from a computational and analytical point of view; it is convenient as a frailty distribution and it is easy to derive the closed form expressions of survival and the hazard function. This is due to the simplicity of the Laplace transform. Therefore, most published work on frailty analysis assumes the Gamma distribution because it is mathematically attractive. This includes both the frequentist approach as well as the Bayesian approach. (See Clayton (1978), Clayton and Cuzick (1985), Vaupel et al. (1979), Oakes (1982), Crowder (1985), Scallan (1987), Yashin et al. (1995), dos Santos et al. (1995), Congdon (1995), Shih and Louis (1995), Sahu et al. (1997), Hougaard (1995, 2000), Yin and Ibrahim (2005), Perperoglou et al. (2006), Balakrishnan and Peng (2006), Duchateau and Janssen (2008), Peng and Zhang (2008), Jonker et al. (2009), Xu and Zhang (2010) and Molenberghs and Verbeke (2011)). For a comparison between the Bayesian approach and the frequentist approach see David et al. (2007).

Weibull hazard with Gamma frailty

Assume the survival times follow the Weibull distribution $T \sim Weib(\alpha, \lambda)$, and the frailty follows a Gamma distribution with unit mean and variance τ^2, $Z \sim \Gamma(1/\tau^2, \tau^2)$. (Without loss of generality, any other value for the expectation could be absorbed into the baseline hazard function). For the Weibull distribution, the baseline hazard is $h_0(t) = \alpha\lambda t^{\alpha-1}$. The effect of covariates is modelled through the scale parameter of the Weibull distribution,

$\lambda = \exp(\boldsymbol{x_i'}\boldsymbol{\beta})$. According to (3.4.2) the conditional survival function is given by

$$S(t_i, \boldsymbol{x_i}|z) = \exp\left(-zH_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}\right) = \exp\left(-zt_i^{\alpha}e^{\boldsymbol{x_i'}\boldsymbol{\beta}}\right).$$

The unconditional survival and hazard functions are given by

$$S(t_i, \boldsymbol{x_i}) = [1 + \tau^2 H_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}]^{-(1/\tau^2)} = [1 + \tau^2 t_i^{\alpha}e^{\boldsymbol{x_i'}\boldsymbol{\beta}}]^{-(1/\tau^2)}.$$

$$h(t_i, \boldsymbol{x_i}) = \frac{h_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}}{1 + \tau^2 H_0(t_i)e^{\boldsymbol{x_i'}\boldsymbol{\beta}}} = \frac{\alpha t_i^{\alpha-1}e^{\boldsymbol{x_i'}\boldsymbol{\beta}}}{1 + \tau^2 t_i^{\alpha}e^{\boldsymbol{x_i'}\boldsymbol{\beta}}}. \tag{3.4.5}$$

3.4.2 Inverse Gaussian frailty model

The inverse Gaussian distribution is named so because it satisfies the inverse relationship with the Gaussian distribution. There are many similarities between the statistics derived from this distribution and those of the Normal distribution. It is a member of the exponential family and like the Gamma distribution it is mathematically attractive. It was presented as an alternative to the Gamma distribution by Hougaard (1984) since it makes the population homogeneous with time, whereas for Gamma the relative heterogeneity is constant. It is not popular like Gamma frailty especially in multivariate frailty framework since the summation of Inverse Gaussian usually is not an Inverse Gaussian (reproductivity property). However, many authors have considered it. (See, Manton et al. (1986), Whitmore and Lee (1991), Klein et al. (1992a), Lam and Kuk (1997), Keiding et al. (1997), Price and Manatunga (2001), Economou and Caroni (2005), Jeong and Oakes (2005), Kheiri et al. (2007), Duchateau and Janssen (2008) and Chen and Lio (2008)). The probability density function of the Inverse

Gaussian distribution $T \sim IG(\alpha, \lambda)$ with location parameter μ and scale parameter λ is

$$f(t) = \sqrt{\frac{\lambda}{2\pi}} t^{-3/2} \exp\left\{-\frac{\lambda}{2\mu^2 t}(t-\mu)^2\right\}$$

The mean and the variance are

$$E[T] = \mu, \qquad V[T] = \frac{\mu^3}{\lambda}$$

Figure 3.1 describes the density of the Inverse Gaussain distribution with a scale parameter which equals one with different values of the shape parameter.

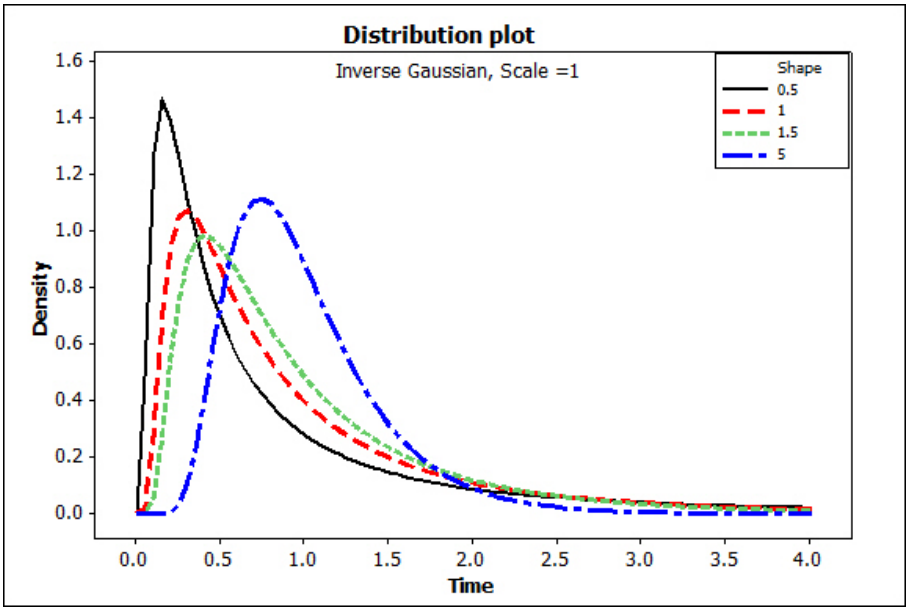


Figure 3.1: Inverse Gaussian densities with scale parameter $\lambda = 1$ and shape parameter $\alpha = (0.5, 1, 1.5, 5)$

Weibull hazard with Inverse Gaussian frailty

Assume the survival times follow the Weibull distribution, $T \sim Weib(\alpha, \lambda)$, and the frailty model is an Inverse Gaussian distribution with unit mean and variance τ^2, $Z \sim IG(1, 1/\tau^2)$.

The unconditional survival and hazard functions are given by

$$\begin{aligned} S(t_i, \boldsymbol{x_i}) &= \exp\left(\frac{1}{\tau^2}(1 - \sqrt{1 + 2\tau^2 H_0(t_i) e^{\boldsymbol{x_i'\beta}}}\right) \\ &= \exp\left(\frac{1}{\tau^2}(1 - \sqrt{1 + 2\tau^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}}}\right). \end{aligned}$$

$$h(t_i, \boldsymbol{x_i}) = \frac{h_0(t_i) e^{\boldsymbol{x_i'\beta}}}{\left(1 + 2\tau^2 H_0(t_i) e^{\boldsymbol{x_i'\beta}}\right)^{1/2}} = \frac{\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'\beta}}}{\left(1 + 2\tau^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}}\right)^{1/2}}. \tag{3.4.6}$$

3.4.3 Log-Normal frailty models

Because of its relation to the Normal distribution, the Log-Normal distribution is frequently used for frailty in the literature. Assuming a Log-Normal distribution is equivalent to assuming Normal distribution for the additive frailty model incorporated in the exponent of the hazard function of the Cox model. For models (3.4.1), the frailty distribution is assumed to follow the Log-Normal distribution, whilst for models (3.4.4), the frailty distribution is Normal. One of the difficulties of the Log-Normal frailty distribution is that the Laplace transform does not have a simple form and hence no explicit form of the unconditional likelihood exists. The Log-Normal distribution was mainly developed by McGilchrist and Aisbett (1991). Many authors considered the Log-Normal frailty models in multivariate frailty models. (See, McGilchrist (1993), Lillard (1993), Lillard et al. (1995), Xue and Brookmeyer (1996), Sastry (1997), Gustafson (1997), Vaida and Xu (2000), Ripatti and Palmgren (2000), Ripatti et al. (2002), Huang and Wolfe (2002) Stefanescu and Turnbull (2006) and Duchateau and Janssen (2008)).

3.4.4 Weibull hazard with Log-Normal frailty

Assume the survival times follow the Weibull distribution and the frailty has a Log-Normal random variable Z with mean μ and variance τ^2, $Z \sim LogN(\mu, \tau^2)$. In Log-Normal frailty, the inclusion of the frailty in the model is usually done by using $W = LN(Z)$ which has Normal distribution, $W \sim N(\mu^*, \sigma^2)$. In this case, the mean and the variance of the frailty are related to those of the normal distribution through the following relations:

$$\begin{aligned} \mu &= E[Z] = e^{\mu^* + \sigma^2/2}, \\ \tau^2 &= V[Z] = e^{2\mu^* + \sigma^2}(e^{\sigma^2} - 1). \end{aligned} \tag{3.4.7}$$

There are two forms of Log-Normal frailty in the literature. Depending on the restriction on the frailty expected value, either the mean of frailty is one, i.e., $\boldsymbol{E}[Z]{=}\mu = 1$ or the mean of the log of frailty is zero, i.e., $E[W] = E[LN(Z)] = \mu^* = 0$. These restrictions are set to assure model identifiability. If the effect of covariates is modelled through the scale parameter of the Weibull distribution $\lambda = \exp(\boldsymbol{x_i'\beta} + w)$, then the conditional survival function is given by

$$S(t_i, \boldsymbol{x_i}|z) = \exp\left(-zH_0(t_i)e^{\boldsymbol{x_i'\beta}}\right) = \exp\left(-t^{\alpha}e^{\boldsymbol{x_i'\beta}+w}\right).$$

Unfortunately, the unconditional survival and hazard functions do not have a closed form and numerical integration is needed to integrate out the frailty variable. The contribution of the i^{th} individual to the conditional likelihood is given by

$$L_i(t_i, \delta_i, \boldsymbol{x_i}|z) = (zh_0(t_i)e^{\boldsymbol{x_i'\beta}})^{\delta_i} e^{-zH_0(t_i)e^{\boldsymbol{x_i'\beta}}}.$$

where t_i is the survival time or the censoring time of the i^{th} individual, δ_i is the censoring indicator, Z is the unobserved random effects (frailties), $\boldsymbol{x_i}$ is the vector of the observed

covariate, and $h_0(t_i)$ is the baseline hazard. Assuming the conditional independence of the survival times given the frailty, the unconditional (marginal) likelihood function is

$$L_i(t_i, \delta_i, \boldsymbol{x_i}) = \int_{R^+} (z h_0(t_i) e^{\boldsymbol{x_i'\beta}})^{\delta_i} e^{-z H_0(t_i) e^{\boldsymbol{x_i'\beta}}} f(z, \tau) dz.$$

where $f(z, \tau)$ is the p.d.f of the frailty distribution. In the case of Log-Normal frailty, the marginal likelihood of the i^{th} individual is given by

$$L_i(t_i, \delta_i, \boldsymbol{x_i}) = \int_R (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'\beta}+w})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'\beta}+w}) \frac{1}{\tau\sqrt{2\pi}} e^{-\frac{w^2}{2\tau^2}} dw.$$

A numerical integration such as Gauss quadrature integration could be used to integrate out the frailty

$$\int_{-\infty}^{\infty} f(x) dx \approx \sum_{k=1}^{K} \pi_k e^{x_k^2} f(x_k).$$

where x_k and π_k are the zeros of Hermite polynomials and their corresponding weight factors respectively. To make this integration simpler and less time consuming from a computational point of view one can set $Y = \frac{W}{2\tau}$, the simplified likelihood is given by

$$\begin{aligned} L_i(t_i, \delta_i, \boldsymbol{x_i}) &= \int_R (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'\beta}+\tau y\sqrt{2}})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'\beta}+\tau y\sqrt{2}}) \frac{1}{\sqrt{\pi}} e^{-y^2} dy \\ &\approx \sum_{k=1}^{K} \pi_k^* (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'\beta}+\tau y_k^*})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'\beta}+\tau y_k^*}), \end{aligned} \tag{3.4.8}$$

where $y_k^* = y_k\sqrt{2}$ and $\pi_k^* = \pi_k/\sqrt{\pi}$. Either the likelihood is maximise directly using an iterative method, say Newton-Raphson or using the EM-algorithm by considering (3.4.8) as a finite mixture. To use the EM-algorithm the vector of survival time $\mathbf{T}$ is assumed to be observed part whilst the vector $\boldsymbol{\zeta} = (\boldsymbol{\zeta_1}, ..., \boldsymbol{\zeta_n})$ to be unobservable random variables, where

$\boldsymbol{\zeta}_i = (\zeta_{i1}, ..., \zeta_{iK})$ such that ζ_{ik} is unity if t_i comes from component k and 0 otherwise. So, given all of the data $\mathbf{Y} = (\mathbf{T}, \boldsymbol{\zeta})$ and the set of the parameter of interest $\boldsymbol{\phi} = (\boldsymbol{\beta}, \tau, \alpha)$, the complete likelihood of is

$$L(\boldsymbol{Y}, \boldsymbol{\phi}) = \prod_{i=1}^{n} \prod_{k=1}^{K} [\pi_k f_{ik}(t_i, \boldsymbol{x_i})]^{\zeta_{ik}},$$

and the complete log-likelihood is

$$\ell(\boldsymbol{Y}, \boldsymbol{\phi}) = \sum_{i=1}^{n} \sum_{k=1}^{K} \zeta_{ik} \log \pi_k + \sum_{i=1}^{n} \sum_{k}^{K} \zeta_{ik} \log f_k(y, \boldsymbol{x_i}). \tag{3.4.9}$$

where $f_{ik}(t_i, \boldsymbol{x_i}) = (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\tau y_k^*})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\tau y_k^*})$. The EM-algorithm starts by estimating the missing quantities in E-step and then maximisation in M-step.

E-Step. Suppose that $\boldsymbol{\phi} = (\boldsymbol{\beta}, \tau, \alpha)$ are known. Then the missing quantities $\boldsymbol{\zeta}$ are replaced by their conditional expectations, conditioned on the parameters and on the observed data $\mathbf{T}$. The conditional expectation of the k^{th} component of $\boldsymbol{\zeta}_i$ is just the conditional probability that the observation t_i comes from the k^{th} component of the mixture, conditioned on the parameters and the observed data. Let the conditional expectation of the k^{th} component of $\boldsymbol{\zeta}$ be $\tilde{\zeta}_{ik}$. Then

$$\tilde{\zeta}_{ik} = \frac{w_k f_{ik}(t_i, \boldsymbol{x_i})}{\sum_{k=1}^{m} w_k f_{ik}(t_i, \boldsymbol{x_i})}.$$

M-Step. Suppose that the missing $\boldsymbol{\zeta}_i$ are now known. The estimates of the parameters $\boldsymbol{\phi} = (\boldsymbol{\beta}, \tau, \alpha)$ can then be obtained by maximising the log-likelihood function ℓ in (3.4.9). This procedure works fine if the score equations have closed form, but the problem here is that the score equations cannot be solved analytically and ℓ needs to be maximised iteratively. Similar procedures can be found in LMMs assuming a Normal random effect

and the conditional distribution belongs to the exponential family Bock and Aitkin (1981), Hsu (2000), McLachlan and Krishnan (2008) and Aitkin et al. (2009).

3.4.5 Non-parametric frailty models

Most of the recently published work about non-parametric frailty is from the Bayesian prospective. The unknown frailty distribution is modelled non-parametrically using a Dirichlet process (see, Manda 2011, Cai 2010, Naskar et al. 2005 and Pennell and Dunson (2006)). Alternatively, a semi-parametric survival frailty model can be obtained by assuming a non-parametric baseline hazard (see, Clayton (1988), Klein (1992b), Li and Lin (2000) and Vaida and Xu (2000)). Naskar (2008) introduced a non-parametric Dirichlet process for the distribution of frailty along with the assumption of a non-parametric baseline hazard function. This thesis focuses on the frequentist approach since no prior distributions of the model parameters is assumed. In the next section, simulations will show how the model inferences are not robust against mis-specifying of the frailty distribution. This is also supported by the literature. Heckman and Singer (1982a) induced interest in non-parametric representation of frailty (also see, Laird 1978, Heckman and Singer 1984, Davies and Crouchley 1984). The theoretical result of the non-parametric characterisation of the frailty distribution within maximum likelihood estimation is narrowed to a finite number of mass points. Also from the empirical experience it has been shown that the number of mass points tends to be small (Davies, 1993). For univariate frailty models, see (dos Santos et al. 1995, Aitkin 1999 and Aitkin et al. 2009). The Weibull hazard with non-parametric frailty model can be written as

$$L_i(t_i, \delta_i, \boldsymbol{x_i}) \approx \sum_{k=1}^{K} \pi_k (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\gamma_k})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\gamma_k}). \tag{3.4.10}$$

where $0 < \pi_k < 1$ and $\sum_{k=1}^{K} \pi_k = 1$. This model is similar to the Weibull survival time with Log-Normal frailty model given in (3.4.8) except that K and (γ, π) are no longer given

and have to be estimated like other parameters. The mean and variance of the frailty have been absorbed into the non-parametric representation of the model. For the EM-algorithm, same argument as for (3.4.8) could be used to estimate the mixture in (3.4.10) as before, except that the zeros of Hermite polynomials and their corresponding weights are considered as parameter and need to be estimated.

3.5 Univariate simulations

Simulation studies are used to assess the performance of a proposed model, especially if there is a lack of theoretical background. In simulations, the researchers know the true parameters values and use them to generate the data. This data is used to fit the parameters using the proposed model; if the original parameters values are retrieved then the proposed model is acknowledged. However, real data sets are used to show the applicability of the proposed model, but it does not evaluate its performance. Recurrence of breast cancer data is used as an application of the proposed models in this thesis. It is an observational study where recurrence is defined as clinical recurrence of breast cancer after remission where the event of interest is the first recurrence of breast cancer. There are five possible outcomes, local recurrence, regional recurrence, metastasis, died from breast cancer or died from other causes. This section test the performance of the models mentioned above using simulated data (using a self-written GAUSS code), while the next section demonstrates the applications of these models to the breast cancer data. The model is fitted by maximum likelihood estimation method based on numerical integration not the EM-algorithm. The simulated data have been generated from a univariate frailty models assuming Weibull baseline hazard and different frailty distributions including Gamma, Inverse Gaussian and Log-Normal distribution. The simulations are divided into three parts: First, simulation of Log-Normal frailty model with Weibull baseline hazard. Second, simulation of Log-Normal, Gamma and Inverse Gaussian

frailty model with Weibull baseline hazard fitted by Log-Normal frailty using model (3.4.8). Third, simulation of Log-Normal, Gamma, Inverse Gaussian and arbitrary (non-parametric) frailty model with Weibull baseline hazard fitted non-parametrically using model (3.4.10). First, a simulated data of failure times that follow a Weibull distributions $T \sim Weib(\alpha, \lambda)$ and censoring times follow a Weibull distribution $C \sim Weib(\theta, \alpha)$ assuming random censoring. The distribution of the log of frailty $W = Log(Z)$ is assumed to be Normal with mean zero and variance σ^2, $W \sim N(0, \sigma^2)$. Two different sets of parameters were used to check the model estimation. Different types of explanatory variables were generated, X_{i1} is a continuous random variable from Uniform distribution $X_{i1} \sim Uni(0, 1)$, X_{i2} is a dichotomous, and X_{i3} is a qualitative variable with three categories which was converted to two dummy variable $X_{i3,1}$ and $X_{i3,2}$.

$$X_{i2} = \begin{cases} 1 & if \quad u_{i1} < 0.3 \\ 0 & if \quad u_{i1} \geq 0.3 \end{cases} \qquad \text{and} \qquad X_{i3} = \begin{cases} 1 & if \quad u_{i2} < 0.4 \\ 2 & if \quad 0.4 \leq u_{i2} < 0.6 \\ 3 & if \quad u_{i2} > 0.6 \end{cases}$$

where $U_{ij} \sim Uni(0, 1)$, j=1,2. The linear predictors are generated as

$$\boldsymbol{x_i'\beta} = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{i3,1}\beta_3 + x_{i3,2}\beta_4.$$

The failure times were generated as $T_i = (-\log(u_{i3})/\lambda_i)^{1/\alpha}$, where $\lambda_i = \exp(\boldsymbol{x_i'\beta} + W)$ and W is the log of frailty with $W \sim N(0, \sigma^2)$ and $u_{i3} \sim Uni(0, 1)$. The censoring times were generated as $C_i = (-\log(u_{i4})/\theta)^{1/\alpha}$, where $u_{i4} \sim Uni(0, 1)$ and finally the survival times are $Y_i = \min(T_i, C_i)$. Two sets of parameters are used to generate the failure and censoring times with censoring rate of 20%. The above models are estimated using the maximisation procedure in Gauss program 'MAXLIK'. A Gaussian quadrature integration with 32 quadrature points was used to integrate out the frailty, the code is in appendix C. For each set of parameters 600 data sets each with sample sizes of 500 and 5000 are simulated.

The results are shown in Table 3.1. Obviously, the estimation method and the GAUSS code managed to retrieve the true parameters values especially with the large sample size. The standard errors are much smaller in case of sample size 5000 than 500.

Parameter	True values	Sample size		True values	Sample size	
		500	5000		500	5000
		Mean (S.e)	Mean (S.e)		Mean (S.e)	Mean (S.e)
α	1	1.066	1.003	0.5	0.550	0.503
		(0.310)	(0.038)		(0.157)	(0.030)
σ^2	1	1.097	1.004	1.5	1.689	1.508
		(0.596)	(0.084)		(0.713)	(0.147)
β_0	-4	-4.212	-4.011	3	3.354	3.019
		(1.024)	(0.148)		(1.150)	(0.215)
β_1	1	1.052	1.007	0.8	0.872	0.808
		(0.382)	(0.090)		(0.415)	(0.111)
β_2	-2	-2.144	-2.007	-1	-1.104	-1.005
		(0.699)	(0.092)		(0.406)	(0.090)
β_3	4	4.259	4.013	0.7	0.762	0.703
		(1.233)	(0.160)		(0.317)	(0.074)
β_4	2	2.128	2.009	-2	-2.209	-2.011
		(0.689)	(0.102)		(0.720)	(0.145)

Table 3.1: Simulation data of Weibull baseline hazard generated with Log-Normal frailty and fitted by Log-normal, 600 data sets each with sample sizes of 500 and 5000.

Second, to test the robustness of the parameters estimate against mis-specifying the frailty distribution, different frailty distribution were generated such as Gamma, Inverse-Gaussian, and Log-Normal distribution and fitted by Log-Normal distribution. Table 3.2 shows these simulations. Clearly, the results are not robust against the mis-specifying of frailty distribution. There is a big difference between Gamma frailty and Log-Normal especially in estimating the frailty variance τ^2. However, there is a big similarity between Log-Normal and Inverse Gaussian frailties. More simulations that support these conclusions are given in chapter five. The standard errors are smaller in case of Log-Normal and Inverse Gaussian frailty than the Gamma frailty. This is also another indication of the similarity between

Log-Normal and Inverse Gaussian and different from the Gamma distribution.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian	
		500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
α	1	1.066	1.003	1.596	1.282	1.030	0.979
		(0.310)	(0.038)	(0.633)	(0.102)	(0.261)	(0.031)
τ^2	1	1.097	1.004	2.427	1.847	0.855	0.786
		(0.596)	(0.084)	(1.220)	(0.208)	(0.510)	(0.072)
β_0	-4	-4.212	-4.011	-6.931	-5.680	-4.470	-4.285
		(1.024)	(0.148)	(2.561)	(0.418)	(0.983)	(0.133)
β_1	1	1.052	1.007	1.537	1.288	1.011	0.978
		(0.382)	(0.090)	(0.819)	(0.157)	(0.308)	(0.080)
β_2	-2	-2.144	-2.007	-3.212	-2.564	-2.065	-1.955
		(0.699)	(0.092)	(1.352)	(0.218)	(0.561)	(0.082)
β_3	4	4.259	4.013	6.363	5.113	4.121	3.918
		(1.233)	(0.160)	(2.550)	(0.409)	(1.089)	(0.132)
β_4	2	2.128	2.009	3.205	2.568	2.060	1.959
		(0.689)	(0.102)	(1.369)	(0.219)	(0.537)	(0.086)

Table 3.2: Log-Normal, Gamma and Inverse Gaussian frailty model with Weibull baseline hazard and four covariates simulated data fitted by Log-Normal frailty, 600 data sets each with sample sizes of 500 and 5000.

Third, to check the ability of non-parametric frailty model in capturing the parameters estimates, a simulated data with different frailty distribution is generated and fitted by non-parametric frailty. Four different frailty distribution are generated, Log-Normal, Gamma, Inverse Gaussian and arbitrary (a discrete random variable with expect value equals one). Table 3.3 presents the parameters' estimates using five mass points of the non-parametric frailty. These results are very close when the model is fitted with four mass points. Appendix A gives the parameters estimates using one, two, three, four or five mass points along with their log-likelihood. Obviously, the non-parametric frailty is capable of capturing the parameters' estimates regardless of the original distribution of frailty. The mass points and their corresponding weights are represented by γ_i and π_i, $(i = 1, \ldots, 5)$ respectively. The standard errors are smaller in case of sample size of 5000.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian		Arbitrary	
		Sample size		Sample size		Sample size		Sample size	
		500	5000	500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_1	1	1.041	0.982	1.018	0.991	1.051	0.987	1.048	0.981
		(0.330)	(0.097)	(0.333)	(0.105)	(0.347)	(0.095)	(0.372)	(0.115)
β_2	-2	-2.104	-1.963	-2.072	-1.990	-2.106	-1.973	-2.092	-1.962
		(0.388)	(0.140)	(0.360)	(0.121)	(0.421)	(0.144)	(0.429)	(0.164)
β_3	4	4.181	3.924	4.127	3.96 8	4.220	3.950	4.195	3.923
		(0.705)	(0.266)	(0.596)	(0.230)	(0.741)	(0.266)	(0.789)	(0.318)
β_4	2	2.079	1.963	2.065	1.983	2.103	1.975	2.093	1.961
		(0.386)	(0.144)	(0.378)	(0.128)	(0.428)	(0.145)	(0.463)	(0.173)
α	1	1.046	0.981	1.034	0.992	1.055	0.988	1.048	0.982
		(0.171)	(0.064)	(0.144)	(0.054)	(0.180)	(0.064)	(0.186)	(0.077)
γ_1		-1.351	-1.381	-1.329	-0.850	-1.344	-1.332	-1.284	-1.214
		(1.202)	(1.267)	(2.044)	(6.282)	(1.137)	(1.141)	(1.333)	(1.050)
γ_2		-1.314	-1.346	-1.354	-0.632	-1.219	-1.170	-1.241	-1.105
		(1.097)	(2.541)	(2.113)	(6.226)	(0.984)	(0.971)	(1.190)	(1.059)
γ_3		-1.430	-1.461	-1.534	-1.184	-1.476	-1.273	-1.534	-1.448
		(1.293)	(1.312)	(2.231)	(6.495)	(1.212)	(1.103)	(1.509)	(1.273)
γ_4		-1.462	-1.413	-1.773	-1.288	-1.398	-1.427	-1.490	-1.489
		(1.407)	(1.381)	(2.325)	(6.371)	(1.315)	(1.104)	(1.281)	(1.196)
γ_5		-1.607	-1.439	-1.797	-1.348	-1.599	-1.508	-1.701	-1.577
		(1.364)	(1.382)	(2.380)	(6.197)	(1.264)	(1.177)	(1.418)	(1.225)
π_1		0.195	0.196	0.212	0.207	0.200	0.190	0.198	0.196
		(0.137)	(0.136)	(0.169)	(0.153)	(0.144)	(0.134)	(0.134)	(0.116)
π_2		0.193	0.202	0.191	0.222	0.196	0.204	0.197	0.197
		(0.138)	(0.128)	(0.160)	(0.143)	(0.139)	(0.138)	(0.129)	(0.121)
π_3		0.203	0.207	0.206	0.202	0.201	0.204	0.194	0.199
		(0.141)	(0.137)	(0.168)	(0.154)	(0.152)	(0.142)	(0.123)	(0.112)
π_4		0.200	0.197	0.195	0.189	0.197	0.204	0.203	0.204
		(0.146)	(0.136)	(0.168)	(0.152)	(0.145)	(0.141)	(0.127)	(0.119)
π_5		0.208	0.197	0.195	0.181	0.206	0.198	0.208	0.205
		(0.145)	(0.135)	(0.160)	(0.146)	(0.150)	(0.141)	(0.131)	(0.116)

Table 3.3: Log-Normal, Gamma, Inverse Gaussian and arbitrary frailty model with Weibull baseline hazard and four covariates simulated data, fitted by non-parametric frailty, 600 data sets each with sample sizes of 500 and 5000.

3.6 Results on breast cancer recurrence data

This is to illustrate the relationship of the models discussed above to the breast cancer data given in chapter two and appendix A. The study included 2850 patients. Their status at the time of first recurrence is given in 3.4. Around 38% of them experienced one of the five types of failures (competing risks). Metastasis and regional recurrence are the most frequent failure type with 15.8% and 9.2% respectively.

Failure Type	N	Percent
LOCAL RECURRENCE	169	5.9%
REGIONAL RECURRENCE	261	9.2%
METASTASIS	451	15.8%
DIED FROM BREAST CANCER	185	6.5%
DIED FROM OTHER CAUSES	128	4.5%
CENSORED	1656	58.1%
Total	2850	100.0%

Table 3.4: Patients status at time of first recurrence.

stage(I) tumour size less than 2cm and no nodes involved.

stage(II) tumour size between 2cm and 5cm and no nodes involved, or tumour size less than 5cm with nodes.

stage(III) tumour size more than 5cm with or without nodes.

stage(IV) any tumour size with the presence of distant metastasis, i.e., disease elsewhere other than the breast or local nodes.

Variable	Categories	N	%
STAGE	STAGE1	2107	73.9
	STAGE2	457	16.0
	STAGE3	179	6.3
	STAGE4	107	3.8
SURGERY TYPE	NONE	45	1.6
	INCISION BIOPSY	186	6.5
	EXCISION BIOPSY	712	25.0
	SIMPLE MASTECTOMY	543	19.1
	RADICAL MASTECTOMY	36	1.3
	WIDE LOCAL EXCISION AND AXILLARY CLEARANCE	479	16.8
	SURGERY AFTER NEO ADJUVANT CHEMOTHERAPY	73	2.6
	RADICAL MAST AND AXILLARY CLEARANCE	776	27.2
HISTOLOGY	DUCTAL	1856	65.1
	LOBULAR	336	11.8
	DCIS (Ductal Carcinoma In Situ)	412	14.5
	OTHER	246	8.6
DATE OF PRIMARY SURGERY	BEFORE 1990	1130	39.6
	1990+	1720	60.4
ANY NEO OR ADJUVANT CHEMOTHERAPY	NO	1525	53.5
	YES	1325	46.5
MENOPAUSAL STATUS	PRE	788	27.6
	POST	2062	72.4
ANY ADJUVANT RADIOTHERAPY	NO	1813	63.6
	YES	1037	36.4
SIDE OF THE BODY	RIGHT	1354	47.5
	LEFT	1496	52.5
Age		Mean ± SD	
		58.11 ± 12.97	

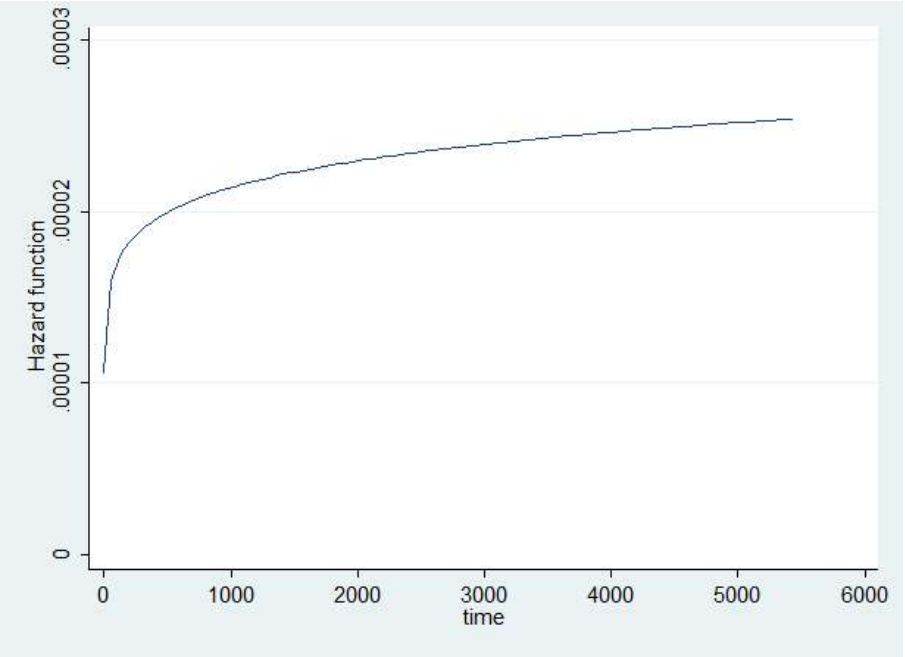
Table 3.5: Independent variables included in the models.

For the models used to fit this data, the qualitative variables are converted into dummy variables and set the first category as the reference category except for *Surgery type* the last category is used as the reference, see appendix A. Since this chapter focuses on univariate frailty, death due to breast cancer is considered as the event of interest and all other recurrences as independent censored observations. For simple models, three different software packages are used to fit the models, Gauss, R and STATA. For advanced models, a self-written code in both Gauss and R are used, see appendix C. STATA was used only to check the codes written in Gauss and R since it contains a build in univariate frailty models especially for Gamma and Inverse Gaussian. In multivariate case only Gauss program was used to fit the model parameters. The optimisation procedure in Gauss showed robustness in reaching the maximum likelihood estimates, while the optimisation procedure in R was sensitive to the parameters' initial values. Table 3.6 summarises the regression analysis of the data using three different models. The first model is the Cox proportional model which does not assume any parametric distribution for the baseline hazard function, model (2.6.1). The second model is an independent Weibull model where the distribution of baseline hazard function is Weibull and no frailty, model (2.6.3). The third model assumes Weibull distribution baseline hazard function and Gamma frailty distribution as model (3.4.5). It is clear that the results in the first and second columns of Table 3.6 for the Cox proportional hazard as close to those of the Wiebull hazard which is a well-known result in the literature. To compare the second model (i.e. reference model) with the third model (i.e. frailty model), a test of $\tau^2 = 0$ vs $\tau^2 \neq 0$, where τ^2 is the variance of the frailty distribution is used. The likelihood ratio test is

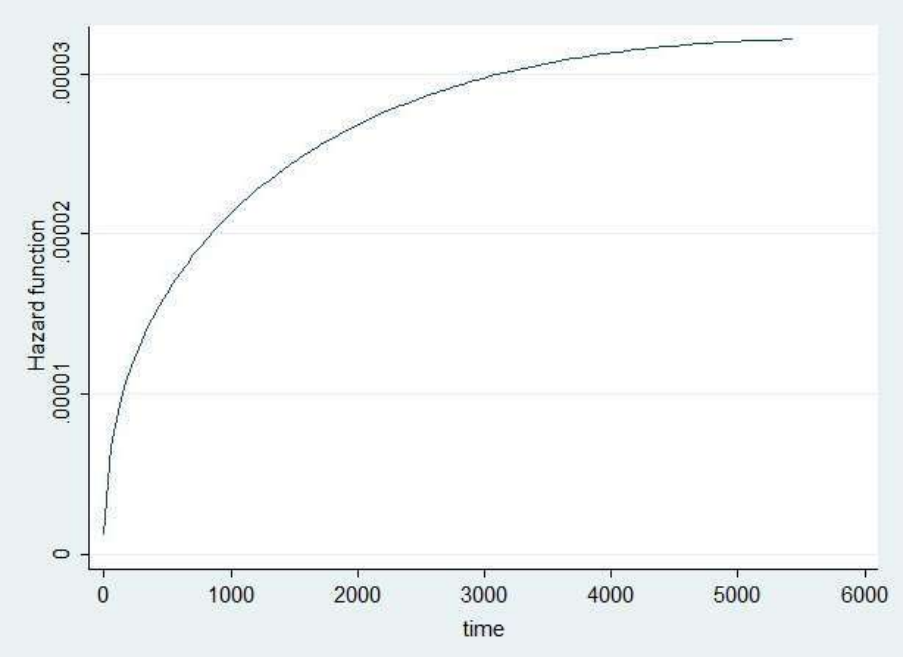
$$-2log\left(\frac{likelihood_{(reference)}}{likelihood_{(frailty)}}\right) = (3785.24 - 3764.17) = 21.07,$$

which has a chi-square distribution with one degree of freedom. In this case the frailty model is preferred to the reference model since (P-value < 0.0005). Figure 3.2 shows the hazards of death due to breast cancer after initial treatment, for an average individual (function evaluated at the mean value of the covariates), for each of the independent and the frailty models. In the case of the independent model, the hazard of death due to breast cancer seems unrealistic as it reaches its peak directly after initial treatment and then decreases over time. In contrast, the frailty model displays a more realistic hazard function for death due to breast cancer, whereas the lowest hazard level is directly after initial treatment and then the hazard increases with time. The hazard of death due to breast cancer is always lower when employing the frailty model which is expected as the effects of any unobserved frailty on prognostic factors which have been appropriately controlled.

An important concept used in survival analysis is hazard ratio (HR) which can be defined as the ratio of the hazard rate of one set of covariates to another set of covariates in the model. For example, the point estimate of the HR of *Stage2* compared to *Stage1* (i.e. the reference category) is $\widehat{HR} = e^{\beta_2} = e^{0.545} = 1.725$; in particular, the hazard for patients in *Stage2* is 1.725 times the hazard for patients in *Stage1*. Table 3.6 shows that applying frailty increases the HR for most of the factors in the model. Among the factors entered in the three models, *Age, stage2, stage3, stage4, surgtype1, surgtype2, hist3* and *cohort* have a significant effect on breast cancer mortality. On the other hand, only *hist2* has a significant effect in the Cox PH model and the independent Wiebull model. The Weibull scale parameter is different between the second model and the third model. In the independent Weibull model, α is close to one which means constant hazard (i.e. Exponential baseline hazard), while in the Gamma frailty, $\alpha = 1.41$ indicating an increasing hazard which is more realistic in human studies.



(a) Independent (no frailty)



(b) Gamma frailty

Figure 3.2: Weibull hazards of died from breast cancer for independent and frailty models.

Variable	Cox PH	Weibull hazard	Wiebull-Gamma frailty
AGE	**0.030 (0.009)	**0.030 (0.008)	**0.037 (0.011)
STAGE2	**0.545 (0.189)	**0.548 (0.211)	*0.613 (0.248)
STAGE3	*0.728 (0.305)	*0.726 (0.289)	*0.933 (0.373)
STAGE4	**2.152 (0.283)	**2.163 (0.270)	**3.722 (0.510)
SURGTYPE1	**1.981 (0.526)	**2.078 (0.416)	**2.668 (0.606)
SURGTYPE2	**1.354 (0.416)	**1.387 (0.311)	**1.599 (0.395)
SURGTYPE3	-0.073 (0.356)	-0.056 (0.297)	-0.089 (0.337)
SURGTYPE4	-0.253 (0.391)	-0.243 (0.297)	-0.363 (0.343)
SURGTYPE5	-44.78 (181.8)	-8.152 (48.73)	-10.901 (181.1)
SURGTYPE6	-0.552 (0.436)	-0.549 (0.389)	-0.644 (0.417)
SURGTYPE7	0.499 (0.495)	0.530 (0.514)	0.543 (0.577)
HIST2	*0.465 (0.206)	*0.472 (0.207)	0.439 (0.246)
HIST3	*-1.472(4.690)	*-1.471 (0.614)	*-1.484 (0.637)
HIST4	0.146 (0.238)	0.121 (0.211)	0.094 (0.267)
COHORT	**0.578 (0.200)	**0.620 (0.192)	**0.703 (0.242)
CHEMO	0.019 (0.233)	0.043 (0.191)	0.114 (0.216)
MENO	0.211 (0.271)	0.204 (0.284)	0.181 (0.333)
RADIO	-0.455 (0.328)	-0.457 (0.298)	-0.421 (0.326)
SIDE	0.242 (0.153)	0.248 (0.151)	0.277 (0.184)
CONSTANT		**-13.700 (0.834)	**-16.488 (1.235)
LN(α)		0.097 (0.060)	**0.343 (0.080)
(FRAILTY) τ^2			*2.171 (0.750)
-2 Log Likelihood	2421.70	3785.24	3764.17

Table 3.6: Results of breast cancer data: died from breast cancer with the cox proportional hazard, Weibull hazard and Weibull-Gamma frailty. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

P-value for testing $H_0 : \beta_i = 0$ vs $H_1 : \beta_i \neq 0$

Although the standard errors of the third model are higher than the other models, it is more appropriate in making inference about the factors effect on mortality of breast cancer and ignoring the frailty underestimate the model parameters. Using the results of the third model, one can conclude the following remarks. As expected, as *age* increases breast cancer mortality increases. There is a significant difference between the four stages of the disease in breast cancer mortality. The hazard ratio of *stage2* is $e^{0.613} = 1.85$ times as for *stage1* (the reference category). The same for *stage3* and *stage4*, the hazard ratio of *stage3* is $e^{0.933} = 2.54$ times as for *stage1*, and for *stage4* it is $e^{3.72} = 43.3$ times as for *stage1*. For surgery type, *surgtype1* (None) and *surgtype2* (Incision biopsy) have significantly higher hazard ratios compared to *surgtype8* (Radical mast and axillary clearance), the reference category. This seems unexpected, but actually this is because most of these two groups occur either in *stage3* or *stage4* of the disease. For histology, the only significant difference is between *hist3* and *hist1* (reference category). Patients with *hist3* (DCIS , Ductal carcinoma in situ) have a hazard ration of $e^{-1.48} = 0.227$ times as *hist1* (Ductal). Finally, The hazard ratio of mortality due to breast cancer of *Cohort* with the date of primary surgery after 1990 is $e^{0.703} = 2.02$ times as those with before 1990. Unfortunately, the likelihood ratio test cannot be applied to different frailty distributions in Table 3.7 since they are not nested. Two criteria are usually used for model selection among a finite set of models. The first one is AIC which stands for Akaike Information Criterion developed by Akaike (1974). The second one is BIC which stands for Bayesian Information Criterion by Schwarz (1978). These criteria introduce a penalty term for the number of parameters in the model, and the model with the smallest AIC or BIC is preferred. By both criteria, the non-parametric frailty model has the smallest values and hence is preferable when compared to other frailty distributions. The non-parametric model is preferable since it has the smallest standard errors for all parameters estimate.

	Frailty distribution			
Variable	Gamma	Inverse Gaussian	Log-Normal	Non-parametric
AGE	**0.037(0.011)	**0.036(0.011)	**0.041(0.015)	**0.036(0.009)
STAGE2	*0.613(0.248)	*0.647(0.266)	**0.685(0.313)	*0.530(0.221)
STAGE3	*0.933(0.373)	*1.008(0.406)	**1.139(0.502)	**0.826(0.311)
STAGE4	**3.722(0.510)	**3.340(0.490)	**4.069(0.985)	**3.426(0.309)
SURGTYPE1	**2.668(0.606)	**2.725(0.628)	**3.172(0.927)	**2.307(0.463)
SURGTYPE2	**1.599(0.395)	**1.694(0.440)	**1.940(0.582)	**1.388(0.325)
SURGTYPE3	-0.089(0.337)	-0.054(0.347)	-0.092(0.394)	-0.141(0.309)
SURGTYPE4	-0.363(0.343)	-0.355(0.359)	-0.446(0.419)	-0.374(0.311)
SURGTYPE5	-10.9(181.05)	-7.89(38.634)	-7.89(31.485)	-7.889(41.46)
SURGTYPE6	-0.644(0.417)	-0.657(0.437)	-0.792(0.510)	-0.616(0.397)
SURGTYPE7	0.543(0.577)	0.555(0.609)	0.521(0.695)	0.521(0.538)
HIST2	0.439(0.246)	0.474(0.260)	0.490(0.303)	0.403(0.218)
HIST3	*-1.484(0.637)	*-1.573(0.656)	**-1.740(0.745)	*-1.411(0.619)
HIST4	0.094(0.267)	0.100(0.270)	0.062(0.310)	0.115(0.230)
COHORT	**0.703(0.242)	**0.722(0.250)	**0.839(0.315)	**0.609(0.205)
CHEMO	0.114(0.216)	0.098(0.227)	0.145(0.257)	0.099(0.199)
MENO	0.181(0.333)	0.201(0.342)	0.275(0.391)	0.210(0.305)
RADIO	-0.421(0.326)	-0.477(0.339)	-0.516(0.383)	-0.397(0.306)
SIDE	0.277(0.184)	0.295(0.190)	*0.365(0.224)	0.239(0.162)
CONSTANT	**-16.49(1.235)	**-16.29(1.462)	**-19.91(3.934)	
LN(α)	**0.343(0.080)	**0.332(0.104)	**0.466(0.194)	**0.290(0.060)
(FRAILTY)τ^2	*2.171(0.750)	4.968(4.081)	**1.998(0.487)	
-2 Log Likelihood	3764.17	3770.04	3802.20	3756.78
AIC	3808.17	3814.04	3846.20	**3796.78**
BIC	3939.18	3945.05	3977.21	**3915.88**

Table 3.7: Results of breast cancer data: Died from breast cancer assuming different frailty distributions. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

Number of Points	Number of parameters	-2log-likelihood	AIC	BIC
1	23	3785.2368	3831.2	3968.2
2	25	3758.7958	**3808.8**	**3957.7**
3	27	3756.0952	3810.1	3970.9
4	29	3756.1028	3814.1	3986.8
5	31	3756.1026	3818.1	4002.7

Table 3.8: Non-parametric models with different mass points.

For the Log-Normal frailty 128 quadrature mass points were used to fit the model whereas for the non-parametric model only five mass points were used. Table A.4 lists the estimates of the Log-Normal frailty with different number of quadrature points. The results are very close when using nine points or more. One important issue about the non-parametric frailty is the number of mass points which should be used to fit the model. Since these models are not nested the log-likelihood ratio test can not be used the number of mass points. Table 3.8 lists the values of AIC and BIC of mortality due to breast cancer with different number of mass points. Both AIC and BIC suggest that two points as the suitable number of mass points. The parameters estimates with three mass points are very similar to those with four and five point points (see Table A.3).

3.7 Summary

This chapter described the main features of univariate frailty models and reviewed the most commonly used frailty distributions. Gamma, Inverse Gaussian and the Log-Normal distributions are intensively studied in the literature. Some software like STATA can be used to fit both Gamma and Inverse Gaussian frailty models. The Log-Normal frailty is available in SAS-system through the PHREG Procedure, but in this thesis a self-written Gauss code

is used. Other frailty distributions are not available in commercial software and need self-written code in some programming languages. The simulation studies showed that the model fits are sensitive to the choice of the frailty distribution. To overcome this problem a non-parametric frailty model is also presented. The non-parametric frailty model was capable to fit the model irrespective to the frailty distribution. Different models were used to analyse the breast cancer data in which the event of interest is died from breast cancer. These models include: Cox proportional hazard, Weibull hazard, Weibull-Gamma frailty, Weibull-Inverse Gaussian frailty, Weibull-Log-Normal frailty and Weibull-nonparametric frailty. By applying these modes one can conclude the following: First, ignoring the frailty underestimates the model parameters. Second, the standard errors of parameter estimates are the smallest in the case of non-parametric frailty. Only small numbers of mass points are needed in the non-parametric frailty to reach the maximum likelihood estimates of the model parameters which mean less time is required to fit the model compared to the Log-normal frailty. All models discussed in this chapter are univariate frailty models where the main goal of adding the frailty component is to take into account the unobserved risk factors. In many studies, individual are clustered in subgroups or the individual could face more than one failure type. The next chapter discusses these types of studies using multivariate frailty models.

Chapter 4

Multivariate frailty in competing risks models

4.1 Introduction

Multivariate frailty models are used when there are repeated measures or clustering. Repeated data occur in case of longitudinal data, concerning multiple recurrences of an event for the same individual. Clustered data occur when individuals fall into groups like families or hospitals. The difficulties of working with this kind of data arise from the dependence of individuals upon the social context of their groups, or of repeated measures upon the individuals concerned. Such dependence usually arises because individuals in the same group are related to each other or because of the recurrence of an event for the same individual. Multivariate frailty models have been used frequently for modelling dependence in multivariate time-to-event data (Clayton 1978, Oakes 1982, Hougaard 2000 and Yashin et al. 1995).

As mentioned, linear mixed models are usually used to analyse repeated measures and cluster data for one outcome (response), but when multiple outcomes are measured at each time-point or for each cluster, a joint model is required which allows for a correlation

structure between the different outcomes. Many authors have studied such models; for example, Thum (1997) studied models with multivariate clustered data in the context of hierarchical modelling, Gueorguieva (2001) used joint modelling of a continuous- and a binary-outcome measure in a developmental toxicity study on mice. Also in a longitudinal setting, Chakraborty et al. (2003) obtained estimates of the correlation between blood and semen HIV-1 RNA using a joint random-effects model. Although Models in all of these examples can be applied to any number of responses, the computational procedure of model fitting is only feasible for a limited number of responses. A recent series of papers proposed a model-fitting procedure that is applicable irrespective of the number of responses. Fieuws and Verbeke (2004) used a joint model of random-effects in a bivariate setting with longitudinally measured continuous outcomes by fitting different mixed models joined by specifying a common distribution for their random-effects. Fieuws and Verbeke (2006a) proposed a method that allows joint analysis of multivariate repeated measures of a relatively high dimension. The method is based on fitting bivariate mixed models for all pairs of outcomes. As long as each bivariate mixed model can be fitted, estimates can be obtained for the full multivariate mixed model. Fieuws et al. (2006b) applied the pairwise modelling strategy proposed by Fieuws and Verbeke (2006a), to obtain parameter estimates of high dimensional LMMs for binary questionnaire data, where all possible bivariate mixed models were fitted and where the inference that follows from pseudo-likelihood theory has been proposed as a solution. Fieuws et al. (2007) combined the proposed methods by Fieuws and Verbeke (2006a) and Fieuws et al. (2006b); they used the pairwise modelling of repeated measures for continuous as well as for binary questionnaire data, so that the approach is sufficiently flexible to allow for different types of models for the different outcomes (i.e. linear, non-linear, and generalised linear).

Despite the big similarity between LMMs and multivariate frailty, many methods that could be applicable to both models, the multivariate frailty models need special treatment. The aim of the frailty model is to take into account the presence of the correlation between the multivariate survival times. The first extension of univariate frailty to multivariate frailty is the shared frailty model. In shared frailty models, individuals in the same group or cluster are assumed to share the same frailty. Thus, it accommodates the heterogeneity among clusters rather than among individuals. Correlated frailty models are an extension of shared frailty. In correlated frailty models, frailties are correlated and have the same set of marginal distributions. If the frailties follow a multivariate distribution with a general correlation structure, then it is a full multivariate frailty model.

4.2 Shared frailty models

In many studies individuals are grouped into clusters, such as families, hospitals or schools. Similar to univariate frailty, shared frailty models are used to take into account the heterogeneity between clusters due to exclusion of important shared covariates. In these models, it is assumed that all individuals in the same cluster share the same frailty. It was introduced by Clayton (1978) and extensively studied by Hougaard (2000). Many authors have studied shared frailty models with different distributions, including Gamma. Clayton (1978), Gill (1989), Klein (1992b), Yu (2006), inverse Gaussian, Whitmore and Lee (1991), Henderson and Oman (1999) and Log-Normal, McGilchrist (1993), Ripatti et al. (2002), Vaida and Xu (2000). For the Bayesian approach see, Sinha (1993), Sahu et al. (1997) and Yin and Ibrahim (2005).

In a shared frailty model there are two main assumptions: First, the failure times are conditionally independent given the frailties. Second, the random effect Z of the j^{th} cluster

$(j = 1, \cdots, k)$ is assumed to be constant over time and common to all individuals in the same cluster. The hazard function of the i^{th} individual in the j^{th} cluster conditional on the random effect is given by

$$h(t_{ij}, \boldsymbol{x_{ij}}|z_j) = z_j h(t_{ij}, \boldsymbol{x_{ij}}) = z_j h_0(t_{ij}) exp(\boldsymbol{x_{ij}'}\boldsymbol{\beta}),$$

where z_j are the frailties, assumed to be iid (identically and independently distributed) with the density $g(z, \theta)$ and parameter vector θ. $h_0(t_{ij})$ is the baseline hazard function, $\boldsymbol{x_{ij}}$ is the vector of covariate of the i^{th} individual in the j^{th} cluster and $\boldsymbol{\beta}$ is the fixed effect vector. Here the frailty term is included to take into account the heterogeneity between clusters and not individuals. Assuming that the failure times are conditionally independent given the frailties, the conditional joint survival function is given by

$$S(t_{ij}, \boldsymbol{x_{ij}}|z_j) = \prod_{j=1}^{k} S_j(t_{ij}, \boldsymbol{x_{ij}}|z_j).$$

The unconditional joint survival function is given by

$$S(t_{ij}, \boldsymbol{x_{ij}}) = E_Z[S(t_{ij}, \boldsymbol{x_{ij}}|z_j)] = E_Z[\exp(-z_j \sum_{j=1}^{k} H_{0j}(t_{ij}) e^{\boldsymbol{x_{ij}'}\boldsymbol{\beta}})] = \mathcal{L}_Z \left[\sum_{j=1}^{k} H_{0j}(t_{ij}) e^{\boldsymbol{x_{ij}'}\boldsymbol{\beta}} \right].$$

where $\mathcal{L}_Z[.]$ is the Laplace transformation with respect to the frailty random variable Z and $H_{0j}(t_{ij})$ is the cumulative baseline hazard function of the j^{th} cluster. Thus, the unconditional joint survival function can be expressed as the Laplace transform of the sum of the cumulative baseline hazards with respect to the frailty distribution. The unconditional joint density function is given by differentiation of the survival function with respect to $t_{i1}, \cdots, t_{ik}$

$$f(t_{ij}, \boldsymbol{x_{ij}}) = (-1)^k \prod_{j=1}^{k} h_{0j}(t_{ij}) e^{\boldsymbol{x_{ij}'}\boldsymbol{\beta}} \mathcal{L}_Z^{(k)} \left[\sum_{j=1}^{k} H_{0j}(t_{ij}) e^{\boldsymbol{x_{ij}'}\boldsymbol{\beta}} \right].$$

Another formulation of the above procedure is through the marginal survival functions

$$S_j(t_{ij}, \boldsymbol{x_{ij}}) = \mathcal{L}_Z[H_{0j}(t_{ij})e^{\boldsymbol{x}_{ij}'\boldsymbol{\beta}}].$$

The unconditional joint survival function is given by

$$S(t_{ij}, \boldsymbol{x_{ij}}) = E_Z[S(t_{ij}, \boldsymbol{x_{ij}}|z_j)] = \mathcal{L}_Z\left[\sum_{j=1}^{k} \mathcal{L}_Z^{-1}[S_j(t_{ij}, \boldsymbol{x_{ij}})]\right].$$

There are some limitations with shared frailty models as mentioned by Xue and Brookmeyer (1996). Firstly, it forces the unobserved factors to be the same within the cluster, which is generally not appropriate. Secondly, the dependence between survival times within the cluster is based on marginal distributions of survival times. Thirdly, in most cases shared frailty will only induce positive association within the cluster.

4.3 Correlated frailty models

Due to their limitations, shared frailty models can be extended to correlated frailty models, where each cluster has its own frailty distribution. Most of the correlated frailty models developed until now are bivariate frailty models or restricted multivariate models. Such frailties are often constructed using independent additive components with one common component for both frailties to create the correlation between the frailties. There exists a need for more flexibility in modelling correlation. The difficulty in these models is that related outcomes have different but dependent frailties. (Yashin et al. (1995), Yashin and Iachine (1999)) introduced a correlated Gamma frailty model and discussed its identifiability conditions. Vaida and Xu (2000) suggested a bivariate frailty model in a slightly different setting, dos Santos et al. (1995) used a combination of a shared Log-Normal and a

Gamma frailty model on breast cancer data. Wienke et al. (2005) compared different bivariate correlated frailty models and possible estimation strategies. Zahl (1997) used several correlated Gamma frailty models to model the excess hazard. Li (2002) proposed a multivariate Gamma frailty model in a genetic situation. For different frailty distributions see Wienke (2007, 2010a) and Duchateau and Janssen (2008). A number of authors have used a Bayesian approach for analysis of correlated frailty models; for instant, Xue and Ding (1999), Kheiri et al. (2005), Wienke et al. (2005) Locatelli et al. (2004, 2007), Yin (2008) and Cai (2010).

In contrast to a shared frailty model, correlated frailty models allow different frailties between clusters and individuals in the same cluster. The hazard function of the i^{th} individual in the j^{th} cluster conditional on the random effect is given by

$$h(t_i, \boldsymbol{x_{ij}}|z_{ij}) = z_{ij}h(t_{ij}, \boldsymbol{x_{ij}}) = z_{ij}h_0(t_{ij})exp(\boldsymbol{x_{ij}^{'}\beta_j}).$$

where z_{ij} are the frailties, assumed to be iid with a joint density $g(z_{i1}, \cdots, z_{ik}, \theta_j)$ with parameter vector θ_j. The Laplace transformation is more complicated in this case and a different formulation is needed. Assuming the conditional independence of failure times given the frailty, the unconditional likelihood function is given by

$$L(t_{ij}, \boldsymbol{x_{ij}}) = \prod_i \int_{R^+} \cdots \int_{R+} \prod_j \left[(z_{ij}h_{0j}(t_{ij})e^{\boldsymbol{x_{ij}^{'}\beta_j}})^{d_{ij}} \exp(-z_{ij}H_{0j}(t_{ij})e^{\boldsymbol{x_{ij}^{'}\beta_j}}) \right] \times f_{Z_j}(z_{i1}, \cdots, z_{ik}) dz_{i1} \cdots dz_{ik}. \tag{4.3.1}$$

For bivariate correlated Gamma frailty, see Yashin et al. (1995), restricted high dimension of correlated Gamma frailties, Yashin and Iachine (1999), for four dimension Gamma frailty, Giard et al. (2002) and different correlated frailty distributions Wienke (2007, 2010a). As mentioned, correlated frailty models are usually constructed using independent positive

components and hence they generate a restricted positive correlation coefficient which in many cases is not appropriate. A general class of multivariate frailty models is when the frailties are assumed to follow a multivariate distribution with a general structure of covariance matrix. The most frequent distribution used in multivariate frailty is the Log-Normal distribution. Other distributions are limited in the literature. The multivariate Log-Normal frailty models are more flexible than other distributions but numerical integration is needed to calculate the joint survival function.

4.4 Competing risks models

In studies of survival, subjects may be at risk of failure due to more than one cause, the so-called competing risks analysis or multiple causes of failure. The objective of such analysis is usually to determine risk rates of failure due to one cause while taking into account the presence of the other causes. There are two approaches in the literature to analyse competing risks: the first emphasises cause-specific hazard functions and sub-distribution functions for a particular kind of failure, (Prentice et al. 1978, Fine and Gray 1999, Lunn and McNeil 1995). The second, approaches the subjects through the concept of latent failure times, where there is an inherent failure time for each type of failure and only one such time - the smallest - is observable, (Slud et al. (1988), Kalbfleisch and Prentice (2002)). In this thesis, the emphasis is on the concept of latent failure time. Parametric competing risk models, in which it is assumed that the failure times are coming from a known distribution with monotonically increasing or decreasing baseline hazard, is widely used in practice perhaps due to its simplicity, for example (Hougaard (2000), Lambert et al. (2004), Oskrochi and Crouchley (2004)) perhaps due to its simplicity. In conventional survival analysis the competing risks are usually assumed independent, Han and Hausman (1990). Semi-parametric and non-parametric competing risks models are discussed by many authors. For instance (Gelfand

et al. 2000, Abbring and van den Berg 2003, Hudgens et al. 2001), Jewell et al. (2003)). In univariate survival data, there is usually the following data for the i^{th} individual: the failure time (Y_i), the censoring variable (δ_i) and some covariates ($\boldsymbol{x_i}$). In multivariate survival data, there is either have cluster or recurrence data and the following data for the i^{th} individual in the j^{th} cluster is observed: the failure times (Y_{ij}), the censoring variable (δ_{ij}) and some covariates ($\boldsymbol{x_{ij}}$). In competing risks data, the following data for the i^{th} individual is observed: the failure time (time to the first failure) (Y_i), the failures indicator variables ($\delta_{ik}, k = 1, ..., K$) and some covariates ($\boldsymbol{x_i}$). See Table 4.1.

4.5 Frailty in Competing Risks Models

Random effects or frailty in competing risk models consist of two underlying distributions: the conditional distribution of the response variables (failure times), given the random effect depending on the explanatory variables each with a failure type specific random effect; and the distribution of the random effect in the population (frailty distribution). In this situation, the distribution of interest is the unconditional distribution of the response variables which may or may not have a tractable form. Bandeen-Roche and Liang (2002) described the association of time to multivariate failure in the presence of a competing risk. As mentioned above frailty models in the presence of competing risks involve the conditional distribution failure time given the frailty and the distribution of the frailty. Fahrmeir and Tutz (1994) and Oskrochi and Davies (1997) implemented the Cholesky decomposition for multivariate frailty models. The Cholesky decomposition decomposes the frailty variance-covariance matrix into triangular matrices to convert the integration over a multivariate distribution to multiple integrals over independent univariate Normal distributions.

Univariate data

ID	Y_i	δ_i	x_{i1}	$\cdots$	x_{ip}
1	Y_1	δ_1	x_{11}	$\cdots$	x_{1p}
2	Y_2	δ_2	x_{21}	$\cdots$	x_{2p}
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
n	Y_n	δ_n	x_{n1}	$\cdots$	x_{np}

Multivariate data

ID	Y_{ij}	cluster	δ_{ij}	x_{ij1}	$\cdots$	x_{ijp}
1	Y_{11}	1	δ_{11}	x_{111}		x_{11p}
2	Y_{21}	1	δ_{21}	x_{211}		x_{21p}
.	.	.	.	.		.
.	.	.	.	.		.
.	.	.	.	.		.
n_1	$Y_{n_1 1}$	1	$\delta_{n_1 1}$	$x_{n_1 11}$	$\cdots$	$x_{n_1 1p}$
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
1	Y_{1J}	J	δ_{1J}	x_{1J1}		x_{1Jp}
2	Y_{2J}	J	δ_{2J}	x_{2J1}		x_{2Jp}
.	.	.	.	.		.
.	.	.	.	.		.
.	.	.	.	.		.
n_J	$Y_{n_J J}$	J	$\delta_{n_J J}$	$x_{n_J J1}$	$\cdots$	$x_{n_J Jp}$

Competing risk data

ID	Y_i	δ_{i1}	$\cdots$	δ_{iK}	x_{i1}	$\cdots$	x_{ip}
1	Y_1	δ_{11}	$\cdots$	δ_{1K}	x_{11}	$\cdots$	x_{1p}
2	Y_2	δ_{21}	$\cdots$	δ_{2K}	x_{21}	$\cdots$	x_{2p}
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
n	Y_n	δ_{n1}	$\cdots$	δ_{nK}	x_{n1}	$\cdots$	x_{np}

Table 4.1: Univariate, multivariate and competing risks data presentation.

Fine and Gray (1999) proposed a novel semi-parametric proportional hazards model for the sub-distribution. Using the partial likelihood principle and weighting techniques, they derived estimation and inference procedures for the finite-dimensional regression parameters under a variety of censoring scenarios. They gave a uniformly consistent estimator for the predicted cumulative incidence for an individual with certain covariates, confidence intervals and bands can be obtained analytically or with an easy-to-implement simulation technique. Fine et al. (2001) considered semi-competing risk models, in which a terminal event censors a non-terminal event but not vice versa. The joint distribution of the events is formulated via Gamma frailty model with marginal distributions unspecified. They showed that the dependence between morbidity and mortality can be estimated separately from their marginals under a Gamma frailty copula in the region of observable data. Jiang et al. (2004) considered semi-competing risk models where mortality and morbidity may be correlated and mortality may censor morbidity. They proposed a semi-parametric estimator for the survival function based on a joint model for the two time-to-event variables, which utilises the Gamma frailty specification in the region of the observable data. They extended the methods of Fine et al. (2001) for the left truncated semi-competing risk problem, developed an estimator for the Gamma frailty parameter under left truncation and derived a closed form estimator for the marginal distribution of the non-terminal event. Naskar et al. (2005) proposed a dependent competing risks model where dependency is induced through the mixture of various failure types and a frailty component. The frailty term is modelled non-parametrically using a Dirichlet Process (DP). They considered a semi-parametric mixture model for analysing clustered competing risks data. Conditional on cluster-specific quantities, the joint distribution of the failure time and event indicator can be expressed as a mixture of the distribution of time to failure due to a certain type (or specific cause), and the failure type distribution. They assumed that the marginal probabilities of various failure types

(competing risks) are logistic functions of some covariates. The cluster-specific quantities are subject to some unknown distribution that causes frailty. Lu and Tsiatis (2005) compared two approaches for the competing risks model with missing cause of failure. Under the assumption that the cause of death is missing at random, they compared the Goetghebeur and Ryan (1995) partial likelihood approach with the one by Dewanji (1992). They showed that the Dewanji partial likelihood estimator for the regression coefficients is consistent and asymptotically Normal. While the Goetghebeur and Ryan estimator is more robust estimator against mis-specification of proportional baseline hazards. Finkelstein and Esaulovac (2006, 2008) discussed the asymptotic behaviour of competing risks models with correlated frailty in univariate and bivariate cases. They consider a set of absolutely continuous distributions of a lifetime random variable.

In the next section, a review the correlated Gamma frailty and its application to competing risks framework is given. In section 4.7, a new proposed correlated Inverse Gaussian frailty model and its application to competing risks are presented. A general multivariate Inverse Gaussian frailty model without any restriction on the correlation structure between the frailties is given in section 4.8. In addition, a flexible multivariate frailty model that can be applied whatever the original distribution of the frailty in section 4.10 is proposed. In this model, a non-parametric multivariate frailty presented.

4.6 Correlated Gamma frailty model

In this section, a summary of the main aspects of the correlated Gamma frailty, which has been intensively studied in the literature, is listed. Yashin et al. (1995) introduced model of bivariate survival based on the notion of correlated individual frailty. This model usually refereed to as the additive model since the frailty variable is constructed as the sum

of independent variables. Many authors have studied this model and its extensions and applications, for example see (Yashin and Iachine 1995a,b, 1999, Wienke et al. 2002, 2003, Kheiri et al. 2005, Abbring and van den Berg 2007 and Duchateau and Janssen 2008) among others. The procedure starts by defining a system of equations (frailty) as the sum of independent Gamma distributions a common variable to create the dependency and then derive the unconditional survival function as the product of their Laplace transformation. For the bivariate case, the results shown by Yashin et al. (1995) and Wienke (2007, 2010a) are given.

Let $Z_1 = \frac{\lambda_0}{\lambda_1} X_0 + X_1$ and $Z_2 = \frac{\lambda_0}{\lambda_2} X_0 + X_2$, where X_0, X_1, X_2 are independent random variables with Gamma distribution $X_i \sim \Gamma(\alpha_i, \lambda_i), i = 0, 1, 2$. The distribution of Z_1 and Z_2 is Gamma with mean one and variance $V[Z_i] = 1/\lambda_i = \tau_i^2, i = 1, 2$, the correlation coefficient between the two frailties is given by

$$\rho_z = \frac{\alpha_0}{(\alpha_0 + \alpha_1)(\alpha_0 + \alpha_2)}.$$

The unconditional survival function is

$$S(t_{i1}, t_{i2}) = \frac{S_1(t_{i1})^{1-\rho\frac{\sigma_1}{\sigma_2}} S_2(t_{i2})^{1-\rho\frac{\sigma_2}{\sigma_1}}}{\left(S_1(t_{i1})^{-\sigma_1^2} + S_2(t_{i2})^{-\sigma_1^2} - 1\right)^{\frac{\rho}{\sigma_1\sigma_2}}}. \tag{4.6.1}$$

This procedure is restricted since it can only create a positive correlation between the frailties which may limits its applications. Moreover this correlation must satisfy the following condition.

$$0 \leq \rho \leq \min\left(\frac{\tau_1}{\tau_2}, \frac{\tau_2}{\tau_1}\right).$$

To build the likelihood function, all partial derivatives of the bivariate survival function in

(4.6.1) must be calculated which has the following form

$$L(t_{i1},t_{i2}) = \begin{cases} S(t_{i1},t_{i2}) & \text{if} \quad \delta_{i1}=0,\delta_{i2}=0 \\ \left(-\frac{\partial}{\partial t_{i1}}\log S(t_{i1},t_{i2})\right)\times S(t_{i1},t_{i2}) & \text{if} \quad \delta_{i1}=1,\delta_{i2}=0 \\ \left(-\frac{\partial}{\partial t_{i2}}\log S(t_{i1},t_{i2})\right)\times S(t_{i1},t_{i2}) & \text{if} \quad \delta_{i1}=0,\delta_{i2}=1 \\ \left(\frac{\partial}{\partial t_{i2}\partial t_{i1}}S(t_{i1},t_{i2})\right) & \text{if} \quad \delta_{i1}=1,\delta_{i2}=1. \end{cases}$$

The last term of the likelihood can be written in logarithmic form

$$\frac{\partial}{\partial t_{i2}\partial t_{i1}}S(t_{i1},t_{i2}) = \left[\frac{\partial}{\partial t_{i2}\partial t_{i1}}\log S(t_{i1},t_{i2}) + \left(\frac{\partial}{\partial t_{i1}}\log S(t_{i1},t_{i2})\right)\left(\frac{\partial}{\partial t_{i2}}\log S(t_{i1},t_{i2})\right)\right] S(t_{i1},t_{i2}).$$

Hence, The log-likelihood function can be written in terms of the bivariate survival function as

$$\begin{aligned} &\delta_{i1}\delta_{i2}\log\left[\frac{\partial}{\partial t_{i2}\partial t_{i1}}\log S(t_{i1},t_{i2}) + \left(\frac{\partial}{\partial t_{i1}}\log S(t_{i1},t_{i2})\right)\left(\frac{\partial}{\partial t_{i2}}\log S(t_{i1},t_{i2})\right)\right] \\ &\qquad + \textstyle\sum_{j=1}^{2}\delta_{ij}\log\left(-\frac{\partial}{\partial t_{ij}}\log S(t_{i1},t_{i2})\right) + \log[S(t_{i1},t_{i2})]. \end{aligned} \tag{4.6.2}$$

Presenting the log-likelihood in terms of logarithm of the survival function simplifies the calculations of the log-likelihood which is equivalent to the one given by Wienke (2007, 2010a). The partial derivatives of the $\ln S(t_{i1},t_{i2})$ are available in Appendix B. An extensions of the above bivariate correlated frailty are given by Yashin and Iachine (1999). They used the same argument as the bivariate model by defining the frailties as $Z_j = \alpha_j(Y_0+Y_j), j=1,...,k$. The Joint survival function has the following form

$$S(t_{i1},\cdots,t_{ik}) = \left(\sum_{j}^{k} S_j(t_{ij})^{-\sigma_j^2} - n + 1\right)\prod_{j}^{k} S_j(t_{ij})^{1-\sigma_j^2(\rho_{jh}/\sigma_h)}, \tag{4.6.3}$$

where ρ_{jh} are the correlation coefficients between Z_j and Z_h, and σ_j^2 is the variance of Z_j, $i, j = 1, ..., k; j \neq k$. This model is restricted since $\rho_{jh}/\sigma_j\sigma_h$ is assumed to be constant and does not depend on j and k and should satisfies the following constrain

$$\frac{\rho_{jh}}{\sigma_j\sigma_h} \leq \min\left\{\frac{1}{\sigma_j}, j = 1, ..., k\right\},$$

which may be too restrictive for real applications.

4.7 Correlated Inverse Gaussian frailty model

In this section, a general correlated Inverse Gaussian frailty with different variances is proposed. Many authors have studied the univariate Inverse Gaussian frailty model. (Hougaard 1984,Manton et al. 1986, Whitmore and Lee 1991, and Klein et al. 1992a). Although it has a closed form of the Laplace transformation, the correlated Inverse Gaussian frailty is rarely considered in the literature since it doesn't have the reproductivity property (i.e. the summation of Inverse Gaussian is not an Inverse Gaussian). Kheiri et al. (2007) suggested a Bayesian analysis of a correlated Inverse Gaussian frailty with common variance. Wienke et al. (2010b) extended the compound Poisson frailty model to a bivariate model where the correlated Gamma frailty model and the correlated inverse Gaussian frailty model by Kheiri et al. (2007) as special cases.

In this section, a general correlated Inverse Gaussian frailty with different variances is proposed. Let $X_1, ..., X_k$ be independent Inverse Gaussian random variables with $X_i \sim IG(c_i, c_i^2)$. The mean and the variance are equal to $(c_i > 0)$. Define $Y = \sum_{i=1}^{k} X_i$, such that $Y \sim IG(\sum c_i, (\sum c_i)^2)$. To derive the formulation of the correlated Inverse Gaussian frailty, most of researchers start with a bivariate model then say it is straightforward to generalise

it to the multivariate model. To derive a multivariate frailty model with general variance-covariance matrix, a trivariate model is used to get the general form of the multivariate model. In the first step, define the following system of equations (frailties):

$$\begin{aligned} Z_1 &= a_1(X_1 + X_2 + X_4), \quad a_1 = 1/(c_1 + c_2 + c_4) \\ Z_2 &= a_2(X_1 + X_3 + X_5), \quad a_2 = 1/(c_1 + c_3 + c_5) \\ Z_3 &= a_3(X_2 + X_3 + X_6), \quad a_3 = 1/(c_2 + c_3 + c_6). \end{aligned} \tag{4.7.1}$$

It can be shown that $\boldsymbol{E}[Z_i] = 1$, $\boldsymbol{V}[Z_i] = a_i = \tau_i^2$, and the variance-covariance matrix of $\boldsymbol{Z}$ is given by

$$\boldsymbol{\Sigma} = Cov(\boldsymbol{Z}) = \begin{pmatrix} a_1 & a_1a_2c_1 & a_1a_3c_2 \\ a_1a_2c_1 & a_2 & a_2a_3c_3 \\ a_1a_3c_2 & a_2a_3c_3 & a_3 \end{pmatrix} \tag{4.7.2}$$

The number of variables usually needed to define k frailty variables is $\binom{k}{2} + k$. For covariances $\binom{k}{2}$ variables are needed and k variables for the variances. A very important assumption here is that given frailties $Z_i, i = 1, 2, 3$ the survival times $T_i, i = 1, 2, 3$ are conditionally independent. Hence the unconditional survival function can be calculated by

$$\begin{aligned} S(t_{i1}, t_{i2}, t_{i3}) &= \boldsymbol{E}[S(t_{i1}, t_{i3}, t_{i3}|Z_1, Z_2, Z_3)] \\ &= \textstyle\prod_{j=1}^{3} \boldsymbol{E}[S_j(t_{ij}|Z_j)] \\ &= \textstyle\prod_{j=1}^{3} \boldsymbol{E}[\exp(-Z_j H_{0j}(t_{ij})] \\ &= \boldsymbol{E}[\exp(-X_1[a_1H_{01}(t_{i1}) + a_2H_{02}(t_{i2})] - X_2[a_1H_{01}(t_{i1}) + a_3H_{03}(t_{i3})] - \\ &\quad X_3[a_2H_{02}(t_{i2}) + a_3H_{03}(t_{i3})] - X_4[a_1H_{01}(t_{i1})] - X_5[a_2H_{02}(t_{i2})] - \\ &\quad X_6[a_3H_{03}(t_{i3})])]. \end{aligned}$$

Using the fact that the random variables $X_1, ..., X_6$ are independent, the unconditional survival function is given by

$$S(t_{i1}, t_{i2}, t_{i3}) = \mathcal{L}_1[a_1 H_{01}(t_{i1}) + a_2 H_{02}(t_{i2})]\mathcal{L}_2[a_1 H_{01}(t_{i1}) + a_3 H_{03}(t_{i3})]$$
$$\mathcal{L}_3[a_2 H_{02}(t_{i2}) + a_3 H_{03}(t_{i3})]\mathcal{L}_4[a_1 H_{01}(t_{i1})]\mathcal{L}_5[a_2 H_{02}(t_{i2})]\mathcal{L}_6[a_3 H_{03}(t_{i3})].$$

The unconditional survival function can be expressed using the marginal survival functions as follow

$$\begin{aligned}
&S(t_{i1}, t_{i2}, t_{i3}) = \\
&[S_1(t_{i1})]^{(1-\rho_{12}\frac{\tau_1}{\tau_2}-\rho_{13}\frac{\tau_1}{\tau_3})}[S_2(t_{i2})]^{(1-\rho_{12}\frac{\tau_2}{\tau_1}-\rho_{23}\frac{\tau_2}{\tau_3})}[S_1(t_{i3})]^{(1-\rho_{13}\frac{\tau_3}{\tau_1}-\rho_{23}\frac{\tau_3}{\tau_2})} \\
&\times\exp\left\{\frac{\rho_{12}}{\tau_1\tau_2}\left(1-\left[(1-\tau_1^2\log S_1(t_{i1}))^2+(1-\tau_2^2\log S_2(t_{i2}))^2-1\right]^{1/2}\right)\right\} \\
&\times\exp\left\{\frac{\rho_{13}}{\tau_1\tau_3}\left(1-\left[(1-\tau_1^2\log S_1(t_{i1}))^2+(1-\tau_3^2\log S_3(t_{i3}))^2-1\right]^{1/2}\right)\right\} \\
&\times\exp\left\{\frac{\rho_{23}}{\tau_2\tau_3}\left(1-\left[(1-\tau_2^2\log S_2(t_{i2}))^2+(1-\tau_3^2\log S_3(t_{i3}))^2-1\right]^{1/2}\right)\right\}.
\end{aligned} \tag{4.7.3}$$

Since the frailties are defined by a system of non-negative random variables, the correlation coefficients between the frailties are *positive* and they must satisfy the following conditions.

$$\rho_{12}\frac{\tau_1}{\tau_2}+\rho_{13}\frac{\tau_1}{\tau_3}<1, \quad \rho_{12}\frac{\tau_2}{\tau_1}+\rho_{23}\frac{\tau_2}{\tau_3}<1 \quad \text{and} \quad \rho_{13}\frac{\tau_3}{\tau_1}+\rho_{23}\frac{\tau_3}{\tau_2}<1.$$

For more detail about the derivation of the unconditional survival function see appendix B. The following argument is used to derive the likelihood function. It has different components depending on the number of failures. Using the general form of the likelihood as in (2.5.1)

and using the relation in (2.1.2) one can write the likelihood as

$$L(t_{i1}, t_{i2}, t_{i3}) = \begin{cases} S(t_{i1}, t_{i2}, t_{i3}) & \text{if all } \delta_{ij} = 0, j = 1, 2, 3 \\ \left(-\frac{\partial}{\partial t_{ij}} S(t_{i1}, t_{i2}, t_{i3})\right)^{\delta_{ij}} & \text{if only one of } \delta_{ij} = 1, j = 1, 2, 3 \\ \left(\frac{\partial}{\partial t_{ij} t_{ik}} S(t_{i1}, t_{i2}, t_{i3})\right)^{\delta_{ij}\delta_{ik}} & \text{if two of } \delta_{ij} = 1, \overset{(j,k)}{(j \neq k)} = 1, 2, 3 \\ \left(-\frac{\partial^3}{\partial t_{i1} \partial t_{i2} \partial t_{i3}} S(t_{i1}, t_{i2}, t_{i3})\right) & \text{if all } \delta_{ij} = 1, j = 1, 2, 3. \end{cases} \tag{4.7.4}$$

In the case of competing risks models the individual faces only one type of the failures which has the minimum failure time (i.e. only one of $\delta_{ij} = 1, j = 1, 2, 3$), see Table 4.1 . Another possibility is that the individual does not face any of the failure types which means censored from all failure types (i.e. all $\delta_{ij} = 0, j = 1, 2, 3$). In this case only the first order of the partial derivatives of the likelihood function is needed. Consequently, (4.7.4) reduces to

$$L(t_{i1}, t_{i2}, t_{i3}) = \prod_{j=1}^{3} \left(-\frac{\partial}{\partial t_{ij}} \log S(t_{i1}, t_{i2}, t_{i3})\right)^{\delta_{ij}} S(t_{i1}, t_{i2}, t_{i3}).$$

This comes from the fact that

$$\frac{\partial}{\partial t_{ij}} S(t_{i1}, t_{i2}, t_{i3}) = \left(-\frac{\partial}{\partial t_{ij}} \log S(t_{i1}, t_{i2}, t_{i3})\right) S(t_{i1}, t_{i2}, t_{i3}).$$

The log-likelihood is given by

$$\ell(t_{i1}, t_{i2}, t_{i3}) = \sum_{j=1}^{3} \delta_{ij} \log \left(-\frac{\partial}{\partial t_{ij}} \log S(t_{i1}, t_{i2}, t_{i3})\right) + \log[S(t_{i1}, t_{i2}, t_{i3})]. \tag{4.7.5}$$

In the case of competing risks, only the minimum failure time is observed and t_{ij} is replaced by the time to the first recurrence, $t_i = min(t_{i1}, t_{i3}, t_{i3})$.

4.8 Multivariate Inverse Gaussian frailty model

4.8.1 Inverse Gaussian frailty model

In this section, a general multivariate Inverse Gaussian frailty model where the joint distribution of the frailty vector is a multivariate Inverse Gaussian is proposed. The additive model in the previous section has two restrictions. First, it generates a correlated frailty model whose marginal distributions are Inverse Gaussian variables but not a multivariate Inverse Gaussian distribution. Second, it produces restricted correlation coefficients between frailties. In this section, these restrictions are relaxed and a multivariate Inverse Gaussian frailty with a general correlation structure between the frailties is presented. Minami (2003) proposed a multivariate Inverse Gaussian distribution based on the inverse relationship with the multivariate Normal distribution. The proposed distribution has three sets of parameters $\boldsymbol{\beta}, \boldsymbol{\mu}$ and Ω denoted by $\mathbf{Z}_{k\times 1} \sim MIG(\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ define on $\left\{\mathbf{z} : \boldsymbol{\beta}'\mathbf{z} > 0, \mathbf{z} \in \mathbb{R}^k\right\}$. The p.d.f is given by

$$f(\mathbf{z}; \boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\Omega}) = (2\pi)^{-k/2}\boldsymbol{\beta}'\boldsymbol{\mu}\,|\Omega|^{-1/2}\,(\boldsymbol{\beta}'\mathbf{z})^{-(k/2+1)} \times \exp\left\{-\frac{1}{2\boldsymbol{\beta}'\mathbf{z}}(\mathbf{z}-\mu)'\Omega^{-1}(\mathbf{z}-\mu)\right\}. \quad (4.8.1)$$

where $\boldsymbol{\beta}, \boldsymbol{\mu} \in \mathbb{R}^k$, $\boldsymbol{\beta}'\boldsymbol{\mu} > 0$ and Ω is a symmetric positive definite matrix of size $k \times k$. The mean and the covariance matrix are given by

$$\boldsymbol{E}[\boldsymbol{Z}] = \boldsymbol{\mu} \quad \text{and} \quad \boldsymbol{V}[Z] = \boldsymbol{\beta}'\boldsymbol{\mu}\boldsymbol{\Omega}.$$

The cumulant generating function (c.g.f) of the distribution is given by

$$\Phi_{MI}(\boldsymbol{t}) = \ln(\boldsymbol{E}[\exp(-\boldsymbol{Z}\boldsymbol{t})]) = -\boldsymbol{\mu}^{'}(\boldsymbol{t} - b\boldsymbol{\beta}), \tag{4.8.2}$$

where

$$b = \frac{1}{\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}} \left\{ 1 + \boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{t} - \sqrt{(1+\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{t})^2 - \boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}\ \ \boldsymbol{t}'\boldsymbol{\Omega}\boldsymbol{t}} \right\}.$$

For identifiability purpose the frailty distribution is assumed to have a mean vector of ones $\boldsymbol{\mu} = \mathbf{1}$ and $\boldsymbol{\beta} = \mathbf{1}$, $\mathbf{Z} \sim MIG(\mathbf{1}, \mathbf{1}, \boldsymbol{\Omega})$. Using the result of (4.8.2) the unconditional survival function is given by

$$S(\boldsymbol{t}) = \exp\left[-\mathbf{1}^{'}\boldsymbol{H_0}(\boldsymbol{t}) + \frac{k}{\mathbf{1}'\boldsymbol{\Omega}\mathbf{1}} \left\{ 1 + \mathbf{1}^{'}\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t}) - \sqrt{(1+\mathbf{1}'\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t}))^2 - \mathbf{1}'\boldsymbol{\Omega}\mathbf{1}\ \ \boldsymbol{H_0}(\boldsymbol{t})'\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t})} \right\} \right]. \tag{4.8.3}$$

where $\boldsymbol{H_0(t)} = (H_{01}(t_{i1}), ..., H_{0k}(t_{ik}))^{'}$ is the cumulative baseline hazard. This is more flexible than the previous model with no restriction on the correlation coefficients. The same argument as in (4.7.4) could be used to derive the likelihood function. The first order partial derivative of the log-survival function with respect to one of the survival times say t_r is given by

$$\frac{\partial ln S(\boldsymbol{t})}{\partial t_r} = -\mathbf{1}^{'}\boldsymbol{h_{0r}}(\boldsymbol{t}) + \frac{k}{\mathbf{1}'\boldsymbol{\Omega}\mathbf{1}} \times \left\{ \mathbf{1}^{'}\boldsymbol{\Omega}\boldsymbol{h_{0r}}(\boldsymbol{t}) - \frac{\mathbf{1}^{'}\boldsymbol{\Omega}\boldsymbol{h_{0r}}(\boldsymbol{t})(1+\mathbf{1}^{'}\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t})) - \mathbf{1}'\boldsymbol{\Omega}\mathbf{1}\boldsymbol{h}'_{\boldsymbol{0r}}(\boldsymbol{t})\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t})}{\sqrt{(1+\mathbf{1}'\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t}))^2 - \mathbf{1}'\boldsymbol{\Omega}\mathbf{1}\boldsymbol{H_0}(\boldsymbol{t})'\boldsymbol{\Omega}\boldsymbol{H_0}(\boldsymbol{t})}} \right\}. \tag{4.8.4}$$

where $\boldsymbol{h_{0r}(t)} = (0, ..., h_{0r}(t_r), ..., 0)^{'}$ and k is number of individuals in the cluster or number

repeated measure of an individual. The general likelihood of the multivariate survival data either for cluster or repeated measures involves k orders of the partial derivatives of $-\log S(\boldsymbol{t})$. In the case of competing risks only the first order of partial derivatives with respect to $t_{i1}, ..., t_{ik}$ are required. Hence the log-likelihood function of competing risk model with k possible failures is given by

$$\ell(t_{i1}, ..., t_{ik}) = \sum_{j=1}^{k} \delta_{ij} \log \left(-\frac{\partial}{\partial t_{ij}} \log S(t_{i1}, ..., t_{ik}) \right) + \log[S(t_{i1}, ..., t_{ik})]. \tag{4.8.5}$$

4.9 Multivariate Log-Normal frailty model

The Laplace transformation of the Log-Normal variable does not have a closed-form expression. Hence it doesnt follow the same methodology as the Gamma or the Inverse Gaussian distributions to derive the unconditional survival function. The likelihood could still be maximised directly using numerical integration methods. However, The Log-Normal frailty model gives more flexibility in the multivariate case with general variance-covariance matrix. A number of Authors have discussed the correlated Log-Normal frailty (Xue and Brookmeyer 1996, Ripatti and Palmgren 2000 and Pankratz et al. 2005). For the Bayesian analysis of multivariate Log-Normal frailty models see, (Locatelli et al. 2004 and Stefanescu and Turnbull 2006).

4.9.1 Cholesky decomposition

This section describes the use of Cholesky decomposition in analysing competing risks model for recurrence of the breast cancer. Given a symmetric positive definite matrix Σ, the Cholesky decomposition creates a lower triangular matrix $\boldsymbol{L}$ with strictly positive diagonal entries such that $\Sigma = \boldsymbol{L}\boldsymbol{L}'$ or equivalently crates an upper triangle matrix $\boldsymbol{D}$ such that

$\Sigma = \boldsymbol{D}'\boldsymbol{D}$. Fahrmeir and Tutz (1994) and Oskrochi and Davies (1997) implemented the Cholesky decomposition for multivariate frailty models. First, define a multivariate Normal distribution of three variables such that the vector $\boldsymbol{U} \sim MVN(\mathbf{0}, \Sigma)$ is a trivariate Normal distribution

$$\begin{pmatrix} U_{i1} \\ U_{i2} \\ U_{i3} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{pmatrix}\right)$$

Let $Z_i = exp(U_i)$, then (Z_1, Z_2, Z_3) have a trivariate Log-Normal

$$\begin{pmatrix} Z_{i1} \\ Z_{i2} \\ Z_{i3} \end{pmatrix} \sim LogN\left(\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \tau_{12} & \tau_{13} \\ \tau_{21} & \tau_2^2 & \tau_{23} \\ \tau_{31} & \tau_{32} & \tau_3^2 \end{pmatrix}\right)$$

The mean, variance and the correlation of the frailties Z_{ij} are

$$\begin{aligned} \mu_j &= E[Z_{ij}] = E[\exp(U_{ij})] = e^{0.5\sigma_j^2} \\ \tau_j^2 &= V[Z_{ij}] = e^{\sigma_j^2}(e^{\sigma_j^2} - 1) \\ \tau_{jk} &= Cov(Z_{ij}, Z_{ik}) = e^{0.5(\sigma_1^2+\sigma_2^2)}(e^{\sigma_{12}} - 1) \end{aligned}$$

The unconditional likelihood function

$$L(t_{i1}, t_{i2}, t_{i3}) = \prod_i^n \iiint_{R^+} \prod_j^3 \left[(z_{ij} h_{0j}(t_{ij}) e^{\boldsymbol{x}_{ij}'\boldsymbol{\beta}_j})^{d_{ij}} \exp(-z_{ij} H_{0j}(t_{ij}) e^{\boldsymbol{x}_{ij}'\boldsymbol{\beta}_j})\right] \times f_{Z_j}(z_{i1}, z_{i2}, z_{i3}) dz_{i1} dz_{i2} dz_{i3},$$

or equivalently,

$$L(t_{i1},t_{i2},t_{i3}) = \prod_i^n \iiint_R \prod_j^3 \left[(h_{0j}(t_{ij})e^{\boldsymbol{x_{ij}'\beta_j}+u_{ij}})^{d_{ij}} \exp(-H_{0j}(t_{ij})e^{\boldsymbol{x_{ij}'\beta_j}+u_{ij}})\right] \times f_{U_j}(u_{i1},u_{i2},u_{i3})du_{i1}du_{i2}du_{i3}. \tag{4.9.1}$$

Using the Cholesky decomposition, the multivariate Normal random vector $\boldsymbol{U} \sim MVN(\boldsymbol{0},\Sigma)$ can be written as $\boldsymbol{U} = \boldsymbol{LU^*}$, where $\boldsymbol{U^*} \sim MVN(\boldsymbol{0},\boldsymbol{I})$ and $\boldsymbol{L}$ is the lower Cholesky triangle

$$L = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}$$

The covariance matrix of $\boldsymbol{U}$ is $\Sigma = \boldsymbol{LIL'}$.

$$U = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \times \begin{pmatrix} U_1^* \\ U_2^* \\ U_3^* \end{pmatrix} = \begin{pmatrix} l_{11}U_1^* \\ l_{21}U_1^* + l_{22}U_2^* \\ l_{31}U_1^* + l_{32}U_2^* + l_{33}U_3^* \end{pmatrix} \tag{4.9.2}$$

Hence, the random variable u_{ij} in model (4.9.1) are replaced by their corresponding values in (4.9.2). Since the joint distribution of $\boldsymbol{U^*}$ is a multivariate standard Normal distribution, then it can be written as the product of its marginals.

$$L(t_{i1},t_{i2},t_{i3}) = \prod_i^n \iiint_R \prod_j^3 \left[(h_{0j}(t_{ij})e^{\boldsymbol{x_{ij}'\beta_j}+[LU^*]_{ij}})^{d_{ij}} \exp(-H_{0j}(t_{ij})e^{\boldsymbol{x_{ij}'\beta_j}+[LU^*]_{ij}})\right] \times \prod_j^3 f_{U_j^*}(u_{ij}^*)du_{i1}^*du_{i2}^*du_{i3}^*.$$

where $[LU^*]_{ij}$ is the j^{th} row of the vector LU^*. To evaluate the triple integrals the Gaussian quadrature is used and then replacing the vector of the standard Normal variable $(u_{i1}^*,u_{i2}^*,u_{i3}^*)'$ by quadrature mass points $(y_{i1}^*,y_{i2}^*,y_{i3}^*)'$ with quadrature weights $(w_{i1}^*,w_{i2}^*,w_{i3}^*)'$,

see section (3.3.3). The unconditional likelihood is given by

$$L=\prod_{i}^{n}\left(\sum_{m_3}\sum_{m_2}\sum_{m_1}\left[\prod_{j}^{3}\left(h_{0j}(t_{ij})e^{\boldsymbol{x_{ij}'\beta_j}+[Ly^*]_j}\right)^{d_{ij}}\exp\left(-H_{0j}(t_{ij})e^{\boldsymbol{x_{ij}'\beta_j}+[Ly^*]_j}\right)\right]w_{i1}^*w_{i2}^*w_{i3}^*\right). \tag{4.9.3}$$

where m_1, m_2 and m_3 are the number of quadrature points of y_{i1}^*, y_{i2}^* and y_{i3}^* respectively. The likelihood function in (4.8.3) contains nested loops which may increase the time needed to obtain the optimal solution of the model. Most software used to get the maximum likelihood estimates are matrix oriented and working with matrices is much faster than loops. For example, Gauss software has an element by element product procedure which can be used to replace the loop by a vector. The summation in (4.8.3) is over the quadrature points and the total number of iterations needed to get a single outcome is $m_1 \times m_2 \times m_3$. One can transform the loops into vectors by creating the following vectors

$$\begin{aligned}
\boldsymbol{v_1} &= \mathbf{1}_{m_2m_3\times 1}\otimes \boldsymbol{y}_1^* & \boldsymbol{\omega_1} &= \mathbf{1}_{m_2m_3\times 1}\otimes \boldsymbol{w}_1^*\\
\boldsymbol{v_2} &= \mathbf{1}_{m_3\times 1}\otimes (\boldsymbol{y}_2^*\otimes \mathbf{1}_{m_1\times 1}) & \boldsymbol{\omega_2} &= \mathbf{1}_{m_3\times 1}\otimes (\boldsymbol{w}_2^*\otimes \mathbf{1}_{m_1\times 1})\\
\boldsymbol{v_3} &= \boldsymbol{y}_3^*\otimes \mathbf{1}_{m_1m_2\times 1} & \boldsymbol{\omega_3} &= \boldsymbol{w}_3^*\otimes \mathbf{1}_{m_1m_2\times 1},
\end{aligned}$$

where $\otimes$ is the Kronecker product. This procedure showed a significant decrease in time to reach the optimal solution of the model (i.e. the maximum likelihood estimators of the model parameters). Converting the nested loops to vectors decreases the time needed to fit the model by 50%. For the breast cancer data, fitting the model with five quadrature points took ten days while using vectors only five days are required to fit the model. Further reduction in time is by using the proposed non-parametric frailty model given in the next section which decreases the time needed to fit the model by more than 80% of the nested loops. Around two days are needed to fit the model of the breast cancer data using the purposed non-parametric frailty.

4.9.2 Weibull competing risks with Log-Normal frailty model

Assume that the failure times of the i^{th} individual of j^{th} failure (risk) have Weibull distribution ($T_{ij} \sim Weib(\lambda_j, \alpha_j)$), $i = 1, .., n$,$j = 1, .., k$. The baseline hazards are $h_{ij}(t_i) = \alpha_j t_i^{\alpha_j - 1}$, where $t_i = \min_j(t_{ij})$. The likelihood of the trivariate competing risks model is given by

$$L = \prod_i^n \left(\sum_{m_3} \sum_{m_2} \sum_{m_1} \left[\prod_j^3 \left(\alpha_j t_i^{\alpha_j - 1} e^{\boldsymbol{x'_{ij}}\boldsymbol{\beta_j} + [Ly^*]_j} \right)^{d_{ij}} \exp\left(-t_i^{\alpha_j} e^{\boldsymbol{x'_{ij}}\boldsymbol{\beta_j} + [Ly^*]_j} \right) \right] w_{i1}^* w_{i2}^* w_{i3}^* \right).$$

4.10 Competing risks with non-parametric frailty model

In the previous chapter, it was shown that the model estimates are not robust against the mis-specifying of the frailty distribution. In multivariate frailty models, the choice of the frailty distribution is crucial to obtain correct estimates of the dependence structure (Duchateau and Janssen, 2008). This section makes use of the results of the previous section of Cholesky decomposition and the non-parametric frailty in section (3.3.4). The frailty variable is assumed to follow some distribution say $g(\boldsymbol{z})$ with a variance-covariance matrix $\Sigma = \boldsymbol{QIQ'}$ where $\boldsymbol{Q}$ is lower triangle of the Cholesky decomposition.

$$Z = \begin{pmatrix} q_{11} & 0 & 0 \\ q_{21} & q_{22} & 0 \\ q_{31} & q_{32} & q_{33} \end{pmatrix} \times \begin{pmatrix} Z_1^* \\ Z_2^* \\ Z_3^* \end{pmatrix} = \begin{pmatrix} q_{11} Z_1^* \\ q_{21} Z_1^* + q_{22} Z_2^* \\ q_{31} Z_1^* + q_{32} Z_2^* + q_{33} Z_3^* \end{pmatrix}$$

$Z = QZ^*$ where Z^* has some p.d.f, $g^*(\boldsymbol{z^*})$ and the variance-covariance matrix is the identity matrix $\boldsymbol{I}$. There are two differences between this model and the Log-Normal model. First, the identity matrix of the variance-covariance of Z^* does not necessary imply independence

and hence the joint density can not be written as the product of the marginal distributions. Second, the diagonal of the variance-covariance matrix of Z will be absorbed in the non-parametric representation and the frailty variable QZ^* will be replaced by the quadrature mass points.

$$L(t_{i1}, t_{i2}, t_{i3}) = \prod_i^n \iiint_R \prod_j^3 \left[(h_{0j}(t_{ij}) e^{\boldsymbol{x'_{ij}}\boldsymbol{\beta_j} + [QZ^*]_j})^{d_{ij}} \exp(-H_{0j}(t_{ij}) e^{\boldsymbol{x'_{ij}}\boldsymbol{\beta_j} + [QZ^*]_j}) \right] \times f_{Z_j^*}(z_{i1}^*, z_{i2}^*, z_{i3}^*) dz_{i1}^* dz_{i2}^* dz_{i3}^*$$

$$\begin{pmatrix} q_{11} Z_1^* \\ q_{21} Z_1^* + q_{22} Z_2^* \\ q_{31} Z_1^* + q_{32} Z_2^* + q_{33} Z_3^* \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ r_{21}\gamma_1 + \gamma_2 \\ r_{31}\gamma_1 + r_{32}\gamma_2 + \gamma_3 \end{pmatrix} = \boldsymbol{\gamma}.$$

Hence, the likelihood function is given by

$$L = \prod_i^n \left(\sum_{m_3} \sum_{m_2} \sum_{m_1} \left[\prod_j^3 \left(\alpha_j t_i^{\alpha_j - 1} e^{\boldsymbol{x'_{ij}}\boldsymbol{\beta_j} + [\gamma]_{ij}} \right)^{d_{ij}} \exp\left(-t_i^{\alpha_j} e^{\boldsymbol{x'_{ij}}\boldsymbol{\beta_j} + [\gamma]_{ij}} \right) \right] \pi_{i1} \pi_{i2} \pi_{i3} \right).$$

where $[\gamma]_{ij}$ is the j^{th} row of the vector $\boldsymbol{\gamma}$ and $(\gamma_1, .., \gamma_3,\ \pi_{i1}, ..., \pi_{i3})$ are also vector of the quadrature points and weights respectively. The terms $(q_{ij}, i = j)$ are absorbed in the non-parametric representation and $(q_{ij}, i \neq j)$ are replaced by r_{ij}. The relation between $(q_{ij}, i \neq j)$ and r_{ij} can be found algebraically. For example, $r_{21} = (q_{21}/q_{11})(\gamma_1 - \beta_{10})$. Note that r_{ij} are not the correlation coefficients between the frailties, but they were added to the model to account for the association between frailties. In non-parametric analysis, the main interest is to fit the regression coefficients not these associations. Here the quadrature mass points and their corresponding weights are unknown and need to be estimated. In this model, there are two assumptions. First, the marginal distributions of the frailty are identical and can be estimated by the same vector of quadrature mass points, $\gamma_1 = \gamma_2 = \gamma_3$ and same weights, $\pi_{i1} = \pi_{i2} = \pi_{i3}$. For example, if three quadrature points are used, then the number

of parameters needed to approximate the integrations is six parameters, three mass points and three weights.

Second, using the Cholesky decomposition generates independence between the frailties. But these two assumptions can be relaxed. First, one can assume different mass points and different weights but independent marginal. In this case the total number of parameters needed to approximate the integrations is eighteen parameters, nine mass points and nine weights. Second, if the Cholesky decomposition does not generate independence between the frailties, the weights at each pair of mass point cannot be written as the product of the corresponding weights. In this case each combination of three points need a weight. The number of parameters needed is nine mass points and one twenty-seven weights. Fortunately, simulations showed that if there is no restriction on these parameters, the total number of parameters needed to get acceptable result decreases.

4.11 Multivariate simulations

This is to test the performance of the above proposed models through simulated data. Four simulation studies are conducted. First, a bivariate competing risks model with two failure times following a Weibull distribution with same shape parameter α and different scale parameters λ. The bivariate frailty is assumed to follow Inverse Gaussian distribution. Second, same as the first data but the frailty is assumed to follow Log-Normal distribution fitted using Cholesky decomposition. Third, an extension of the previous data by adding a third failure type, a competing risks model with multivariate frailty. Fourth, a bivariate competing risks model is generated with a Weibull baseline hazard and three different frailty distributions, Log-Normal, Gamma, and Inverse Gaussian and fitted by non-parametric frailty.

4.11.1 Bivariate Inverse Gaussian frailty of competing risks model

This is to check the proposed bivariate correlated Inverse Gaussian frailty model given in section 4.7. A bivariate competing risks model with two failure times following a Weibull distribution is generated, $T_1 \sim Weib(\lambda_1, \alpha)$ and $T_2 \sim Weib(\lambda_2, \alpha)$. Censoring times are assumed to be independent and follow a Weibull distribution $C \sim Weib(\theta, \alpha)$. The frailties Z_{ij} are assumed to follow a correlated bivariate Inverse Gaussian distribution with mean one and variance τ_i^2 generated using the argument given in section 4.7. The failure times were generated as $T_{ij} = (-\log(u_{ij})/\lambda_{ij})^{1/\alpha}$, where $\lambda_{ij} = Z_{ij}\exp(\boldsymbol{x_i'\beta_j})$ and $u_{ij} \sim Uni(0,1)$, $j = 1, 2$. Two different sets of parameters were used to check the model estimation. Three different types of explanatory variables were generated, X_{i1} is a continuous random variable from Uniform distribution, X_{i2} is a dichotomous, and X_{i3} is a qualitative variable with three categories which was converted to two dummy variable $X_{i3,1}$ and $X_{i3,2}$. See section 3.5. The regression model is constructed using the following predictors,

$$\boldsymbol{x_i'\beta_1} = \beta_{10} + x_{i1}\beta_{11} + x_{i2}\beta_{12} + x_{i3,1}\beta_{13} + x_{i3,2}\beta_{14}.$$

$$\boldsymbol{x_i'\beta_2} = \beta_{20} + x_{i1}\beta_{21} + x_{i2}\beta_{22} + x_{i3,1}\beta_{23} + x_{i3,2}\beta_{24}.$$

The censoring times were generated as $C_i = (-\log(u_i)/\theta)^{1/\alpha}$, where $u_i \sim Uni(0,1)$ and finally the survival times are $Y_i = \min(T_{i1}, T_{i2}, C_i)$. Table 4.2 shows the simulated data of two sets of parameters. The censoring rate is set to 20% and failure rate to 40% for each of the two failure types to get a representative sample of each group. For each set of parameters, 500 data sets were generated each with sample sizes of 1000 and 5000. Using sample size of 1000 instead of 500 like other simulation is due to the difficulty of getting the maximum likelihood estimation of the Inverse Gaussian distribution. Different values of α, τ^2, ρ, and different regression coefficients were used. The simulated data showed high levels of accuracy

in retrieving the true values of model parameters with small standard errors particularly when large sample size is used. The estimation method was capable to accommodate both weak and strong correlation between frailties with small and large variances. The parameters used in the first data set are, $\tau_1^2 = 0.8,\ \tau_2^2 = 1.25,\ \rho = 0.3$ and in the second data set are, $\tau_1^2 = \tau_2^2 = 1,\ \rho = 0.8$.

Parameter	True values	Sample size		True values	Sample size	
		1000	5000		1000	5000
		Mean (S.e)	Mean (S.e)		Mean (S.e)	Mean (S.e)
α_1	0.5	0.510	0.501	1.0	1.028	1.003
		(0.057)	(0.021)		(0.112)	(0.041)
α_2	0.5	0.507	0.501	1.0	1.030	1.001
		(0.059)	(0.022)		(0.128)	(0.049)
τ_1^2	0.8	1.317	0.838	1.0	1.710	1.045
		(1.611)	(0.330)		(1.895)	(0.350)
τ_2^2	1.25	1.929	1.311	1.0	2.017	1.080
		(2.046)	(0.449)		(2.084)	(0.494)
ρ	0.3	0.197	0.294	0.8	0.690	0.815
		(0.689)	(0.273)		(0.511)	(0.186)
β_{10}	-4.0	-4.027	-4.003	-2.0	-1.978	-2.002
		(0.272)	(0.116)		(0.237)	(0.089)
β_{11}	9.0	9.201	9.013	6.0	6.214	6.019
		(1.023)	(0.383)		(0.698)	(0.251)
β_{12}	3.0	3.050	3.007	4.0	4.101	4.010
		(0.400)	(0.150)		(0.460)	(0.179)
β_{13}	2.0	2.053	2.003	2.0	2.066	2.005
		(0.258)	(0.102)		(0.247)	(0.093)
β_{14}	1.0	0.987	0.998	-1.0	-1.015	-1.004
		(0.339)	(0.143)		(0.198)	(0.085)
β_{20}	-3.0	-2.977	-3.006	-1.0	-0.895	-0.991
		(0.201)	(0.087)		(0.320)	(0.111)
β_{21}	7.0	7.077	7.015	2.0	1.999	1.988
		(0.888)	(0.348)		(0.668)	(0.262)
β_{22}	4.0	4.071	4.013	5.0	5.165	5.008
		(0.477)	(0.185)		(0.595)	(0.234)
β_{23}	1.0	0.996	1.003	1.0	1.020	1.003
		(0.218)	(0.091)		(0.261)	(0.109)
β_{24}	3.0	3.069	3.011	-2.0	-2.063	-2.004
		(0.386)	(0.142)		(0.338)	(0.138)

Table 4.2: Bivariate Inverse Gaussian frailty model with Weibull baseline hazard and two sets of covariates simulated data, 500 data sets each with sample sizes of 1000 and 5000.

4.11.2 Bivariate Log-Normal frailty of competing risks model

In this section, the accuracy of Cholesky decomposition in estimating the bivariate frailty models discussed in section 4.9 is checked. Similar to the previous section, a bivariate competing risks model with two failure times following a Weibull distribution is generated. The log of frailty distributions associated with each failure are assumed to be Normal with mean zero and variance σ_i^2, $W_i \sim N(0, \sigma_i^2), i = 1, 2$. For each failure type, only two types of explanatory variables were generated, X_{i1} is a continuous random variable from Uniform distribution $X_{i1} \sim Uni(0, 1)$, and X_{i2} is a dichotomous variable generated as follows

$$X_{i2} = \begin{cases} 1 & if \quad u_i < 0.3 \\ 0 & if \quad u_i \geq 0.3 \end{cases}$$

where $u_i \sim Uni(0, 1)$. The regression model was constructed using the following predictors,

$$\boldsymbol{x_i'\beta_1} = \beta_{10} + x_{i1}\beta_{11} + x_{i2}\beta_{12}, \qquad \text{and} \qquad \boldsymbol{x_i'\beta_2} = \beta_{20} + x_{i1}\beta_{21} + x_{i2}\beta_{22}.$$

The failure times were generated as $T_{ij} = (-\log(u_{ij})/\lambda_{ij})^{1/\alpha}$, where $\lambda_{ij} = \exp(\boldsymbol{x_i'\beta_j} + W_j)$ and $u_{ij} \sim Uni(0, 1)$, $j = 1, 2$. The log of the frailty distributions associated with each failure are assumed to be Normal with mean zero and variance σ_i^2. $\boldsymbol{W} \sim BVN(\boldsymbol{0}, \Sigma)$ is a bivariate Normal distribution

$$\begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right).$$

The censoring times were generated as $C_i = (-\log(u_{i3})/\theta)^{1/\alpha}$, where $u_{i3} \sim Uni(0, 1)$ and finally the survival times are $Y_i = \min(T_{i1}, T_{i2}, C_i)$. Table 4.3 shows the simulated data of two sets of parameters with censoring rate of 20%. A Gaussian quadrature integration with 32

quadrature points was used to integrate out the frailty, codes are in appendix C. For each set of parameters, 600 data sets each with sample sizes of 500 and 5000 are simulated. Different values of α, σ^2, ρ, and different regression coefficients are used. The simulated data showed high levels of accuracy in retrieving the true values of model parameters using Cholesky decomposition. In first data set, the parameter used are $\sigma_1^2 = \sigma_2^2 = 1, \rho = 0.3$ and in the second data set the parameters used are, $\sigma_1^2 = 0.7, \sigma_2^2 = 1.2, \rho = 0.8$. The frailty variances have bigger standard errors than other parameters in the model.

Parameter	True values	Sample size		True values	Sample size	
		500	5000		500	5000
		Mean (S.e)	Mean (S.e)		Mean (S.e)	Mean (S.e)
α_1	1	1.088	1.010	0.5	0.527	0.501
		(0.265)	(0.051)		(0.130)	(0.023)
α_2	1	1.069	1.007	0.5	0.610	0.506
		(0.203)	(0.047)		(0.218)	(0.029)
σ_1^2	1	1.799	1.072	0.7	0.905	0.710
		(2.506)	(0.354)		(0.618)	(0.127)
σ_2^2	1	1.668	1.058	1.2	1.738	1.236
		(1.917)	(0.326)		(1.016)	(0.164)
ρ	0.3	0.137	0.274	0.8	0.488	0.772
		(0.926)	(0.281)		(0.567)	(0.250)
β_{10}	-3	-3.224	-3.027	-0.2	-0.318	-0.205
		(0.738)	(0.161)		(0.303)	(0.090)
β_{11}	1	1.038	1.009	0.5	0.476	0.486
		(0.454)	(0.120)		(0.393)	(0.101)
β_{12}	2	2.145	2.013	1	1.050	0.997
		(0.613)	(0.134)		(0.341)	(0.077)
β_{20}	-4	-4.305	-4.030	0.2	0.201	0.202
		(0.878)	(0.216)		(0.348)	(0.076)
β_{21}	2	2.154	2.019	0.7	0.907	0.701
		(0.555)	(0.138)		(0.542)	(0.113)
β_{22}	3	3.230	3.024	1	1.223	1.017
		(0.681)	(0.156)		(0.527)	(0.086)

Table 4.3: Bivariate Log-Normal frailty model with Weibull baseline hazard and two sets of covariates simulated data, 600 data sets each with sample sizes of 500 and 5000.

4.11.3 Multivariate Log-Normal frailty of competing risks model

This section extends the pervious data by adding another failure time to have a trivariate frailty model. The purpose of this addition is to test the performance of the Choleskey decomposition of the Log-Normal frailty model in high dimensions. Three competing risks with failure times, T_1, T_2, and T_3 following the Weibull distribution are generated. Similar to bivariate simulations, for each failure type two types of explanatory variables were generated, X_{i1} is a continuous random variable and X_{i2} is a dichotomous variable. The log of the frailty distributions associated with each failure are assumed to be Normal with mean zero and variance σ_i^2. $\boldsymbol{W} \sim MVN(\mathbf{0}, \Sigma)$ is a multivariate Normal distribution

$$\begin{pmatrix} W_1 \\ W_2 \\ W_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix}\right)$$

Table 4.4 shows the simulated data of three dimensions competing risks. The failure and censoring times are generated at censoring rate of 20%. Gaussian quadrature integration with 8 quadrature points was used to integrate out the frailty. 500 data sets each with sample sizes of 1000 and 5000 are simulated. Because of high dimensionality of the model a sample size of 1000 is used instead of 500. These simulations used the following parameters: $\sigma_1^2 = 1.25, \sigma_2^2 = 0.8, \sigma_3^2 = 1.75$, and the following correlations: $\rho_{12} = -0.7, \rho_{13} = -0.1, \rho_{23} = 0.25$. In conclusion, the estimation method using the Cholesky decomposition can be used in all situations of positive and negative correlation between the log of frailties and with small and large variances. The standard errors of parameters estimates in case of sample size 5000 are smaller than those for 1000.

Parameter	True values	Sample size	
		1000	5000
		Mean(S.e)	Mean(S.e)
β_{10}	-4.0	-4.260 (0.596)	-3.992 (0.266)
β_{11}	3.0	3.199 (0.541)	2.990 (0.219)
β_{12}	-1.0	-1.034 (0.426)	-1.000 (0.189)
α_1	0.5	0.531 (0.078)	0.501 (0.030)
β_{20}	-3.0	-3.049 (0.377)	-3.011 (0.128)
β_{21}	-2.0	-2.095 (0.409)	-2.006 (0.148)
β_{22}	5.0	5.164 (0.706)	5.024 (0.206)
α_2	0.5	0.521 (0.074)	0.503 (0.023)
β_{30}	-2.0	-2.141 (0.366)	-2.008 (0.133)
β_{31}	1.0	1.095 (0.355)	1.005 (0.141)
β_{32}	-2.0	-2.111 (0.595)	-2.015 (0.253)
α_3	0.50	0.539 (0.095)	0.503 (0.029)
σ_1^2	1.25	1.754 (1.098)	1.275 (0.418)
σ_2^2	0.80	1.038 (0.875)	0.835 (0.196)
σ_3^2	1.75	2.536 (1.741)	1.846 (0.513)
ρ_{12}	-0.70	-0.647 (0.333)	-0.706 (0.216)
ρ_{13}	-0.10	-0.086 (0.455)	-0.092 (0.248)
ρ_{23}	0.25	0.247 (0.385)	0.258 (0.200)

Table 4.4: Trivariate Log-Normal frailty model with Weibull baseline hazard and two covariates simulated data, 500 data sets each with sample sizes of 1000 and 5000.

4.11.4 Bivariate non-parametric frailty of competing risks model

This section tests the ability of the non-parametric frailty model proposed in section 4.10 in capturing the model fits whatever the original frailty distribution. The simulation procedure is similar to section 4.11.2 with different frailty distributions and there were fitted non-parametrically. Assuming Weibull baseline hazard, the failure times were generated as $T_{ij} = (-\log(u_{ij})/\lambda_{ij})^{1/\alpha}$, where $\lambda_{ij} = Z_j \exp(\boldsymbol{x_i'}\boldsymbol{\beta_j})$ and $u_{ij} \sim Uni(0,1)$, $j = 1,2$ with censoring rate of 20%. The distribution of frailties Z_j are assumed to be either Log-Normal, Gamma, or Inverse Gaussian. Table 4.5 shows the simulated data of 500 data sets each with sample sizes of 500 and 5000 of Log-Normal, Gamma and Inverse Gaussian frailty with two covariates fitted non-parametrically. The constant terms were absorbed in the non-parametric representation. The component r was included in the model to capture the association between frailties. The variances and the correlation used in these simulations are $\tau_1^2 = 0.8, \tau_2^2 = 1.25, \rho = 0.3$. Only three mass points of the quadrature integration and three corresponding weights were used for each frailty distribution and it was enough to fit the models adequately whatever the original distribution. These mass points and their weight are represented by γ_{ij} and π_{ij}, $(i = 1,2; j = 1,2,3)$ respectively. To increase the flexibility of the non-parametric frailty model, the mass points of each frailty distribution are assumed to be different (i.e. γ_{1j} and γ_{2j}, $j = 1,2,3$ are different). The same applies for their corresponding weights. The standard errors are almost the same whatever the original frailty distribution.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian	
		500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_{11}	9	9.327	8.936	9.259	8.750	9.434	9.032
		(1.468)	(0.441)	(1.375)	(0.478)	(1.531)	(0.470)
β_{12}	3	3.136	3.010	3.142	2.932	3.193	3.031
		(0.574)	(0.186)	(0.521)	(0.189)	(0.619)	(0.188)
α_1	0.5	0.520	0.498	0.518	0.487	0.527	0.503
		(0.080)	(0.026)	(0.076)	(0.027)	(0.086)	(0.027)
β_{21}	7	7.600	6.975	7.958	7.128	7.761	7.098
		(1.770)	(0.450)	(2.032)	(0.539)	(1.784)	(0.511)
β_{22}	4	4.327	3.957	4.475	4.056	4.423	4.032
		(0.971)	(0.243)	(1.104)	(0.296)	(0.975)	(0.270)
α_2	0.5	0.542	0.496	0.565	0.508	0.555	0.505
		(0.123)	(0.030)	(0.140)	(0.036)	(0.122)	(0.034)
r		1.032	0.584	1.078	0.585	0.964	0.516
		(2.014)	(0.403)	(1.542)	(0.317)	(1.402)	(0.494)
γ_{11}		-4.531	-4.149	-5.798	-5.215	-4.582	-4.767
		(6.269)	(1.423)	(7.983)	(2.130)	(1.893)	(5.957)
γ_{12}		-4.283	-4.218	-5.392	-5.180	-4.427	-4.551
		(3.863)	(2.114)	(6.445)	(2.251)	(1.608)	(3.592)
γ_{13}		-4.330	-4.333	-4.972	-5.122	-4.597	-4.686
		(4.610)	(1.759)	(2.203)	(1.874)	(2.209)	(3.961)
γ_{21}		0.009	-0.805	-0.059	-1.252	-0.064	-1.109
		(6.559)	(2.338)	(6.084)	(3.125)	(5.122)	(2.348)
γ_{22}		0.263	-0.695	-0.074	-1.470	0.198	-1.114
		(6.658)	(2.411)	(5.673)	(2.743)	(5.332)	(2.585)
γ_{23}		0.264	-0.904	-0.210	-1.545	-0.215	-1.196
		(7.680)	(2.464)	(5.989)	(2.588)	(5.941)	(2.392)
π_{11}		0.333	0.330	0.322	0.321	0.359	0.350
		(0.218)	(0.207)	(0.242)	(0.234)	(0.223)	(0.204)
π_{12}		0.348	0.338	0.342	0.336	0.320	0.309
		(0.209)	(0.212)	(0.239)	(0.239)	(0.207)	(0.195)
π_{13}		0.319	0.332	0.336	0.343	0.322	0.341
		(0.215)	(0.220)	(0.237)	(0.235)	(0.204)	(0.199)
π_{21}		0.335	0.342	0.321	0.318	0.356	0.323
		(0.245)	(0.225)	(0.259)	(0.225)	(0.252)	(0.228)
π_{22}		0.338	0.318	0.350	0.334	0.321	0.336
		(0.255)	(0.217)	(0.264)	(0.214)	(0.255)	(0.223)
π_{23}		0.327	0.340	0.330	0.348	0.323	0.341
		(0.246)	(0.219)	(0.253)	(0.215)	(0.249)	(0.232)

Table 4.5: Log-Normal, Gamma and Inverse Gaussian frailty model with Weibull baseline hazard and two covariates simulated data fitted non-parametrically, 500 data sets each with sample sizes of 500 and 5000.

4.12 Results on breast cancer recurrence data

In this section, the proposed multivariate competing risks frailty models is applied on the breast cancer data discussed in previous chapters. This includes the Weibull hazard models with multivariate Log-Normal frailty using Cholesky decomposition and the multivariate non-parametric frailty along with the univariate models. The following tables summarise the regression analysis for each model of the competing risks assuming Weibull hazard with four different frailty models. First, the independent Log-Normal frailty model. Second, a multivariate Log-Normal frailty model. Third, an independent non-parametric frailty model. Fourth, a multivariate non-parametric frailty model. The discussion of the results is based on the last model since it doesn't assume a specific frailty distribution and it has the lowest standard errors of the parameters estimates. In chapter six discusses the advantages and disadvantages of these models.

4.12.1 Analysis and conclusions

Through Table 4.6 to Table 4.10, the emphasis is on the results of the multivariate non-parametric frailty models since they have the smallest standard errors. Table 4.6 displays the multivariate non-parametric frailty model of local recurrence. It shows that there is no significant effect of *age* on the hazard of local recurrence, but the direction of the relation was as expected, i.e. young patients have a higher chance of local recurrence. Patients in *stage2* and *stage3* of the disease have higher hazard than patients in *stage1* (reference category). There is no difference in the hazard of local recurrence between patients in *stage4* and *stage1* due to the fact that none of stage4 patients has local recurrence, see appendix A Table A.1.

LOCAL RECURRENCE

Variable	Frailty distribution			
	Univarite Log-Normal	Multivariate Log-Normal	Univarite non-parametric	Multivariate non-parametric
AGE	-0.017(0.012)	*-0.026(0.011)	-0.016(0.010)	-0.017(0.010)
STAGE2	**0.808(0.288)	**0.827(0.246)	**0.664(0.227)	**0.758(0.242)
STAGE3	*0.935(0.409)	**1.196(0.368)	*0.726(0.317)	*0.758(0.343)
STAGE4	-10.895(52.30)	-13.818(425.2)	-11.720(98.65)	-11.760(88.21)
SURGTYPE1	**3.314(0.738)	**3.044(0.703)	**3.088(0.605)	**2.888(0.618)
SURGTYPE2	**3.604(0.511)	**3.936(0.471)	**3.201(0.368)	**3.327(0.378)
SURGTYPE3	**1.457(0.344)	**1.616(0.393)	**1.322(0.287)	**1.355(0.286)
SURGTYPE4	0.280(0.358)	0.374(0.392)	0.226(0.307)	0.286(0.310)
SURGTYPE5	0.353(0.876)	0.712(0.777)	0.231(0.771)	0.469(0.799)
SURGTYPE6	-0.019(0.507)	-0.094(0.462)	0.043(0.418)	-0.032(0.416)
SURGTYPE7	0.275(0.878)	0.458(0.776)	0.313(0.767)	0.338(0.772)
HIST2	-0.416(0.330)	-0.482(0.283)	-0.404(0.275)	-0.391(0.278)
HIST3	-0.469(0.379)	*-0.885(0.395)	-0.459(0.317)	-0.570(0.315)
HIST4	*-0.655(0.321)	*-0.654(0.273)	*-0.702(0.280)	*-0.651(0.276)
COHORT	0.125(0.237)	0.026(0.414)	0.154(0.199)	0.053(0.212)
CHEMO	0.280(0.237)	0.304(0.205)	0.234(0.197)	0.277(0.198)
MENO	0.258(0.323)	0.170(0.278)	0.237(0.270)	0.218(0.275)
RADIO	**-1.196(0.317)	**-1.132(0.289)	**-1.099(0.272)	**-1.061(0.273)
SIDE	-0.053(0.195)	0.012(0.400)	-0.066(0.164)	-0.043(0.186)
LN(α)	0.084(0.116)	0.142(0.098)	-0.069(0.068)	0.009(0.100)
CONSTANT	**-11.913(1.590)	**-11.159(1.255)		
-2 Log Likelihood	3628.10		3628.52	

Table 4.6: Results of breast cancer data: local recurrence. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

The hazard of local recurrence for the first three surgery types *no surgery, incision biopsy and excision biopsy* is significantly (P-value < 0.01) higher than hazard of *radical mastectomy and axillary clearance* (reference category). The hazard of the other four types of surgery has no significant difference from the *radical mastectomy and axillary clearance*. Patients with *ductal* histology (reference category) have a higher but not significant hazard of local recurrence than *lobular* and *dcis (ductal carcinoma in situ)*, while it is significantly higher than *other* histology. Patients with *radiotherapy* have a significantly lower hazard of local recurrence than those without. Other variables such as *date of primary surgery*, *chemotherapy*, *menopausal status*, and *side* of the body affected have no significant effect on the hazard of local recurrence. The log of the shape parameter of Weibull, $Ln(\alpha)$ is not significantly different from zero, which means local recurrence has a constant hazard.

Table 4.7 displays the multivariate non-parametric frailty model of regional recurrence. The results show that there is a significant inverse effect of *age* on the hazard of region recurrence. Similar to local recurrence, patients in *stage2* and *stage3* of the disease have a higher hazard of regional recurrence than patients in *stage1*. There is no difference in the hazard of local recurrence between patients in *stage4* and *stage1* due to the fact that only few patients with regional recurrence are in *stage4*. The hazard of regional recurrence for the first five surgery types *no surgery, incision biopsy, excision biopsy, simple mastectomy* and *radical mastectomy* is significantly higher than hazard of *radical mastectomy and axillary clearance*. The hazard of *wide local excision and axillary clearance*, and *surgery after neo adjuvant chemotherapy* has no significant difference from the *radical mastectomy and axillary clearance*. The hazard recurrence *ductal* histology is significantly higher than all other histology types especially *Dcis (ductal carcinoma in situ)* which has a much lower hazard than *ductal* histology. Patients with primary surgery *before 1990* have a significantly lower hazard of regional recurrence than those *after 1990*.

REGIONAL RECURRENCE

Variable	Frailty distribution			
	Univarite Log-Normal	Multivariate Log-Normal	Univarite non-parametric	Multivariate non-parametric
AGE	-0.023(0.012)	**-0.140(0.023)	**-0.047(0.012)	*-0.023(0.009)
STAGE2	**1.318(0.282)	**4.344(0.554)	**2.065(0.342)	**1.213(0.220)
STAGE3	**2.086(0.382)	**7.042(0.812)	**2.959(0.413)	**1.732(0.288)
STAGE4	0.796(0.513)	**2.327(0.589)	0.846(0.590)	0.604(0.423)
SURGTYPE1	**3.183(0.820)	**6.115(0.876)	**4.220(0.935)	**2.491(0.590)
SURGTYPE2	**3.604(0.604)	**11.029(1.129)	**5.205(0.532)	**2.910(0.410)
SURGTYPE3	**2.278(0.421)	**5.813(0.696)	**3.065(0.427)	**1.756(0.309)
SURGTYPE4	**1.437(0.383)	**3.127(0.545)	**1.837(0.402)	**1.185(0.281)
SURGTYPE5	**2.738(0.651)	**6.799(0.880)	**3.873(0.669)	**2.204(0.571)
SURGTYPE6	-0.870(0.633)	**-3.196(0.789)	-1.121(0.651)	-0.923(0.497)
SURGTYPE7	1.277(0.854)	**2.969(0.836)	0.752(0.883)	0.870(0.612)
HIST2	**-1.518(0.400)	**-3.537(0.756)	**-1.621(0.413)	**-1.136(0.318)
HIST3	**-4.151(0.894)	**-11.138(1.355)	**-5.077(0.881)	**-3.413(0.744)
HIST4	*-0.755(0.306)	**-1.620(0.355)	**-1.361(0.363)	**-0.707(0.239)
COHORT	**-0.731(0.271)	**-1.559(0.308)	*-0.644(0.271)	**-0.572(0.190)
CHEMO	-0.332(0.268)	0.438(0.324)	-0.304(0.318)	-0.016(0.239)
MENO	0.529(0.341)	0.003(0.710)	0.063(0.378)	0.236(0.258)
RADIO	-0.507(0.326)	**-1.071(0.338)	*-0.809(0.390)	-0.347(0.226)
SIDE	0.211(0.205)	*0.819(0.326)	0.125(0.231)	0.151(0.154)
LN(α)	*0.200(0.088)	**1.295(0.094)	**0.511(0.047)	0.063(0.073)
CONSTANT	**-12.995(1.321)	**-31.481(3.384)		
-2 Log Likelihood	5248.06		5227.96	

Table 4.7: Results of breast cancer data: regional recurrence. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

Other variables such as *chemotherapy, menopausal status, radiotherapy*, and *side* of the body affected have no significant effect on the hazard of regional recurrence. Similar to local recurrence, the log of the shape parameter of Weibull for regional recurrence is not significantly different from zero, which means regional recurrence has constant hazard.

Table 4.8 displays the multivariate non-parametric frailty model of metastasis. Similar to regional recurrence, there is a significant inverse effect of *age* on the hazard of metastasis. As patients move from one stage to another of the disease, the hazard of metastasis increases significantly. The hazard ratios of metastasis of *stage2*, *stage3*, and *stage4* compared to *stage1* are 4.95, 11.81, and 65.24, respectively. The hazard of metastasis of surgeries, *incision biopsy* and *radical mastectomy* is significantly higher than hazard of *radical mastectomy and axillary clearance*, meanwhile the hazard is significantly lower for *excision biopsy and wide local excision and axillary clearance* than hazard of *radical mastectomy and axillary clearance*. There is no significant difference between the metastasis hazard of *ductal* and *lobular histology*. While the hazard is significantly lower for *ductal* than the other two histology types, *Dcis (ductal carcinoma in situ)* and *other*. The hazard ratios of *Dcis* and *other* compared to Ductal are 0.019 and 0.51 respectively. Patients with primary surgery *after 1990* have significantly lower hazard of metastasis than those *before 1990*. Namely, the hazard ratio of cohort surgery *after 1990* is about half of cohort surgery *before 1990*. The hazard of metastasis of patients with *any neo or adjuvant chemotherapy* is significantly 2.28 higher than those without. Patients who are *post menopausal* have significantly lower metastasis hazard than *pre menopausal*. The hazard of metastasis of patients with *any adjuvant radiotherapy* is significantly 1.92 higher than those without. The *side of the body affected* has no significant effect on the hazard of metastasis. In contrast to local and regional recurrence the log of the shape parameter of Weibull for metastasis is significantly more than zero, which means the hazard of metastasis increases by time.

METASTASIS

Variable	Frailty distribution			
	Univarite Log-Normal	Multivariate Log-Normal	Univarite non-parametric	Multivariate non-parametric
AGE	**-0.054(0.010)	**-0.201(0.026)	**-0.062(0.010)	**-0.061(0.010)
STAGE2	**1.185(0.189)	**5.182(0.586)	**1.508(0.197)	**1.599(0.225)
STAGE3	**2.008(0.321)	**8.917(0.971)	**2.537(0.293)	**2.469(0.339)
STAGE4	**3.885(0.527)	**13.404(1.340)	**4.486(0.350)	**4.178(0.468)
SURGTYPE1	-1.370(0.933)	*-2.887(1.357)	-0.883(1.084)	-0.916(1.188)
SURGTYPE2	0.737(0.388)	**4.642(0.983)	*0.904(0.423)	**1.135(0.390)
SURGTYPE3	**-0.735(0.280)	**-1.328(0.492)	**-0.961(0.300)	*-0.783(0.318)
SURGTYPE4	-0.126(0.239)	*0.994(0.428)	-0.145(0.258)	-0.085(0.264)
SURGTYPE5	0.451(0.405)	0.496(0.720)	**1.365(0.467)	*1.235(0.511)
SURGTYPE6	**-1.137(0.302)	**-4.171(0.797)	**-1.368(0.316)	**-1.453(0.337)
SURGTYPE7	-0.881(0.483)	**-2.125(0.613)	-0.839(0.486)	-0.775(0.486)
HIST2	-0.218(0.220)	-0.576(0.420)	-0.171(0.215)	-0.209(0.226)
HIST3	**-3.279(0.824)	**-12.354(1.668)	**-3.875(0.802)	**-3.990(0.816)
HIST4	-0.418(0.228)	**-2.011(0.729)	*-0.588(0.267)	**-0.672(0.261)
COHORT	**-0.512(0.178)	**-1.122(0.327)	**-0.650(0.214)	**-0.668(0.209)
CHEMO	**0.780(0.173)	**2.711(0.394)	**0.930(0.183)	**0.825(0.195)
MENO	*-0.545(0.235)	**-2.047(0.579)	*-0.574(0.253)	*-0.562(0.253)
RADIO	**0.753(0.229)	**1.684(0.388)	**0.684(0.208)	**0.652(0.210)
SIDE	0.006(0.129)	0.615(0.321)	0.069(0.156)	0.090(0.161)
LN(α)	**0.228(0.077)	**1.585(0.102)	**0.421(0.037)	**0.465(0.074)
CONSTANT	**-9.294(0.920)	**-35.571(4.180)		
-2 Log Likelihood	8808.40		8769.62	

Table 4.8: Results of breast cancer data: metastasis. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

Table 4.9 displays the multivariate non-parametric frailty model of died from breast cancer. In contrast with the pervious recurrences, there is a significant proportional effect of *age* on the hazard of died from breast cancer. There is no significant difference between died from breast cancer hazard of *stage2* and *stage3*, and *stage1* of the disease. While the hazard of *stage4* is significantly 70 times higher than *stage1*, the hazard of died from breast for patients with surgery, *none* and *incision biopsy* is significantly higher than hazard of *radical mastectomy and axillary clearance*, meanwhile there is no significant difference between other surgeries and *radical mastectomy and axillary clearance*. The hazard ratio of died from breast cancer is significant 0.25 times lower for *ductal* than *Dcis (ductal carcinoma in situ)*. There is no significant difference between the hazards of *ductal* and the other two types of histology. Patients with primary surgery *after 1990* have significantly 2.12 times higher hazard ratio of dying due breast than those of surgery *before 1990*. Other variables such as *chemotherapy, menopausal status, radiotherapy*, and *side* of the body affected have no significant effect on the hazard of died from breast cancer. Similar to metastasis the Log of the shape parameter of Weibull for died from breast cancer is significantly more than zero, which means the hazard of died from breast cancer increases by time.

DIED FROM BREAST CANCER

Variable	Frailty distribution			
	Univarite Log-Normal	Multivariate Log-Normal	Univarite non-parametric	Multivariate non-parametric
AGE	**0.059(0.014)	**0.114(0.023)	**0.036(0.009)	**0.040(0.012)
STAGE2	**0.939(0.344)	**1.899(0.411)	*0.530(0.221)	0.525(0.284)
STAGE3	**1.568(0.467)	**2.246(0.524)	**0.826(0.311)	0.848(0.438)
STAGE4	**5.511(0.680)	**12.139(1.346)	**3.426(0.309)	**4.245(0.533)
SURGTYPE1	**3.889(0.626)	**9.334(1.135)	**2.307(0.463)	**2.881(0.788)
SURGTYPE2	**2.910(0.535)	**2.760(0.490)	**1.388(0.325)	**1.566(0.508)
SURGTYPE3	-0.393(0.480)	-0.491(0.495)	-0.141(0.309)	-0.262(0.369)
SURGTYPE4	-0.328(0.449)	-0.098(0.702)	-0.374(0.311)	-0.503(0.371)
SURGTYPE5	-12.646(216.3)	-17.414(49.29)	-7.889(41.46)	-8.092(44.35)
SURGTYPE6	*-1.119(0.556)	**-2.760(0.662)	-0.616(0.397)	-0.722(0.446)
SURGTYPE7	0.867(0.665)	**-3.226(0.879)	0.521(0.538)	0.463(0.617)
HIST2	*0.712(0.299)	0.563(0.345)	0.403(0.218)	0.471(0.293)
HIST3	*-1.733(0.847)	**-5.591(0.970)	*-1.411(0.619)	*-1.375(0.662)
HIST4	0.313(0.317)	**-1.037(0.356)	0.115(0.230)	0.122(0.279)
COHORT	**0.962(0.293)	**2.267(0.394)	**0.609(0.205)	**0.753(0.260)
CHEMO	0.104(0.274)	**1.220(0.422)	0.099(0.199)	0.157(0.233)
MENO	0.747(0.428)	0.050(0.721)	0.210(0.305)	0.206(0.369)
RADIO	-0.210(0.466)	-1.040(0.533)	-0.397(0.306)	-0.335(0.343)
SIDE	*0.537(0.224)	**0.705(0.234)	0.239(0.162)	0.238(0.220)
LN(α)	**0.715(0.106)	**1.369(0.117)	**0.290(0.060)	**0.377(0.119)
CONSTANT	**-25.928(2.782)	**-52.334(6.182)		
-2 Log Likelihood	3801.12		3758.80	

Table 4.9: Results of breast cancer data: died from breast cancer. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

Table 4.10 displays the multivariate non-parametric frailty model of died from other causes. Similar to died from breast cancer, there is a significant proportional effect of *age* on the hazard of died from other causes. The hazard ratio of *stage2* is significantly 1.69 times higher than *stage1*. Whilst there is no significant difference between died from other causes hazard of *stage3* and *stage4*, and *stage1* of the disease. The hazard of died from other causes for patients with surgery, *none*, *incision biopsy*, and *excision biopsy* is significantly higher than hazard of *radical mastectomy and axillary clearance*. In the meantime, there is no significant difference between other surgeries and *radical mastectomy and axillary clearance*. There is no significant difference between the hazard of *ductal* and all other types of histology. The hazard ratio of died from other causes of patients with *any adjuvant radiotherapy* is significantly 0.38 lower than those without. Other variables such as *cohort, chemotherapy, menopausal status*, and *side* of the body affected have no significant effect on the hazard of died from other causes. Similar to local recurrence the log of the shape parameter of Weibull is not significantly different from zero, which means died from other causes has constant hazard.

DIED FROM OTHER CAUSES

Variable	Frailty distribution			
	Univarite Log-Normal	Multivariate Log-Normal	Univarite non-parametric	Multivariate non-parametric
AGE	**0.086(0.012)	**0.086(0.015)	**0.086(0.012)	**0.083(0.012)
STAGE2	0.350(0.232)	0.381(0.259)	0.346(0.232)	*0.527(0.267)
STAGE3	0.401(0.326)	0.405(0.362)	0.399(0.326)	0.509(0.357)
STAGE4	-1.202(1.029)	-1.683(1.076)	-1.204(1.029)	-1.526(1.070)
SURGTYPE1	*1.293(0.576)	1.315(0.886)	*1.295(0.574)	*1.317(0.597)
SURGTYPE2	*1.050(0.459)	1.335(0.802)	*1.050(0.457)	**1.289(0.487)
SURGTYPE3	**0.841(0.319)	*0.897(0.348)	**0.843(0.319)	**0.943(0.337)
SURGTYPE4	0.489(0.325)	0.502(0.391)	0.489(0.325)	0.594(0.343)
SURGTYPE5	1.025(0.754)	1.154(0.802)	1.088(0.754)	1.326(0.798)
SURGTYPE6	0.415(0.461)	0.449(0.487)	0.415(0.461)	0.418(0.470)
SURGTYPE7	0.303(1.040)	0.432(1.066)	0.301(1.040)	0.393(1.061)
HIST2	0.102(0.266)	0.117(0.282)	0.099(0.266)	0.089(0.276)
HIST3	-0.448(0.441)	-0.476(0.482)	-0.453(0.441)	-0.557(0.454)
HIST4	-0.143(0.279)	-0.132(0.289)	-0.146(0.278)	-0.133(0.283)
COHORT	-0.282(0.228)	-0.298(0.235)	-0.280(0.228)	-0.352(0.234)
CHEMO	0.015(0.209)	-0.017(0.751)	0.013(0.204)	0.032(0.214)
MENO	0.394(0.453)	0.419(0.482)	0.392(0.453)	0.357(0.458)
RADIO	**-0.944(0.353)	*-0.979(0.412)	**-0.947(0.352)	**-0.977(0.358)
SIDE	-0.097(0.179)	-0.080(0.195)	-0.098(0.178)	-0.070(0.187)
LN(α)	0.063(0.076)	0.084(0.087)	0.062(0.076)	0.125(0.089)
CONSTANT	**-16.831(1.118)	**-17.209(1.235)		
-2 Log Likelihood	2747.36		2746.92	

Table 4.10: Results of breast cancer data: died from other causes. Parameters' estimates with their standard error in parentheses.

*.P-value < 0.05

**.P-value < 0.01

4.12.2 Merging failure types

Dimensionality in multivariate analysis is a very important issue. Removing one failure type from a competing risks model could minimise the model fitting time remarkably. One would be interested in testing for merge some of failure types

$$\begin{aligned} H_0 &: h_{0j}(t_{ij})e^{\boldsymbol{x}_{ij}'\boldsymbol{\beta_j}} = h_{0k}(t_{ik})e^{\boldsymbol{x}_{ik}'\boldsymbol{\beta_k}} \\ H_1 &: h_{0j}(t_{ij})e^{\boldsymbol{x}_{ij}'\boldsymbol{\beta_j}} \neq h_{0k}(t_{ik})e^{\boldsymbol{x}_{ik}'\boldsymbol{\beta_k}} \end{aligned} \quad j \neq k \ .$$

For example, can local and regional recurrence be merged? To answer these types of questions, the chi-square distribution is used with degrees of freedom equivalent to the difference in number of parameters between the full (before merging) and reduced models (after merging). Table 4.11 lists the deviance of testing merging between each pair of competing risks. Since all deviance values are very large and by using the chi-square distribution with 30 degrees of freedom, the null hypothesis is rejected and none of the competing risks pairs can be merged.

	Local	Regional	Metastasis	Died from breast cancer
Local				
Regional	463.9			
Metastasis	506.1	771.9		
Died from breast cancer	332.9	365.7	546.0	
Died from other causes	293.6	300.4	301.3	295.0

Table 4.11: Deviances of testing for merging competing risks.

4.12.3 Interpretation of the frailties

Using the Cholesky decomposition of the multivariate Log-Normal frailty, the lower triangle of the variance-covariance matrix of the random effects is given by

$$\begin{array}{c} \\ LR \\ RR \\ MT \\ DB \\ DO \end{array}\begin{array}{c} \begin{array}{ccccc} LR & RR & MT & DB & DO \end{array} \\ \begin{pmatrix} l_{11} & 0 & 0 & 0 & 0 \\ l_{12} & l_{22} & 0 & 0 & 0 \\ l_{13} & l_{23} & l_{33} & 0 & 0 \\ l_{14} & l_{24} & l_{34} & l_{44} & 0 \\ l_{15} & l_{25} & l_{35} & l_{45} & l_{55} \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{ccccc} LR & RR & MT & DB & DO \end{array} \\ \begin{pmatrix} 1.377 & 0 & 0 & 0 & 0 \\ 6.001 & 3.653 & 0 & 0 & 0 \\ 4.491 & 2.396 & 6.511 & 0 & 0 \\ -6.358 & 3.688 & 0.095 & 0.0004 & 0 \\ 0.431 & -0.456 & -0.033 & -0.0001 & 0 \end{pmatrix} \end{array}$$

and hence, the variance-covariance matrix of the random effects is $\Sigma = \boldsymbol{L}\boldsymbol{L}'$

$$\begin{array}{c} \\ LR \\ RR \\ MT \\ DB \\ DO \end{array}\begin{array}{c} \begin{array}{ccccc} LR & RR & MT & DB & DO \end{array} \\ \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} & \sigma_{35} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 & \sigma_{45} \\ \sigma_{15} & \sigma_{25} & \sigma_{35} & \sigma_{45} & \sigma_5^2 \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{ccccc} LR & RR & MT & DB & DO \end{array} \\ \begin{pmatrix} 1.90 & 8.26 & 6.18 & -8.76 & 0.59 \\ 8.26 & 49.36 & 35.70 & -24.68 & 0.92 \\ 6.18 & 35.70 & 68.30 & -19.10 & 0.63 \\ -8.76 & -24.68 & -19.10 & 54.03 & -4.43 \\ 0.59 & 0.92 & 0.63 & -4.43 & 0.39 \end{pmatrix} \end{array}$$

The variance of the frailty distribution is very small in case of died from other causes, $\sigma_5^2 = 0.39$ which is an indicator of no important risks factors are omitted from the model. Whilst, the variances of other frailty are very big especially for regional recurrence, metastasis and died from breast cancer with $\sigma_2^2 = 49.36, \sigma_3^2 = 68.3$ and $\sigma_4^2 = 54.03$ respectively. This means that there are many important risks factor are not included in the model. Although, the interpretation of the nature of frailty in multivariate competing risks models is not straight forward and more complex than in univariate cases, still one can get a clear idea about the way they are correlated. In multivariate cases, it is frequently encountered with more than one omitted risk factor. This causes the interpretation of the nature of frailty to be more complex than a univariate cases. The above variance-covariance matrix suggests that

- the frailty of time to local recurrence (LR) is positively correlated with the frailty of time to regional recurrence (RR), metastasis (MT), and died from other causes (DO). The correlation coefficients are 0.85, 0.54, and 0.68 respectively. It is negatively correlated with the frailty of time to died from breast cancer (DB) (ρ = -0.87);
- the frailty of regional recurrence is positively correlated with the frailty of time to metastasis ($\rho = 0.62$), but it is negatively associated with frailty of time to died from breast cancer (ρ = -0.48). It is weakly correlated to died from other causes;
- the frailty of metastasis is weakly correlated with frailty of time to died from breast cancer ($\rho = 0.31$), but it is not correlated with the frailty of time to died from other causes ($\rho = 0.12$); and
- the frailty of died from breast cancer is negatively and highly correlated with frailty of time to died from other causes (ρ = -0.96)

Using the multivariate non-parametric frailty the association coefficients are:

$$\begin{array}{c} \\ LR \\ RR \\ MT \\ DB \\ DO \end{array} \begin{array}{c} \begin{array}{ccccc} LR & RR & MT & DB & DO \end{array} \\ \begin{pmatrix} & & & & \\ r_{12} & & & & \\ r_{13} & r_{23} & & & \\ r_{14} & r_{24} & r_{34} & & \\ r_{15} & r_{25} & r_{35} & r_{45} & \end{pmatrix} \end{array} = \begin{array}{c} \begin{array}{ccccc} LR & RR & MT & DB & DO \end{array} \\ \begin{pmatrix} & & & & \\ 1.15 & & & & \\ 1.07 & 0.07 & & & \\ -1.46 & 0.02 & 0.64 & & \\ 0.50 & 0.01 & -0.15 & 0.01 & \end{pmatrix} \end{array}$$

The diagonal of the association matrix was absorbed in the non-parametric representation of the model and merged with the constant part β_{j0}. As mentioned before, this is not a correlation matrix, and it was included in the model to account for the association between competing risks frailties. As a summary conclusions, died from breast cancer is informative

(highly correlated) for both local and regional recurrence while died from other causes is informative for local recurrence only. It seems that the nature of frailty for local recurrence (non-aggressive cancer), which definitely exists, is highly and positively correlated with other frailties except died from breast cancer, which is negatively correlated. With regard to the time of regional recurrence (aggressive cancer), the frailty was related to the susceptibility of patients to more aggressive types of breast cancer (metastasis). On the other hand, died from breast cancer is negatively associated with all other frailties, especially the frailty of died from other causes. None of these competing risks can be merged with another and it seems that each failure type has its own characteristics and different risk factors.

4.12.4 Clinical results

The following points summarise the effect of the risk factors on the hazard of each type of outcome.

- **Age of patient**. Young patients have higher chance of local, regional and metastasis recurrence than old patients, whereas they have lower hazard of "died from breast cancer" than old patients.

- **Stage of the disease**. Patients in *stage2* and *stage3* have significantly higher hazard of local and regional than patient in *stage1*. The hazard of metastasis of *stage1* is significantly lower than the other three stages. Only patients in *stage4* have significantly higher hazard of "died from breast cancer" than *stage1*.

- **Surgery type**. Patients without surgery or with *incision biopsy* surgery have higher hazard of all recurrence than patients with *radical mastectomy and axillary clearance* except metastasis no difference. Patients with *excision biopsy* surgery have higher hazard of local and regional recurrence than patients with *radical mastectomy and*

axillary clearance and lower hazard for metastasis no difference for "died from breast cancer".

- **Histology**. There is more chance for patients with *ductal* than patients with *Other* histology for local recurrence and higher hazard of regional than all other three types of histology. The hazard of metastasis for patients with *ductal* is the same for *lobular* but higher than the other two types. The hazard of metastasis for patients with *ductal* is higher than *dcis (ductal carcinoma in situ)* but the same for the other two types.

- **Cohort**. Patients with primary surgery *before 1990* have the same change of local recurrence as those *after 1990* but they have lower hazard of regional recurrence and metastasis. However, the hazard of "died from breast cancer" is higher for patients with primary surgery *before 1990* than *after 1990*.

- **Chemotherapy**. Patients with chemotherapy have higher hazard of metastasis than patients without. There is no significant effect of chemotherapy on the hazard of local, regional and "died from breast cancer".

- **Menopausal status**. Patients pre-menopausal have lower hazard of metastasis than patients post-menopausal. There is no significant effect of menopausal status on the hazard of local, regional and "died from breast cancer".

- **Radiotherapy**. Patients with radiotherapy have lower hazard of local recurrence than patients without. There is no significant effect of radiotherapy on the hazard of regional, metastasis and "died from breast cancer".

- **Side of the body affected**. There is no significant effect of the side of body with cancer (right or left) on the hazard of all types of outcome (competing risks).

4.13 Summary

In this chapter, different types of frailty models are discussed, shared frailty, correlated frailty and multivariate frailty. Most of the existent models are correlated frailty models rather than multivariate model. One of the limitations of the correlated frailty model is that they have restricted correlation coefficients between frailties; e.g. Correlated Gamma and Inverse Gaussian frailty discussed in sections 4.6 and 4.7. Whilst the multivariate frailty models have unrestricted correlations, but they have the limitation that the marginal likelihood function does not have a closed form and numerical integration is needed to get the maximum likelihood estimator; e.g. Log-Normal frailty discussed in section 4.9. The interpretation of the frailty is not straight forward in multivariate frailty models. A competing risks model with multivariate frailty is introduced and employed in the analysis of the time to the first recurrence of breast cancer. The analysis was carried out by a Cholesky decomposition of multivariate Log-Normal frailty and by a non-parametric multivariate frailty. The non-parametric frailty model is much less time consuming in fitting the data with the smallest standard errors of parameters estimates. The simulations showed that only a few numbers of mass points are needed to fit the data using non-parametric frailty compared to multivariate Log-normal where at least eight points are needed to get acceptable results. In the analysis of competing risks models, including frailty is important to take into account the potential relation between different failure types. Ignoring this fact and employing the commonly used estimation procedures underestimate the parameters of interest and could lead to inaccurate inference about relevant risk factors. Another way to overcome the problem of mis-specifying the frailty distribution is breaking the frailty distribution in sub-distribution the so-called finite mixtures. in the next chapter, several simulation studies of mixture of different frailly distribution are conducted.

Chapter 5

Frailty and finite mixture

5.1 Introduction

In previous chapters, it was shown that the choice of the frailty distribution is crucial for making valid inferences. Fitting the model using a non-parametric frailty is one way to overcome this problem. In this chapter, a different prospective is used to solve the problem. The model is fitted by breaking the frailty distribution into a finite number of components, the so-called *finite mixture*. Mixture models are usually used to model data that come from a heterogeneous population. Using a mixture of frailty distributions increases the flexibility of modelling the unobserved heterogeneity, especially if the frailty distribution is not unimodal. The finite mixture models of a parametric frailty distribution can be viewed as a semi-parametric models, as they can be written in terms of J components of a specific distribution. In general, a random variable T with a probability density function $f(t)$ can be decomposed into a sum of J class probability density functions. Let $f_j(t)$ denote the j^{th} class probability density function. The finite mixture model with J-component has the following general form

$$f(t|\theta_1, \cdots, \theta_J; \pi_1, \cdots, \pi_{J-1}) = \sum_{j=1}^{J} \pi_j f_j(t|\theta_j),$$

where π_j represents the probability that the realisation t is coming from the j^{th} component. Furthermore, these probabilities must satisfy the following constraints

$$0 < \pi_j < 1 \qquad \text{and} \qquad \sum_{j=1}^{J} \pi_j = 1.$$

The mean and the variance of the finite mixture are

$$\mu = E[T] = \sum_{j=1}^{J} \pi_j \mu_j \qquad \sigma^2 = V[T] = \sum_{j=1}^{J} \pi_j (\mu_j^2 + \sigma_j^2) - \mu^2 \tag{5.1.1}$$

For more information see (Everitt and Hand, 1981) and (Frhwirth-Schnatter, 2006).

5.2 Frailty as a finite mixture

Finite mixture models with and without covariates are extensively studied in the literature. Most of the published work has concentrated on mixtures of normal distribution, with less emphasis on non-normal mixtures. Recently, many studies have factored the random effects into a wide variety of regression models. For example, Hall and Wang (2005) considered a finite mixtures of generalized linear mixed effect models. van Duijn and Bockenholt (1995) presented mixture models for the analysis of repeated count data. Olsen and Schafer (2001) considered regression models with mixed effects for clustered continuous data. A Finite mixture of bivariate Poisson regression models in the presence of random effect was considered by BermDez and Karlis (2012).

In this chapter, a simulation study of finite mixture of frailty models is conducted where the frailty distribution is constructed as a mixture of distributions especially for those with a closed-form of the unconditional hazard such as Gamma and Inverse Gaussian. The purpose of these simulations is to assess the performance of finite mixture of frailty models.

Ravishanker and Dey (2000) considered a multivariate survival data using finite mixtures of positive stable frailty distributions. Hanagal (2008) gave two types of mixture models for survival data with frailty. In the frailty model, the general form of the unconditional likelihood function is

$$L(t_i, \delta_i|\boldsymbol{x_i}) = \int_{R^+} (zh_0(t_i)e^{\boldsymbol{x_i'\beta}})^{\delta_i} e^{-zH_0(t_i)e^{\boldsymbol{x_i'\beta}}} f(z,\tau)dz,$$

where $h_0(t_i)$ is the baseline hazard, $\boldsymbol{x_i}$ is the vector of covariates of the i^{th} subject, $\boldsymbol{\beta}$ is a $p \times 1$ fixed effect vector and $f(z,\tau)$ is the p.d.f of the frailty distribution. Assuming that the frailty distribution can be written as a sum of J class probability density functions, then

$$f(z,\tau) = \sum_{j=1}^{J} \pi_j f_j(z,\tau_j),$$

where, $0 < \pi_j < 1, j = 1, \cdots, J$ and $\sum_{j=1}^{J} \pi_j = 1$. The unconditional likelihood function can be written as

$$L(t_i, \delta_i|\boldsymbol{x_i}) = \int_{R} (-zh_0(t_i)e^{\boldsymbol{x_i'\beta}})^{\delta_i} e^{-zH_0(t_i)e^{\boldsymbol{x_i'\beta}}} \sum_{j=1}^{J} \pi_j f_j(z,\tau_j)dz,$$

consequently,

$$L(t_i, \delta_i|\boldsymbol{x_i}) = \sum_{j=1}^{J} \pi_j \int_{R} (-zh_0(t_i)e^{\boldsymbol{x_i'\beta}})^{\delta_i} e^{-zH_0(t_i)e^{\boldsymbol{x_i'\beta}}} f_j(z,\tau_j)dz \tag{5.2.1}$$

The set of parameters need to be estimated is $\boldsymbol{\theta} = (\beta_1, ..., \beta_p, \tau_1^2, ..., \tau_J^2, \pi_1, ..., \pi_{J-1})'$, which can be estimated either by EM-algorithm or by direct the maximisation of likelihood function.

5.2.1 Finite mixture of Gamma frailty model

Assuming the survival times follow the Weibull distribution $T \sim Weib(\alpha, \lambda)$, and the j^{th} class of the frailty distribution is a Gamma distribution with a unit mean and variance τ_j^2, $f_j \sim \Gamma(1/\tau_j^2, \tau_j^2)$, the frailty distribution can be written as

$$f(z, \tau^2) = \sum_{j=1}^{J} \pi_j \Gamma(1/\tau_j^2, \tau_j^2).$$

The unconditional (marginal) likelihood function has a close-form and given by

$$\begin{aligned} L(t_i, \delta_i | \boldsymbol{x_i}) &= \sum_{j=1}^{J} \pi_j \left(\frac{\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'\beta}}}{1 + \tau_j^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}}} \right)^{\delta_i} [1 + \tau_j^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}}]^{-(1/\tau_j^2)} \\ &= (\alpha t^{\alpha-1} e^{\boldsymbol{x_i'\beta}})^{\delta_i} \sum_{j=1}^{J} \pi_i (1 + \tau_j^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}})^{-(\delta_i + (\frac{1}{\tau_j}))}. \end{aligned} \tag{5.2.2}$$

5.2.2 Finite mixture of Inverse Gaussian frailty model

In this section, a Weibull regression model with a finite mixture of Inverse Gaussian frailty is proposed. Assuming the survival times follow the Weibull distribution, $T \sim Weib(\alpha, \lambda)$, and j^{th} class of the frailty distribution is an Inverse Gaussian distribution with a unit mean and variance τ_j^2, $f_j \sim IG(1, 1/\tau_j^2)$, the frailty distribution can be written as

$$f(z, \tau^2) = \sum_{j=1}^{J} \pi_j IG(1, 1/\tau_j^2).$$

The unconditional likelihood function has a close-form and given by

$$L(t_i, \delta_i | \boldsymbol{x_i}) = \sum_{j=1}^{J} \pi_j \left(\frac{\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'\beta}}}{(1 + 2\tau_j^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}})^{1/2}} \right)^{\delta_i} exp\left(\frac{1}{\tau_j} (1 - (1 + 2\tau_j^2 t_i^{\alpha} e^{\boldsymbol{x_i'\beta}})^{1/2}) \right). \tag{5.2.3}$$

5.2.3 Finite mixture of Log-Normal frailty model

Assuming the survival times follow the Weibull distribution and j^{th} class of the frailty distribution is a Log-Normal random variable with mean μ and variance τ_j^2, the frailty distribution can be written as

$$f(z, \tau^2) = \sum_{j=1}^{J} \pi_j LogN(1/\tau_j^2, \tau_j^2).$$

In this case the unconditional likelihood function does not have a closed-form, but using the results of section 3.4.4, it can be expressed as

$$\begin{aligned} L(t_i, \delta_i, \boldsymbol{x_i}) &= \sum_{j=1}^{J} \pi_j \int_R (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\tau_j y\sqrt{2}})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\tau_j y\sqrt{2}}) \frac{1}{\sqrt{\pi}} e^{-y^2} dy. \\ &\approx \sum_{j=1}^{J} \sum_{k=1}^{K} \pi_j \pi_k^* (\alpha t_i^{\alpha-1} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\tau y_k^*})^{\delta_i} \exp(t_i^{\alpha} e^{\boldsymbol{x_i'}\boldsymbol{\beta}+\tau y_k^*}). \end{aligned} \tag{5.2.4}$$

where y_k^* and π_k^* are the zeros of Hermite polynomials and their corresponding weight factors respectively while π_j is the mixing probability.

5.3 Finite mixture of correlated Inverse Gaussian frailty model

In this section, a mixture of correlated Inverse Gaussian frailty model is proposed based on the results reported in section 4.7. Assuming the survival times follow the Weibull distribution and j^{th} class of the frailty distribution is a correlated Inverse Gaussian distribution with a unit mean vector and variance-covariance matrix $\boldsymbol{\Sigma}_j$, similar to the one given in 4.7.2,

$f_j \sim BIG(\mathbf{1}, \mathbf{\Sigma}_j)$, the frailty distribution can be written as

$$f(\boldsymbol{z}, \mathbf{\Sigma}) = \sum_{j=1}^{J} \pi_j BIG(\mathbf{1}, \mathbf{\Sigma}_j).$$

In general, the mean and the variance of multivariate mixtures are given by

$$E[\boldsymbol{Z}] = \boldsymbol{\mu} = \sum_{j=1}^{J} \pi_j \boldsymbol{\mu}_j, \qquad V[\boldsymbol{Z}] = \sum_{j=1}^{J} \pi_j \mathbf{\Sigma}_j + \sum_{j=1}^{J} \pi_j (\boldsymbol{\mu}_j - \boldsymbol{\mu})(\boldsymbol{\mu}_j - \boldsymbol{\mu})^{'} \tag{5.3.1}$$

For the mixture in 5.3.1, the variance reduces to

$$V[\boldsymbol{Z}] = \sum_{j=1}^{J} \pi_j \mathbf{\Sigma}_j$$

In the bivariate case, the unconditional log-likelihood function is given by

$$\begin{aligned}
&\ell(t_{i1}, t_{i2}) = \\
&\sum_{j=1}^{J} \pi_j \Big\{ \delta_{i1}\delta_{i2} \log \left[\frac{\partial}{\partial t_{i2} \partial t_{i1}} \log S_j(t_{i1}, t_{i2}) + \left(\frac{\partial}{\partial t_{i1}} \log S_j(t_{i1}, t_{i2}) \right) \left(\frac{\partial}{\partial t_{i2}} \log S_j(t_{i1}, t_{i2}) \right) \right] \\
&+ \sum_{j=1}^{2} \delta_{ij} \log \left(-\frac{\partial}{\partial t_{ij}} \log S_j(t_{i1}, t_{i2}) \right) + \log[S_j(t_{i1}, t_{i2})] \Big\}.
\end{aligned} \tag{5.3.2}$$

where,

$$\begin{aligned}
&S_j(t_{i1}, t_{i2}) = \\
&[S_1(t_{i1})]^{(1-\rho_j \frac{\tau_{1j}}{\tau_{2j}})} [S_2(t_{i2})]^{(1-\rho_j \frac{\tau_{2j}}{\tau_{1j}})} \\
&\times \exp \left\{ \frac{\rho_j}{\tau_{1j}\tau_{2j}} \left(1 - \left[\left(1 - \tau_{1j}^2 \log S_1(t_{i1})\right)^2 + \left(1 - \tau_{2j}^2 \log S_2(t_{i2})\right)^2 - 1 \right]^{1/2} \right) \right\}
\end{aligned}$$

In case of competing risks, either the individual faces one of the failures (i.e. only one of

$\delta_{ij} = 1, j = 1, 2)$ or censored (i.e. both $\delta_{ij} = 0, j = 1, 2)$. In constructing the likelihood function only the first order of the partial derivatives is needed. Hence the log-likelihood in 5.3.2 reduces to

$$\ell(t_{i1}, t_{i2}) = \sum_{j=1}^{J} \sum_{k=1}^{2} \pi_j \delta_{ik} \log \left(-\frac{\partial}{\partial t_{ik}} \log S_j(t_{i1}, t_{i2}) \right) + \log[S_j(t_{i1}, t_{ij})]. \tag{5.3.3}$$

5.4 Simulations

This section examines the performance of a frailty mixture with different distribution in both univariate and bivariate data. In univariate data, three mixtures are tested, Gamma, Inverse Gaussian, and Log-Normal distributions. In bivariate frailty, only a mixture of the correlated Inverse Gaussian frailty proposed in section 4.7 is examined. In each study, data are generated using Weibull baseline hazard and four different frailty distributions, namely, the Log-Normal, Gamma, Inverse Gaussian and an arbitrary distribution. These models are fitted using mixtures of Gamma, Inverse Gaussian, and Log-Normal frailty distributions.

Simulations from the frailty distributions were conducted in the same manner as described in the previous chapters. The arbitrary distribution is generated from a discrete random variable, say Y with an expected value equals to one using the following steps. First, generate a random numbers from a Uniform distribution $U \sim Uni(0, 1)$. Second, cut the range $(0, 1)$ into different segments with different lengths (probabilities). Third, if the generated number from is within the first segment return some value of Y and if it within the second segment return another value Y, and so forth. The returned values and the length of segments are set so that the expect value of the random variable Y is equal to one.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian		Arbitrary	
		Sample size		Sample size		Sample size		Sample size	
		500	5000	500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_0	-4	-3.933	-3.919	-4.049	-4.008	-3.909	-3.878	-3.907	-3.879
		(0.259)	(0.085)	(0.308)	(0.094)	(0.258)	(0.079)	(0.282)	(0.089)
β_1	1	0.943	0.943	1.024	0.999	0.909	0.917	0.949	0.937
		(0.246)	(0.078)	(0.304)	(0.094)	(0.241)	(0.073)	(0.273)	(0.083)
β_2	-2	-1.887	-1.877	-2.006	-2.004	-1.857	-1.839	-1.885	-1.865
		(0.201)	(0.060)	(0.241)	(0.078)	(0.191)	(0.061)	(0.224)	(0.066)
β_3	4	3.77	3.758	4.037	4.010	3.715	3.678	3.753	3.727
		(0.305)	(0.094)	(0.378)	(0.110)	(0.292)	(0.091)	(0.337)	(0.102)
β_4	2	1.879	1.878	2.016	2.008	1.877	1.838	1.879	1.862
		(0.216)	(0.073)	(0.279)	(0.086)	(0.216)	(0.070)	(0.249)	(0.081)
α	1	0.945	0.940	1.009	1.002	0.931	0.921	0.942	0.934
		(0.068)	(0.020)	(0.082)	(0.024)	(0.065)	(0.019)	(0.072)	(0.022)
τ^2	1	0.364	0.377	0.993	1.004	0.337	0.334	0.662	0.661
		(0.147)	(0.043)	(0.238)	(0.068)	(0.137)	(0.043)	(0.188)	(0.053)

Table 5.1: Log-Normal, Gamma, Inverse Gaussian and arbitrary frailty models with Weibull baseline hazard and four covariates simulated data estimated by Gamma frailty, 500 data sets each with sample sizes of 500 and 5000.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian		Arbitrary	
		Sample size		Sample size		Sample size		Sample size	
		500	5000	500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_0	4	-3.929	-3.909	-4.016	-3.988	-3.893	-3.880	-3.889	-3.876
		(0.254)	(0.085)	(0.305)	(0.091)	(0.254)	(0.082)	(0.288)	(0.085)
β_1	1	0.931	0.935	0.996	0.994	0.917	0.922	0.927	0.932
		(0.233)	(0.077)	(0.291)	(0.090)	(0.243)	(0.072)	(0.275)	(0.084)
β_2	-2	-1.879	-1.874	-1.981	-1.990	-1.832	-1.841	-1.873	-1.860
		(0.191)	(0.060)	(0.246)	(0.075)	(0.194)	(0.063)	(0.226)	(0.068)
β_3	4	3.758	3.746	3.983	3.976	3.681	3.678	3.740	3.723
		(0.293)	(0.094)	(0.365)	(0.114)	(0.294)	(0.096)	(0.337)	(0.099)
β_4	2	1.880	1.875	1.978	1.990	1.851	1.835	1.874	1.858
		(0.214)	(0.072)	(0.281)	(0.090)	(0.216)	(0.071)	(0.247)	(0.078)
α	1	0.941	0.938	0.992	0.995	0.924	0.921	0.939	0.933
		(0.064)	(0.020)	(0.081)	(0.026)	(0.066)	(0.021)	(0.072)	(0.021)
τ^2	1	0.382	0.372	0.937	0.981	0.322	0.335	0.655	0.659
		(0.319)	(0.044)	(0.238)	(0.076)	(0.147)	(0.045)	(0.181)	(0.053)
τ_1^2		0.682	0.372	0.644	0.767	0.231	0.287	0.479	0.586
		(7.542)	(0.044)	(0.500)	(0.691)	(0.167)	(0.107)	(0.241)	(0.124)
τ_2^2		1.957	0.372	0.667	0.731	0.229	0.287	0.473	0.586
		(31.033)	(0.044)	(0.535)	(0.450)	(0.239)	(0.103)	(0.243)	(0.127)
τ_3^2		1.855	0.372	0.676	0.768	0.238	0.289	0.480	0.584
		(36.809)	(0.044)	(0.502)	(0.510)	(0.174)	(0.108)	(0.239)	(0.125)
τ_4^2		0.964	0.372	0.721	0.740	0.258	0.289	0.471	0.588
		(15.666)	(0.044)	(0.846)	(0.536)	(0.412)	(0.104)	(0.240)	(0.129)
τ_5^2		1.240	0.373	0.672	0.734	0.238	0.286	0.486	0.589
		(16.411)	(0.048)	(0.604)	(0.393)	(0.341)	(0.107)	(0.249)	(0.127)
π_1		0.190	0.199	0.190	0.185	0.188	0.192	0.187	0.187
		(0.112)	(0.035)	(0.283)	(0.280)	(0.257)	(0.191)	(0.243)	(0.187)
π_2		0.200	0.198	0.208	0.208	0.185	0.202	0.196	0.199
		(0.133)	(0.032)	(0.293)	(0.299)	(0.245)	(0.210)	(0.248)	(0.199)
π_3		0.197	0.200	0.199	0.223	0.208	0.210	0.210	0.194
		(0.135)	(0.028)	(0.296)	(0.323)	(0.273)	(0.209)	(0.273)	(0.198)
π_4		0.200	0.201	0.213	0.186	0.230	0.207	0.193	0.212
		(0.132)	(0.033)	(0.297)	(0.289)	(0.282)	(0.213)	(0.246)	(0.208)
π_5		0.213	0.202	0.189	0.199	0.190	0.190	0.214	0.208
		(0.148)	(0.041)	(0.284)	(0.289)	(0.252)	(0.183)	(0.272)	(0.213)

Table 5.2: Log-Normal, Gamma, Inverse Gaussian and arbitrary frailty models with Weibull baseline hazard and four covariates simulated data estimated by mixture of Gamma frailty, 500 data sets each with sample sizes of 500 and 5000.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian		Arbitrary	
		Sample size		Sample size		Sample size		Sample size	
		500	5000	500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_0	4	-4.040	-3.999	-3.587	-3.563	-4.070	-4.010	-3.843	-3.814
		(0.304)	(0.092)	(0.233)	(0.076)	(0.318)	(0.099)	(0.258)	(0.078)
β_1	1	1.056	1.007	0.890	0.883	1.073	1.007	0.998	0.978
		(0.295)	(0.088)	(0.282)	(0.086)	(0.312)	(0.086)	(0.291)	(0.090)
β_2	-2	-2.079	-2.016	-1.797	-1.771	-2.122	-2.013	-1.991	-1.963
		(0.301)	(0.086)	(0.213)	(0.064)	(0.353)	(0.094)	(0.221)	(0.067)
β_3	4	4.172	4.029	3.564	3.517	4.240	4.025	3.964	3.900
		(0.553)	(0.146)	(0.266)	(0.081)	(0.622)	(0.168)	(0.300)	(0.088)
β_4	2	2.087	2.012	1.803	1.778	2.116	2.011	1.982	1.946
		(0.333)	(0.093)	(0.218)	(0.069)	(0.360)	(0.102)	(0.246)	(0.077)
α	1	1.046	1.008	0.911	0.897	1.061	1.006	1.000	1.001
		(0.134)	(0.034)	(0.054)	(0.016)	(0.149)	(0.041)	(0.068)	(0.478)
τ^2	1	2.160	1.059	1.851	1.715	2.327	1.068	2.346	1.995
		(4.000)	(0.289)	(0.487)	(0.113)	(3.186)	(0.334)	(1.305)	(0.232)

Table 5.3: Log-Normal, Gamma, Inverse Gaussian and arbitrary frailty models with Weibull baseline hazard and four covariates simulated data estimated by Inverse Gaussian frailty, 500 data sets each with sample sizes of 500 and 5000.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian		Arbitrary	
		Sample size		Sample size		Sample size		Sample size	
		500	5000	500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_0	4	-4.050	-4.008	-3.958	-3.971	-4.030	-4.004	-3.995	-3.976
		(0.312)	(0.095)	(0.319)	(0.100)	(0.332)	(0.099)	(0.299)	(0.094)
β_1	1	1.022	1.005	1.052	1.056	1.030	1.004	1.069	1.043
		(0.286)	(0.080)	(0.330)	(0.105)	(0.300)	(0.086)	(0.332)	(0.100)
β_2	-2	-2.057	-2.001	-2.110	-2.110	-2.064	-2.004	-2.103	-2.082
		(0.317)	(0.091)	(0.319)	(0.102)	(0.360)	(0.108)	(0.302)	(0.112)
β_3	4	4.133	4.005	4.194	4.221	4.111	4.008	4.210	4.162
		(0.580)	(0.157)	(0.529)	(0.184)	(0.644)	(0.192)	(0.497)	(0.192)
β_4	2	2.062	2.003	2.111	2.107	2.064	2.002	2.109	2.078
		(0.327)	(0.099)	(0.348)	(0.113)	(0.371)	(0.114)	(0.335)	(0.118)
α	1	1.034	1.001	1.049	1.055	1.032	1.002	1.054	1.041
		(0.137)	(0.037)	(0.120)	(0.044)	(0.155)	(0.045)	(0.117)	(0.046)
τ^2	1	1.846	1.020	2.878	1.952	2.062	1.063	4.302	1.989
		(2.664)	(0.300)	(2.636)	(0.614)	(3.139)	(0.516)	(4.155)	(0.546)
τ_1^2		2.488	1.025	52.384	120.800	2.388	0.754	6.494	4.342
		(3.347)	(1.335)	(76.120)	(165.430)	(3.554)	(1.136)	(7.048)	(2.604)
τ_2^2		0.421	0.398	1.706	17.318	0.407	0.214	0.606	0.575
		(1.892)	(2.966)	(3.520)	(78.278)	(1.557)	(0.311)	(2.937)	(1.208)
τ_3^2		0.098	0.034	0.870	0.995	0.079	0.033	0.126	0.094
		(0.453)	(0.031)	(4.357)	(2.233)	(0.118)	(0.028)	(0.164)	(0.078)
τ_4^2		0.063	0.017	0.257	0.407	0.083	0.014	0.098	0.054
		(0.220)	(0.042)	(1.661)	(1.513)	(0.479)	(0.020)	(0.465)	(0.070)
τ_5^2		0.175	0.906	0.747	24.502	0.187	0.497	0.247	0.464
		(0.957)	(3.019)	(5.499)	(72.074)	(1.633)	(1.131)	(1.204)	(1.195)
π_1		0.600	0.505	0.295	0.149	0.649	0.522	0.602	0.568
		(0.335)	(0.385)	(0.217)	(0.191)	(0.356)	(0.457)	(0.231)	(0.200)
π_2		0.153	0.195	0.366	0.357	0.163	0.123	0.150	0.225
		(0.282)	(0.313)	(0.329)	(0.271)	(0.291)	(0.298)	(0.232)	(0.216)
π_3		0.026	0.009	0.111	0.198	0.021	0.004	0.025	0.032
		(0.099)	(0.035)	(0.229)	(0.259)	(0.084)	(0.018)	(0.095)	(0.082)
π_4		0.046	0.025	0.040	0.090	0.042	0.007	0.033	0.033
		(0.115)	(0.072)	(0.132)	(0.168)	(0.118)	(0.024)	(0.098)	(0.090)
π_5		0.175	0.266	0.189	0.206	0.125	0.344	0.190	0.141
		(0.234)	(0.355)	(0.229)	(0.243)	(0.212)	(0.439)	(0.210)	(0.217)

Table 5.4: Log-Normal, Gamma, Inverse Gaussian and arbitrary frailty models with Weibull baseline hazard and four covariates simulated data estimated by mixture of Inverse Gaussian frailty, 500 data sets each with sample sizes of 500 and 5000.

Parameter	True values	Log-Normal		Gamma		Inverse Gaussian		Arbitrary	
		Sample size		Sample size		Sample size		Sample size	
		500	5000	500	5000	500	5000	500	5000
		Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)	Mean (S.e)
β_0	4	-4.030	-3.986	-4.015	-3.988	-4.016	-3.903	-4.024	-3.965
		(0.317)	(0.094)	(0.360)	(0.100)	(0.308)	(0.108)	(0.279)	(0.080)
β_1	1	1.004	0.987	1.151	1.068	1.002	0.983	1.035	0.996
		(0.276)	(0.080)	(0.475)	(0.104)	(0.273)	(0.081)	(0.280)	(0.081)
β_2	-2	-2.015	-1.989	-2.266	-2.128	-1.992	-1.973	-2.050	-1.991
		(0.286)	(0.084)	(0.568)	(0.106)	(0.321)	(0.075)	(0.235)	(0.071)
β_3	4	4.048	3.975	4.545	4.258	3.997	3.937	4.103	3.984
		(0.514)	(0.140)	(1.120)	(0.181)	(0.537)	(0.136)	(0.366)	(0.111)
β_4	2	2.015	1.986	2.263	2.131	1.989	1.965	2.053	1.994
		(0.316)	(0.089)	(0.601)	(0.116)	(0.325)	(0.087)	(0.283)	(0.080)
α	1	1.010	0.994	1.135	1.065	0.999	0.985	1.025	0.996
		(0.120)	(0.033)	(0.269)	(0.043)	(0.127)	(0.031)	(0.085)	(0.025)
τ^2	1	0.771	0.805	1.292	1.213	0.759	0.795	0.810	0.758
		(0.284)	(0.081)	(0.525)	(0.108)	(0.299)	(0.072)	(0.227)	(0.080)
τ_1^2		0.643	0.688	1.017	1.047	0.605	0.502	0.633	0.615
		(0.448)	(0.274)	(0.808)	(0.647)	(0.403)	(0.233)	(0.524)	(0.412)
τ_2^2		0.663	0.684	0.964	1.097	0.602	0.498	0.673	0.597
		(0.389)	(0.290)	(0.811)	(0.650)	(0.297)	(0.218)	(0.826)	(0.415)
τ_3^2		0.658	0.656	0.985	1.091	0.608	0.496	0.606	0.635
		(0.633)	(0.198)	(0.803)	(0.611)	(0.369)	(0.219)	(0.508)	(0.427)
τ_4^2		0.636	0.692	1.059	1.082	0.628	0.513	0.609	0.589
		(0.595)	(0.340)	(0.957)	(0.646)	(0.444)	(0.228)	(0.480)	(0.402)
τ_5^2		0.608	0.667	1.033	1.092	0.629	0.516	0.603	0.600
		(0.367)	(0.215)	(0.852)	(0.622)	(0.443)	(0.218)	(0.468)	(0.394)
π_1		0.203	0.203	0.212	0.205	0.182	0.177	0.179	0.205
		(0.250)	(0.260)	(0.243)	(0.198)	(0.235)	(0.331)	(0.224)	(0.229)
π_2		0.207	0.209	0.190	0.186	0.192	0.199	0.198	0.214
		(0.252)	(0.261)	(0.231)	(0.186)	(0.252)	(0.358)	(0.223)	(0.231)
π_3		0.199	0.187	0.205	0.196	0.191	0.189	0.209	0.212
		(0.259)	(0.233)	(0.246)	(0.187)	(0.251)	(0.343)	(0.228)	(0.227)
π_4		0.196	0.210	0.192	0.203	0.220	0.220	0.204	0.185
		(0.244)	(0.259)	(0.230)	(0.193)	(0.276)	(0.372)	(0.235)	(0.227)
π_5		0.195	0.190	0.201	0.210	0.215	0.215	0.211	0.183
		(0.248)	(0.256)	(0.230)	(0.207)	(0.270)	(0.369)	(0.228)	(0.223)

Table 5.5: Log-Normal, Gamma, Inverse Gaussian and arbitrary frailty model with Weibull baseline hazard and four covariates simulated data estimated by mixture of Log-Normal frailty, 500 data sets each with sample sizes of 500 and 5000.

		Log-Normal		Gamma		Inverse Gaussian	
		500	5000	500	5000	500	5000
Parameter	True values	Mean	Mean	Mean	Mean	Mean	Mean
		(S.e)	(S.e)	(S.e)	(S.e)	(S.e)	(S.e)
β_{10}	-4	-4.041	-3.984	-4.080	-4.042	-3.995	-3.984
		(0.354)	(0.343)	(0.423)	(0.123)	(0.371)	(0.110)
β_{11}	9	9.603	9.098	9.917	9.164	9.541	9.070
		(1.730)	(0.504)	(1.631)	(0.378)	(1.853)	(0.487)
β_{12}	3	3.176	3.041	3.255	2.967	3.197	3.041
		(0.624)	(0.231)	(0.601)	(0.167)	(0.674)	(0.191)
α_1	0.5	0.530	0.500	0.550	0.500	0.530	0.510
		(0.100)	(0.030)	(0.090)	(0.020)	(0.140)	(0.030)
β_{20}	-3	-2.948	-2.977	-2.864	-2.882	-2.937	-2.993
		(0.314)	(0.332)	(0.360)	(0.098)	(0.370)	(0.086)
β_{21}	7	7.412	7.099	7.645	7.117	7.470	7.081
		(1.394)	(0.430)	(1.330)	(0.366)	(1.675)	(0.423)
β_{22}	4	4.246	4.049	4.452	4.176	4.275	4.024
		(0.767)	(0.279)	(0.781)	(0.188)	(0.909)	(0.232)
α_2	0.5	0.530	0.500	0.550	0.520	0.530	0.500
		(0.090)	(0.040)	(0.090)	(0.020)	(0.150)	(0.030)
τ_1^2	0.8	6.04	1.010	20.61	2.950	6.35	0.740
		(14.64)	(0.670)	(68.01)	(1.330)	(14.75)	(0.840)
τ_2^2	1.25	7.16	1.610	37.78	8.600	8.44	1.580
		(13.42)	(0.880)	(80.79)	(3.410)	(17.11)	(0.980)
ρ	0.3	0.540	0.490	0.570	0.650	0.550	0.520
		(0.340)	(0.280)	(0.330)	(0.230)	(0.320)	(0.290)

Table 5.6: Bivariate Log-Normal, Gamma and Inverse Gaussian frailty model with Weibull baseline hazard and two covariates simulated data fitted by mixture of bivariate Inverse Gaussian, 500 data sets each with sample sizes of 500 and 5000.

The above tables describe four simulation studies of finite frailty mixtures. The simulated data is generated with four different frailty distributions, Log-Normal, Gamma, Inverse Gaussian and an arbitrary distribution. In the first study, data was fitted by Gamma frailty and mixture of Gamma, results are shown in tables 5.1 and 5.2 respectively. Neither the Gamma frailty model nor its mixture was able to capture the model estimates even with five Gamma mixtures except when the original frailty distribution is Gamma. In the second study, the simulated data was fitted by Inverse Gaussian and mixture of Inverse Gaussian distribution. The results are shown in tables 5.3 and 5.4 respectively. Similar to Gamma distribution, the Inverse Gaussian frailty was not able to capture the model estimates, while its mixture managed to fit all of the four frailty models. Only the frailty variance, τ^2 is not close to its true value since it is the mixing parameter. The mixing variances and their corresponding weights are represented by τ_i^2 and π_i, $(i = 1, \ldots, 5)$ respectively. In the third study, the simulated data was fitted by the Log-Normal and the mixture of Log-Normal distributions. The results of Log-Normal frailty are in Table 3.2 in chapter three, while the results of its mixture are presented in in Table 5.5. The Log-Normal mixture model displays similarity to the Inverse Gaussian mixture with respect to capturing the estimates of the four frailty models. These results confirm the conclusions drawn from the previous chapters. Because of quadrature integration in a Log-normal mixture, the Inverse Gaussian mixture is preferable since it is less time consuming. The last study of simulation is the correlated bivariate Inverse Gaussian frailty mixture. The results are summarized in Table 5.6. Obviously, the bivariate Inverse Gaussian mixture is capable of capturing the parameter estimates of other frailty models except the frailty variances since they are the mixing parameters. Only three mixtures are capable to fit the model, the mixing variances and their corresponding weights are not presented since they are not of the main interest.

5.5 Summary

The main goal of this chapter was to check the performance of finite mixture of frailty models through simulation studies. In the univariate frailty, three finite mixtures are considered, Gamma, Inverse Gaussian and Log-Normal. The Gamma mixture was not able to capture the model parameters except when the original frailty distribution is Gamma. Both Inverse Gaussian and Log-Normal finite mixture were capable to fit the model parameters whatever is the original frailty distribution. However, the inverse Gaussian mixture is preferable since it does not involve numerical integrations. In the bivariate frailty, only the Inverse Gaussian mixture is considered. Using only three mixtures, it managed to fit the model parameters very well except the frailty variances and the correlation coefficient and overestimates them which expected since it they are analytically different from the true parameters, see formulae 5.3.1.

Chapter 6

Conclusions

6.1 Introduction

The present thesis discussed a variation of univariate, bivariate and multivariate frailty models in the presence of competing risks. There are two sources of variability in survival data, variability due to observable covariates and variability caused by unknown risk factors which usually is uncontrollable. Estimating the individual hazard rate without taking into account the unobserved heterogeneity will underestimate the hazard function. Competing risks frailty models consist of two underlying distributions: baseline hazard distribution and the random effect distribution. The main emphasis in this thesis is on the frailty distribution rather than the baseline hazard distribution which is assumed to have a Weibull distribution.

6.2 Concluding Remarks

The simulations showed that the right specification of the frailty distribution is crucial for making valid model inferences. There are four proposed frailty models in this thesis. First, a novel non-parametric multivariate frailty model with competing risks which showed a significant decrease in time to model by more than 80%. A model that can accommodate

all multivariate frailty models irrespective of the original frailty distribution. Second, a correlated Inverse Gaussian frailty model, but with restrictions on correlation coefficients. Third, a multivariate Inverse Gaussian frailty model without any restrictions. Fourth, a finite mixture of frailty models especially for those with a close-form of unconditional survival function. One should distinguish between correlated and multivariate models. In correlated models, the marginal distributions are known while the joint distribution is constructed by the sum of marginals. In multivariate models, both the marginals and joint distribution are well defined. Most of the published research in the area of frailty is for correlated models rather than multivariate models. This is what makes the multivariate Log-Normal and the proposed non-parametric frailty models more flexible in modelling frailty.

There are two advantages of the proposed non-parametric multivariate frailty over the Log-Normal frailty. First, it is a distribution free model which does not depends on the original distribution of the frailty. Second, it is much less time consuming compared with the Log-Normal distribution, even when more parameters are added to the model (the quadrature points and their corresponding weights). By estimating the quadrature points and their weights from the model it minimises the number of iterations of the model fit. Latent approach of competing risks is not a full multivariate survival data in the sense that only the minimum failure time is observed, i.e. there is only one dependent variable. The multivariate settings come from the inclusion of the indicator variable of each failure in the model. One of the limitations of correlated frailty models is that they may be not flexible enough in modelling data that are negatively correlated. Another limitation is that in frailty models, the likelihood function is usually expressed in terms of all partial derivatives of the survival function. Hence, generalisation from the bivariate to the multivariate is not straightforward. On the other hand, this could be an advantage of competing risks model over the multivariate survival data since only the first order of the partial derivative is needed to fit the model.

The simulation studies of the non-parametric frailty for univariate and multivariate showed that a small number of quadrature points (around three) is needed to fit the model well. To test the applicability of the proposed models, a real data set of breast cancer was used. It is a complicated data set in which around three thousand patients of breast cancer and five competing risks are included in the model. Both Log-Normal and non-parametric frailty were tried for the data. In the Log-Normal frailty model with five quadrature points a vector with $(5^5 = 3125)$ points is needed to get a single iteration (number of quadrature to the power of number of competing risks).

From an estimation point of view, one important issue in competing risks with frailty model is the initial values of the model parameters to start with. The following procedure is suggested to minimise the fitting time of the model. First, start with individual failure time without frailty and use them as starting values for individual failure with frailty. Second, aggregate these results as starting values of the multivariate model. Third, for the Log-Normal frailty it is better to start with two quadrature points and then use them as starting values for higher number of quadrature points.

6.3 Limitations and future research

Despite their similarity with linear mixed models, frailty models need special treatment. There is a need for additional developments in methods of frailty models analysis especially in multivariate case. One of the limitations of the multivariate Log-Normal frailty is the numerical integration that is needed to fit the model. The time it needs to fit the model depends on the number of quadrature points for the numerical integration and the dimension of the multivariate (number of failure times) which creates nested loops in the estimation process. Converting loops into vectors decreased the time of model fit tremendously, but still

it is a time consuming process. For future research, one may work on relaxing the assumptions used in frailty models such as constant frailty across individuals and the proportionality of hazard to suit practical applications. Throughout this thesis, it was assumed that the baseline hazard is following the Weibull distribution (i.e. parametric baseline hazard). A further research could be more general models where both the baseline hazard and the frailty are distribution free, i.e. a full non-parametric model.

Appendices

Appendix A

Data

A.1 Variables in the model

The variables included in the regression analysis of breast cancer data are as follow:

AGE a continuous variable represents the age of patient in years

STAGE a categorical variable with four level. Stage of the disease which is known as Manchester stage as given in page 43. The code for dummy variables used is as follows

STAGE1	0	0	0
STAGE2	1	0	0
STAGE3	0	1	0
STAGE4	0	0	1

SURGTYPE a categorical variable of surgery type with with eight levels: the code for dummy variables used is as follows

None	1	0	0	0	0	0	0
Incision biopsy	0	1	0	0	0	0	0
Excision biopsy	0	0	1	0	0	0	0
Simple mastectomy	0	0	0	1	0	0	0
Radical mastectomy	0	0	0	0	1	0	0
Wide local excision and axillary clearance	0	0	0	0	0	1	0
Surgery after neo adjuvant chemotherapy	0	0	0	0	0	0	1
Radical mastectomy and axillary clearance	0	0	0	0	0	0	0

HIST a categorical variable of histology status with four levels: the code for dummy variables used is as follows

Ductal	0	0	0
Lobular	1	0	0
Dcis (Ductal Carcinoma In Situ)	0	1	0
Other	0	0	1

COHORT a categorical variable of date of primary surgery with two levels:

(0)Before 1990 (1)After 1990

CHEMO a categorical variable of any neo or adjuvant chemotherapy with two levels:

(0)No (1)Yes

MENO a categorical variable of menopausal status with two levels:

(0)PRE (1)Post

RADIO a categorical variable of any adjuvant radiotherapy with two levels:

(0)No (1)Yes

SIDE a categorical variable of side of the body affected with two levels:

(0)Right (1) Left

A.2 Variables by risks

		Local recurrence		Regional recurrence		Metastasis		Died from breast cancer		Died from other causes	
Variable		N	%	N	%	N	%	N	%	N	%
STAGE	1	110	65.1	147	56.3	232	51.4	91	49.2	87	68
	2	33	19.5	56	21.5	124	27.5	34	18.4	27	21.1
	3	26	15.4	46	17.6	46	10.2	19	10.3	13	10.2
	4	0	0	12	4.6	49	10.9	41	22.2	1	0.8
SURGERY	1	6	3.6	8	3.1	2	0.4	19	10.3	8	6.3
	2	32	18.9	40	15.3	47	10.4	47	25.4	12	9.4
	3	56	33.1	91	34.9	75	16.6	25	13.5	42	32.8
	4	25	14.8	75	28.7	111	24.6	20	10.8	33	25.8
	5	2	1.2	8	3.1	13	2.9	0	0	2	1.6
	6	10	5.9	6	2.3	55	12.2	13	7	8	6.3
	7	2	1.2	4	1.5	10	2.2	5	2.7	1	0.8
	8	36	21.3	29	11.1	138	30.6	56	30.3	22	17.2
HIST	1	105	62.1	202	77.4	342	75.8	93	50.3	76	59.4
	2	18	10.7	15	5.7	55	12.2	32	17.3	18	14.1
	3	19	11.2	2	0.8	2	0.4	3	1.6	7	5.5
	4	27	16	42	16.1	52	11.5	57	30.8	27	21.1
COHORT	1	81	47.9	183	70.1	238	52.8	76	41.1	80	62.5
	2	88	52.1	78	29.9	213	47.2	109	58.9	48	37.5
CHEMO	1	100	59.2	175	67	202	44.8	118	63.8	66	51.6
	2	69	40.8	86	33	249	55.2	67	36.2	62	48.4
MENO	1	113	66.9	159	60.9	282	62.5	159	85.9	120	93.8
	2	56	33.1	102	39.1	169	37.5	26	14.1	8	6.3
RADIO	1	134	79.3	171	65.5	265	58.8	155	83.8	113	88.3
	2	35	20.7	90	34.5	186	41.2	30	16.2	15	11.7
SIDE	1	84	49.7	119	45.6	216	47.9	80	43.2	62	48.4
	2	85	50.3	142	54.4	235	52.1	105	56.8	66	51.6
Age Mean ± SD		58.0 ± 14.3		56.0 ± 14.6		53.9 ± 12.8		65.8 ± 14.0		71.6 ± 11.9	

Table A.1: Independent variables by recurrence type.

A.3 Data analysis without frailty

Variable	Frailure type				
	Local recurrence	Regional recurrence	Metastasis	Died from breast cancer	Died from other causes
AGE	-0.012(0.01)	-0.012(0.01)	-0.038(0.01)	0.03(0.01)	0.086(0.01)
STAGE2	0.533(0.21)	0.772(0.17)	0.844(0.12)	0.548(0.21)	0.346(0.23)
STAGE3	0.565(0.29)	1.373(0.22)	1.291(0.18)	0.725(0.29)	0.399(0.33)
STAGE4	-10.89(70.7)	0.521(0.35)	2.664(0.23)	2.163(0.27)	-1.204(1.03)
SURGTYPE1	2.894(0.56)	2.148(0.49)	-0.904(0.75)	2.077(0.42)	1.296(0.59)
SURGTYPE2	2.92(0.34)	2.177(0.32)	0.472(0.25)	1.385(0.31)	1.051(0.47)
SURGTYPE3	1.23(0.27)	1.522(0.26)	-0.552(0.19)	-0.055(0.3)	0.844(0.32)
SURGTYPE4	0.198(0.3)	1.084(0.25)	-0.08(0.16)	-0.242(0.3)	0.489(0.33)
SURGTYPE5	0.212(0.73)	1.565(0.41)	0.387(0.3)	-7.886(42.66)	1.088(0.75)
SURGTYPE6	-0.001(0.4)	-0.851(0.48)	-0.805(0.2)	-0.551(0.39)	0.417(0.46)
SURGTYPE7	0.273(0.75)	0.725(0.55)	-0.636(0.35)	0.53(0.51)	0.306(1.04)
HIST2	0.603(0.25)	0.564(0.2)	0.319(0.16)	-0.121(0.21)	0.148(0.28)
HIST3	0.273(0.33)	-0.471(0.31)	0.151(0.2)	0.351(0.26)	0.248(0.34)
HIST4	0.182(0.37)	-2.496(0.74)	-2.405(0.73)	-1.592(0.63)	-0.307(0.49)
COHORT	0.123(0.18)	-0.35(0.16)	-0.316(0.12)	0.621(0.19)	-0.28(0.23)
CHEMO	0.252(0.19)	-0.115(0.15)	0.501(0.11)	0.042(0.19)	0.012(0.24)
MENO	0.27(0.26)	0.354(0.21)	-0.386(0.16)	0.204(0.28)	0.392(0.45)
RADIO	-1.026(0.26)	-0.297(0.18)	0.516(0.14)	-0.458(0.3)	-0.95(0.35)
SIDE	-0.046(0.16)	0.108(0.13)	0.005(0.09)	0.248(0.15)	-0.098(0.18)
LN(α)	-0.131(0.07)	-0.253(0.05)	-0.103(0.04)	0.097(0.06)	0.062(0.08)
CONSTANT	-9.87(0.85)	-8.87(0.67)	-6.918(0.53)	-13.578(0.85)	-16.966(1.14)
-2 Log Likelihood	3628.52	5227.95	8769.63	3758.80	2746.91

Table A.2: Weibull baseline hazard model without frailty for all failure types.

A.4 Non-parametric frailty

The parameters estimates of breast cancer data using non-parametric frailty using different number of mass points.

	Number of mass points				
Variable	One	Two	Three	Four	Five
AGE	0.030(0.008)	0.036(0.009)	0.047(0.014)	0.048(0.014)	0.048(0.014)
STAGE2	0.548(0.211)	0.530(0.221)	0.623(0.288)	0.626(0.290)	0.626(0.290)
STAGE3	0.726(0.289)	0.826(0.311)	1.124(0.407)	1.138(0.408)	1.138(0.408)
STAGE4	2.163(0.270)	3.427(0.321)	4.132(0.542)	4.168(0.539)	4.168(0.539)
SURGTYPE1	2.077(0.416)	2.307(0.464)	2.706(0.597)	2.719(0.599)	2.720(0.598)
SURGTYPE2	1.385(0.311)	1.388(0.326)	1.674(0.441)	1.684(0.443)	1.684(0.443)
SURGTYPE3	-0.055(0.296)	-0.141(0.309)	-0.201(0.394)	-0.207(0.397)	-0.207(0.397)
SURGTYPE4	-0.242(0.297)	-0.374(0.311)	-0.638(0.448)	-0.653(0.451)	-0.654(0.451)
SURGTYPE5	-17.873(5136)	-9.769(106.3)	-15.252(1151)	-12.646(339)	-12.646(335)
SURGTYPE6	-0.550(0.389)	-0.616(0.397)	-0.886(0.532)	-0.905(0.536)	-0.905(0.536)
SURGTYPE7	0.531(0.514)	0.521(0.538)	0.434(0.735)	0.424(0.744)	0.424(0.744)
HIST2	0.472(0.207)	0.403(0.219)	0.391(0.286)	0.387(0.289)	0.387(0.289)
HIST3	-1.471(0.613)	-1.411(0.619)	-1.713(0.708)	-1.726(0.709)	-1.726(0.709)
HIST4	0.121(0.210)	0.115(0.230)	0.059(0.280)	0.055(0.283)	0.055(0.283)
COHORT	0.621(0.192)	0.609(0.206)	0.711(0.251)	0.718(0.252)	0.718(0.252)
CHEMO	0.042(0.193)	0.099(0.199)	0.137(0.260)	0.142(0.263)	0.142(0.263)
MENO	0.204(0.285)	0.210(0.306)	0.329(0.390)	0.343(0.394)	0.343(0.394)
RADIO	-0.459(0.297)	-0.397(0.306)	-0.512(0.376)	-0.514(0.378)	-0.514(0.378)
SIDE	0.248(0.151)	0.239(0.163)	0.269(0.198)	0.270(0.199)	0.270(0.199)
LN(α)	0.097(0.060)	0.290(0.063)	0.508(0.132)	0.519(0.130)	0.519(0.130)
-2 Log Likelihood	3785.2368	3758.7958	3756.0952	3756.1028	3756.1026

Table A.3: Breast cancer Weibull hazard with non-parametric frailty using different number of mass points.

A.5 Log-Normal frailty

Table A.4: Breast cancer Weibull hazard with Log-normal frailty using different number of mass points.

	Number of mass points			
Variable	3 points	4 points	5 points	6 points
AGE	0.036(0.012)	0.046(0.017)	0.059(0.014)	0.042(0.013)
STAGE2	0.637(0.267)	0.649(0.312)	0.939(0.344)	0.694(0.333)
STAGE3	1.139(0.417)	1.047(0.393)	1.569(0.467)	1.140(0.429)
STAGE4	3.771(0.531)	3.764(0.524)	5.512(0.679)	4.034(0.576)
SURGTYPE1	3.006(0.617)	2.887(0.724)	3.888(0.626)	3.208(0.797)
SURGTYPE2	1.928(0.509)	1.704(0.435)	2.909(0.535)	2.150(0.566)
SURGTYPE3	-0.088(0.356)	-0.177(0.411)	-0.393(0.479)	-0.057(0.427)
SURGTYPE4	-0.392(0.363)	-0.611(0.477)	-0.326(0.448)	-0.411(0.426)
SURGTYPE5	-7.890(38.45)	-7.89(30.89)	-7.898(20.46)	-7.89(28.26)
SURGTYPE6	-0.698(0.431)	-0.895(0.534)	-1.118(0.555)	-0.865(0.519)
SURGTYPE7	0.524(0.589)	0.415(0.748)	0.869(0.665)	0.531(0.712)
HIST2	0.475(0.263)	0.404(0.300)	0.711(0.299)	0.490(0.321)
HIST3	-1.531(0.645)	-1.762(0.699)	-1.732(0.847)	-1.842(0.751)
HIST4	0.149(0.273)	0.004(0.298)	0.313(0.317)	0.049(0.307)
COHORT	0.756(0.248)	0.740(0.253)	0.963(0.293)	0.833(0.293)
CHEMO	0.115(0.226)	0.161(0.263)	0.104(0.274)	0.168(0.288)
MENO	0.230(0.350)	0.379(0.443)	0.748(0.428)	0.273(0.403)
RADIO	-0.450(0.338)	-0.552(0.379)	-0.210(0.465)	-0.534(0.408)
SIDE	0.394(0.192)	0.301(0.208)	0.537(0.224)	0.327(0.220)
CONSTANT	-18.189(1.824)	-19.87(2.765)	-25.93(2.781)	-20.96(2.628)
$LN(\alpha)$	0.386(0.098)	0.461(0.127)	0.714(0.106)	0.523(0.132)
$LN(\tau^2)$	0.452(0.177)	0.678(0.209)	1.043(0.124)	0.813(0.209)
-2 Log Likelihood	3805.78	3799.3574	3801.1342	3799.3808

Continued on next page

Continued from previous page

	Number of mass points			
Variable	7 points	8 points	9 points	16 points
AGE	0.041(0.017)	0.050(0.016)	0.039(0.013)	0.039(0.013)
STAGE2	0.703(0.326)	0.713(0.383)	0.673(0.308)	0.663(0.293)
STAGE3	1.154(0.874)	1.290(0.474)	1.038(0.440)	1.088(0.476)
STAGE4	4.455(1.883)	4.759(0.656)	4.014(0.999)	3.876(0.760)
SURGTYPE1	3.308(0.894)	4.177(1.282)	3.240(0.916)	2.999(0.764)
SURGTYPE2	1.757(0.478)	2.935(0.805)	1.848(0.524)	1.875(0.519)
SURGTYPE3	-0.109(0.400)	-0.096(0.522)	-0.073(0.385)	-0.085(0.382)
SURGTYPE4	-0.524(0.520)	-0.194(0.514)	-0.419(0.408)	-0.416(0.396)
SURGTYPE5	-7.892(30.265)	-7.902(20.196)	-7.891(32.579)	-7.890(33.451)
SURGTYPE6	-0.835(0.546)	-1.201(0.678)	-0.772(0.504)	-0.754(0.479)
SURGTYPE7	0.498(0.700)	0.686(0.708)	0.532(0.690)	0.514(0.667)
HIST2	0.512(0.349)	0.791(0.478)	0.487(0.287)	0.474(0.288)
HIST3	-1.766(0.791)	-2.112(0.962)	-1.716(0.737)	-1.694(0.709)
HIST4	0.032(0.324)	0.098(0.351)	0.030(0.316)	0.066(0.299)
COHORT	0.886(0.466)	1.130(0.457)	0.815(0.310)	0.811(0.295)
CHEMO	0.152(0.254)	0.230(0.319)	0.160(0.262)	0.135(0.247)
MENO	0.208(0.377)	0.451(0.489)	0.234(0.376)	0.255(0.373)
RADIO	-0.543(0.411)	-0.385(0.464)	-0.506(0.375)	-0.507(0.371)
SIDE	0.410(0.319)	0.399(0.281)	0.344(0.214)	0.351(0.212)
CONSTANT	-20.095(5.502)	-25.80(3.834)	-19.32(3.269)	-19.10(2.839)
$LN(\alpha)$	0.479(0.259)	0.720(0.162)	0.442(0.176)	0.426(0.148)
$LN(\tau^2)$	0.711(0.483)	1.124(0.215)	0.639(0.333)	0.614(0.281)
-2 Log Likelihood	3802.828	3799.447	3802.376	3802.405

Continued on next page

Continued from previous page

	Number of mass points		
Variable	32 points	64 points	128 points
AGE	0.042(0.015)	0.041(0.015)	0.041(0.015)
STAGE2	0.691(0.319)	0.685(0.313)	0.685(0.313)
STAGE3	1.154(0.514)	1.139(0.502)	1.139(0.502)
STAGE4	4.123(1.068)	4.070(0.986)	4.069(0.985)
SURGTYPE1	3.225(1.013)	3.174(0.928)	3.172(0.927)
SURGTYPE2	1.960(0.608)	1.941(0.582)	1.940(0.582)
SURGTYPE3	-0.092(0.398)	-0.092(0.395)	-0.092(0.394)
SURGTYPE4	-0.451(0.425)	-0.445(0.419)	-0.446(0.419)
SURGTYPE5	-7.891(30.966)	-7.891(31.484)	-7.891(31.485)
SURGTYPE6	-0.802(0.520)	-0.791(0.510)	-0.792(0.510)
SURGTYPE7	0.524(0.700)	0.522(0.695)	0.521(0.695)
HIST2	0.495(0.308)	0.489(0.303)	0.490(0.303)
HIST3	-1.755(0.760)	-1.741(0.745)	-1.740(0.745)
HIST4	0.062(0.315)	0.062(0.314)	0.062(0.310)
COHORT	0.850(0.327)	0.839(0.315)	0.839(0.315)
CHEMO	0.148(0.261)	0.145(0.258)	0.145(0.257)
MENO	0.278(0.395)	0.274(0.391)	0.275(0.391)
RADIO	-0.519(0.388)	-0.516(0.383)	-0.516(0.383)
SIDE	0.369(0.228)	0.364(0.224)	0.365(0.224)
CONSTANT	-20.14(4.320)	-19.92(3.940)	-19.91(3.934)
$LN(\alpha)$	0.477(0.211)	0.466(0.194)	0.466(0.194)
$LN(\tau^2)$	0.712(0.376)	0.693(0.352)	0.692(0.352)
-2 Log Likelihood	3802.189	3802.201	3802.202

Appendix B

Correlated frailty

B.1 Correlated Gamma frailty

The logarithm of bivariate survival function of the correlated frailty model and its partial derivatives

$$\begin{aligned}\ln S(t_1,t_2) = \; & \left(1-\rho\tfrac{\tau_1}{\tau_2}\right)\ln S_1(t_1) + \left(1-\rho\tfrac{\tau_2}{\tau_1}\right)\ln S_2(t_2) \\ & -\left(\tfrac{\rho}{\tau_1\tau_2}\right)\ln\left(S_1(t_1)^{-\tau_1^2} + S_2(t_2)^{-\tau_2^2} - 1\right)\end{aligned}$$

$$\begin{aligned}\tfrac{\partial}{\partial t_1}\ln S(t_1,t_2) = \; & -\left(1-\rho\tfrac{\tau_1}{\tau_2}\right)h_1(t_1) + \left\{\left(\tfrac{\rho}{\tau_1\tau_2}\right)\tau_1^2 f_1(t_1)S_1(t_1)^{-\tau_1^2-1}\right. \\ & \left.\times\left(S_1(t_1)^{-\tau_1^2} + S_2(t_2)^{-\tau_2^2} - 1\right)^{-1}\right\}\end{aligned}$$

$$\begin{aligned}\tfrac{\partial}{\partial t_2}\ln S(t_1,t_2) = \; & -\left(1-\rho\tfrac{\tau_2}{\tau_1}\right)h_2(t_2) + \left\{\left(\tfrac{\rho}{\tau_1\tau_2}\right)\tau_2^2 f_2(t_2)S_2(t_2)^{-\tau_2^2-1}\right. \\ & \left.\times\left(S_1(t_1)^{-\tau_1^2} + S_2(t_2)^{-\tau_2^2} - 1\right)^{-1}\right\}\end{aligned}$$

$$\frac{\partial^2}{\partial t_1 \partial t_2} \ln S(t_1,t_2) = \left(\frac{\rho}{\tau_1\tau_2}\right)\left(\tau_1^2 f_1(t_1) S_1(t_1)^{-\tau_1^2-1}\right)\left(\tau_2^2 f_2(t_2) S_2(t_2)^{-\tau_2^2-1}\right) \times\left(S_1(t_1)^{-\tau_1^2} + S_2(t_2)^{-\tau_2^2} - 1\right)^{-2}$$

B.2 Correlated Inverse Gaussian frailty

$$\ln S(t_1,t_2) = \left(1-\rho\frac{\tau_1}{\tau_2}\right)\ln S_1(t_1) + \left(1-\rho\frac{\tau_2}{\tau_1}\right)\ln S_2(t_2) +\left(\frac{\rho}{\tau_1\tau_2}\right)\left(1 - \left[\left(1-\tau_1^2\ln S_1(t_1)\right)^2 + \left(1-\tau_2^2\ln S_2(t_2)\right)^2 - 1\right]^{1/2}\right)$$

$$\frac{\partial}{\partial t_1}\ln S(t_1,t_2) = -\left(1-\rho\frac{\tau_1}{\tau_2}\right)h_1(t_1) - \left(\frac{\rho}{\tau_1\tau_2}\right)\tau_1^2 h_1(t_1)\left(1-\tau_1^2\ln S_1(t_1)\right) \left[\left(1-\tau_1^2\ln S_1(t_1)\right)^2 + \left(1-\tau_2^2\ln S_2(t_2)\right)^2 - 1\right]^{-1/2}$$

$$\frac{\partial}{\partial t_2}\ln S(t_1,t_2) = -\left(1-\rho\frac{\tau_2}{\tau_1}\right)h_2(t_2) - \left(\frac{\rho}{\tau_1\tau_2}\right)\tau_2^2 h_2(t_2)\left(1-\tau_2^2\ln S_2(t_2)\right) \left[\left(1-\tau_1^2\ln S_1(t_1)\right)^2 + \left(1-\tau_2^2\ln S_2(t_2)\right)^2 - 1\right]^{-1/2}$$

$$\frac{\partial^2}{\partial t_2 \partial t_1}\ln S(t_1,t_2) = -\left(\frac{\rho}{\tau_1\tau_2}\right)\tau_1^2 h_1(t_1)\left(1-\tau_1^2\ln S_1(t_1)\right)\tau_2^2 h_2(t_2)\left(1-\tau_2^2\ln S_2(t_2)\right) \left[\left(1-\tau_1^2\ln S_1(t_1)\right)^2 + \left(1-\tau_2^2\ln S_2(t_2)\right)^2 - 1\right]^{-3/2}$$

Appendix C

Gauss code

The software used to analysis the breast cancer data and in conducting the simulations are "Gauss" and "R". STATA was used only to check the codes written in Gauss and R for univariate frailty models. In multivariate case only Gauss program was used to fit the model. The optimisation procedure in Gauss showed robustness in reaching the maximum likelihood estimates, while the optimisation procedure in R was sensitive to the parameters' initial values. The list below is a sample of Gauss code used in the thesis.
GAUSS is commercial statistical software. GAUSS is a matrix programming language for mathematics and statistics developed by Aptech Systems. Its primary purpose is the solution of numerical problems in statistics, econometrics, time-series, optimization and 2D- and 3D-visualization.

R is a free software programming language and a software environment for statistical computing and graphics. The R language is widely used among statisticians and data miners for developing statistical software and data analysis.

STATA is commercial statistical software developed by StataCorp. Its capabilities include data management, statistical analysis, graphics and simulations

Listing C.1: Gauss code of univariate simulation of Log-Normal frailty

```
/*Simulation of Univariate survival time assuming weibull(alpha,lambda) for failure times@
and weibull(alpha,theta) for survival times */
new;
output file = Univariate_Simulation5000.out reset;
for ii(1,600,1);
n=5000;
const=ones(n,1);
alpha=1;
sigma=1;
b0=-4;
b1=1;
b2=-2;
b3=4;
b4=2;
x1=rndu(n,1);
u1=rndu(n,1);
x2=(u1 .<0.3);
u2=rndu(n,1);
x31=(u2 .<0.4);
x32=(0.4 .<= u2).*(u2 .<0.6);
z=sigma*rndn(n, 1);
xb=b0+b1*x1+b2*x2+b3*x31+b4*x32;
lamb=exp(xb+z);
u3=rndu(n,1);
lifetimes =(-ln(u3)./lamb)^(1/alpha);
theta = 0.01;
u4=rndu(n,1);
censtimes =(-ln(u4)/theta)^(1/alpha);
aa=lifetimes~censtimes;
stime = minc(aa');
status= (censtimes .> lifetimes);
y=stime~const~x1~x2~x31~x32~status;
let varname = time const x1 x2 x31 x32 status;
call dstat(0,y);
call dstat(0,z);
b_ =cols(y);
u={
```

```
-10.07742267422950,-9.06439921070241,-8.21972876538224,-7.46075575412152,-6.75593083054070,
-6.08896430907698, -5.45003327362342,-4.83260461324449,-4.23202110999540,-3.64478124988082,
-3.06813516901312,-2.49984041518739,-1.93800490592571,-1.38098019927214,-0.82728490377977,
-0.27554641923028,0.27554641923028,0.82728490377977,1.38098019927214,1.93800490592571,
2.49984041518739,3.06813516901312,3.64478124988082,4.23202110999540,4.83260461324449,
5.45003327362342,6.08896430907698,6.75593083054070,7.46075575412152,8.21972876538224,
  9.06439921070241,10.07742267422950
};
w={
4.124607489018270E-23,5.208449591960860E-19,6.755290223670070E-16,2.378064855777800E-13,
3.347501239801200E-11,2.312518412074240E-09,8.881290713105870E-08,2.059622103953430E-06,
3.055980306089630E-05,3.025570258170630E-04,2.062051051307880E-03,9.903461702320580E-03,
3.410984772609200E-02,8.534480827208050E-02,1.565389937575980E-01,2.117055698804790E-01,
2.117055698804790E-01,1.565389937575980E-01,8.534480827208050E-02,3.410984772609200E-02,
9.903461702320580E-03,2.062051051307880E-03,3.025570258170630E-04,3.055980306089630E-05,
2.059622103953430E-06,8.881290713105870E-08,2.312518412074240E-09,3.347501239801200E-11,
2.378064855777800E-13,6.755290223670070E-16,5.208449591960860E-19,4.124607489018270E-23
};

proc lwei(be,y);
local llikl,I;
    llikl=zeros(n,1);
        I=zeros(n,1);
I=sumr(
exp( y[.,b_].*
(be[b_-1,1]+y[.,2:b_-1]* be[1:b_-2,1] + u'*exp(be[b_,1])+(exp(be[b_-1,1])-1).* ln(y[.,1]))
- (Exp(y[.,2:b_-1] * be[1:b_-2,1] + u'*exp(be[b_,1]) ).* (y[.,1]^(exp(be[b_-1,1])))))
  ).*w'  );

llikl=ln(I + (I.==0.0).*1e-15);
   retp(llikl);
endp;

library maxlik;
#include maxlik.ext;
start={1,1,1,1,1,1,1};
maxset;
__title = "Simulation of Weibull with random effect using cholesky decomposition";
```

```
__output = 1000;

{x0,f,g,cov,ret}=maxlik(y,0,&lwei,start);
call maxprt(x0,f,g,cov,ret);
print "log-likelihood=" f*n;
endfor;
output off;
end;
```

Listing C.2: Gauss code of multivariate simulation of Log-Normal frailty

```
/*Simulation of Biivariate survival time assuming weibull(alpha,lambda1) for failure times
 of type1,weibull(alpha,lambda2) for failure times of type2 and weibull(alpha,theta)
 for survival times */
new;
output file = Bivariate_Simulation700.csv reset;
mm=700;
p_est=zeros(mm,11);
strr=zeros(mm,11);
for ii(1,mm,1);
print ii;
n=500;
const=ones(n,1);
alpha=0.5;
sigma1=0.7;
sigma2=1.2;
rho=0.8;

b10=-0.2;
b11=0.5;
b12=1;

b20=0.2;
b21=0.7;
b22=1;

x11=  rndu(n,1);
unif1=rndu(n,1);
x12=(unif1 .<0.3);
```

```
sigma=sigma1^2~rho*sigma1*sigma2|rho*sigma1*sigma2~ sigma2^2;
sigma;
z=rndmn((0~0)',sigma,n);
vcx(z);

xb1=b10+b11*x11+b12*x12;
xb2=b20+b21*x11+b22*x12;

unif3=rndu(n,1);
unif4=rndu(n,1);

lamb1=exp(xb1+z[.,1]);
lifetimes1 =(-ln(unif3)./lamb1)^(1/alpha);

lamb2=exp(xb2+z[.,2]);
lifetimes2 =(-ln(unif4)./lamb2)^(1/alpha);

theta = 0.3;
unif5=rndu(n,1);
censtimes =(-ln(unif5)/theta)^(1/alpha);

aa=lifetimes1~lifetimes2~censtimes;
stime = minc(aa');
status1 = (lifetimes1 .< censtimes).*(lifetimes1 .< lifetimes2)   ;
status2 = (lifetimes2 .< censtimes).*(lifetimes2 .< lifetimes1)   ;
cens =(status1.==0).*(status2.==0);

y=stime~const~x11~x12~status1~status2;
pr=cols(y);
b_ =pr-1;

u={
-10.07742267422950,-9.06439921070241,-8.21972876538224,-7.46075575412152,-6.75593083054070,
-6.08896430907698, -5.45003327362342,-4.83260461324449,-4.23202110999540,-3.64478124988082,
-3.06813516901312,-2.49984041518739,-1.93800490592571,-1.38098019927214,-0.82728490377977,
-0.27554641923028,0.27554641923028,0.82728490377977,1.38098019927214,1.93800490592571,
2.49984041518739,3.06813516901312,3.64478124988082,4.23202110999540,4.83260461324449,
5.45003327362342,6.08896430907698,6.75593083054070,7.46075575412152,8.21972876538224,
```

```
  9.06439921070241,10.07742267422950
};
w={
4.124607489018270E-23,5.208449591960860E-19,6.755290223670070E-16,2.378064855777800E-13,
3.347501239801200E-11,2.312518412074240E-09,8.881290713105870E-08,2.059622103953430E-06,
3.055980306089630E-05,3.025570258170630E-04,2.062051051307880E-03,9.903461702320580E-03,
3.410984772609200E-02,8.534480827208050E-02,1.565389937575980E-01,2.117055698804790E-01,
2.117055698804790E-01,1.565389937575980E-01,8.534480827208050E-02,3.410984772609200E-02,
9.903461702320580E-03,2.062051051307880E-03,3.025570258170630E-04,3.055980306089630E-05,
2.059622103953430E-06,8.881290713105870E-08,2.312518412074240E-09,3.347501239801200E-11,
2.378064855777800E-13,6.755290223670070E-16,5.208449591960860E-19,4.124607489018270E-23
};
e=ones(rows(u),1);
u2=u.*.e;
u1=e.*.u;
w2=w.*.e;
w1=e.*.w;

proc lwei(be,y);
local llikl,I;
 llikl=zeros(n,1);
     I=zeros(n,1);
I=sumr(
exp(
 y[.,b_].*(be[b_-1,1] +y[.,2:b_-1]* be[1:b_-2,1] + u1'*be[b_,1]+ u2'*be[2*b_+1,1]
+(exp(be[b_-1,1])-1).* ln(y[.,1]))
-(Exp(y[.,2:b_-1]*be[1:b_-2,1]+u1'*be[b_,1]+u2'*be[2*b_+1,1] ).*(y[.,1]^(exp(be[b_-1,1]))))
   +
 y[.,b_+1].* (be[2*b_-1,1] +  y[.,2:b_-1]* be[b_+1:2*b_-2,1] + u2'*be[2*b_,1]
+(exp(be[2*b_-1,1])-1).* ln(y[.,1]))
-(Exp(y[.,2:b_-1]* be[b_+1:2*b_-2,1] + u2'*be[2*b_,1] ).* (y[.,1]^(exp(be[2*b_-1,1]))))

    )
.*w1'.*w2'  );

llikl=ln(I + (I.==0.0).*1e-15);
  retp(llikl);
endp;
```

```
105
106 library maxlik;
107 #include maxlik.ext;
108 start={1,1,1,1,1,1,1,1,1,1,1};
109 maxset;
110 __title = "Bivariate Simulation of Weibull with random effect";
111 __output = 1000;
112 {x0,f,g,cov,ret}=maxlik(y,0,&lwei,start);
113 call maxprt(x0,f,g,cov,ret);
114 print "log-likelihood=" f*n;
115 p_est[ii,.]=x0';
116 if rows(cov) == 11;
117 strr[ii,.]=sqrt(diag(cov))';
118 endif;
119 endfor;
120 output off;
121 output file = Parameters_Bivariate_Simulation700.out reset;
122 format /m1 8,4;
123 outwidth 132;
124 print p_est;;
125 print strr;
126 output off;
127 end;
```

Bibliography

Aalen, O. O. (1988), 'Heterogeneity in survival analysis', *Statistics in Medicine.* **7**, 1121–1137.

Aalen, O. O. (1992), 'Modelling heterogeneity in survival analysis by the compound poisson distribution', *Annals of Applied Probability* **4**, 951 – 972.

Abbring, J. H. and van den Berg, G. J. (2003), 'The identifiability of the mixed proportional hazards competing risks model', *Royal Statistical Society, Series B* **65**(3), 701 – 710.

Abbring, J. H. and van den Berg, G. J. (2007), 'The unobserved heterogeneity distribution in duration analysis', *Biometrika* **94**(1), 87–99.

Aitkin, M. (1999), 'A general maximum likelihood analysis of variance components in generalized linear models', *Biometrics* **55**(1), 117–128.

Aitkin, M., Francis, B., Hinde, J. and Darnell, R. (2009), *Statistical Modelling in R*, Oxford statistical science series, Oxford University Press, Oxford.

Akaike, H. (1974), 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control* **19**, 716723.

Anderson, J. E. and Louis, T. A. (1995), 'Survival analysis using a scale change random effects model', *Journal of the American Statistical Association* **90**, 669 – 679.

Balakrishnan, N. and Peng, Y. W. (2006), 'Generalized gamma frailty model', *Statistics in Medicine* **25**(16), 2797–2816.

Bandeen-Roche, K. and Liang, K. Y. (2002), 'Modelling multivariate failure time associations in the presence of a competing risk', *Biometrika* **89**(2), 299 – 314.

Beamonte, E. and Bermdez, J. D. (2003), 'A bayesian semiparametric analysis for additive hazard models with censored observations', *TEST* **12**(2), 347 – 363.

BermDez, L. and Karlis, D. (2012), 'A finite mixture of bivariate poisson regression models with an application to insurance ratemaking', *Computational Statistics and Data Analysis* **56**(12), 3988–3999.

Bock, R. and Aitkin, M. (1981), 'Marginal maximum likelihood estimation of item parameters: Application of an em algorithm', *Psychometrika* **46**(4), 443–459.

Cai, B. (2010), 'Bayesian semiparametric frailty selection in multivariate event time data', *Biometrical Journal* **52**(2), 171–185.

Cai, J. and Zeng, D. (2011), 'Additive mixed effect model for clustered failure time data', *Biometrics* **67**(4), 1340–1351.

Casella, G. and Berger, R. (2002), *Statistical inference*, Thomson Learning.

Chakraborty, H., Helms, R. W., Sen, P. K. and Cohen, M. S. (2003), 'Estimating correlation by using a general linear mixed model: evaluation of the relationship between the concentration of hiv-1 rna in blood and semen', *Statistics in Medicine* **22**, 1457 – 1464.

Chang, S. H. (2004), 'Estimating marginal effects in accelerated failure time models for serial sojourn times among repeated events', *Lifetime Data Analysis* **10**, 175 – 190.

Chen, D. G. and Lio, Y. L. (2008), 'Comparative studies on frailties in survival analysis', *Communications in Statistics-Simulation and Computation* **37**(8), 1631–1646.

Clayton, D. G. (1978), 'A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence', *Biometrika* **65**, 141 – 151.

Clayton, D. G. (1988), 'The analysis of event history data - a review of progress and outstanding problems', *Statistics in Medicine* **7**(8), 819–841.

Clayton, D. G. and Cuzick, J. (1985), 'Multivariate generalizations of the proportional hazards model (with discussion)', *Journal of the Royal Statistical Society, Series A* **148**(2), 82 – 117.

Congdon, P. (1995), 'Modelling frailty in area mortality', *Statistics in Medicine* **14**, 1859 –1874.

Cox, D. R. (1972), 'Regression models and life-tables', *Journal of the Royal Statistical Society. Series B (Methodological).* **34(2)**, 187–220.

Cox, D. R. and Oakes, D. V. (1984), *Analysis of Survival Data*, Chapman and Hall, London.

Crowder, M. (1985), 'A distributional model for repeated failure time measurements', *Royal Statistical Society, Series B* **47**, 447 – 452.

David, I., Lorino, T. and Sanaa, M. (2007), 'Bayesian versus frequentist approach of the frailty cox model, application to calf gastroenteritis', *Communications in Statistics - Simulation and Computation* **36**(6), 1309–1320.

Davies, R. B. (1993), 'Nonparametric control for residual heterogeneity in modelling recurrent behaviour', *Computational Statistics and Data Analysis* **16**(2), 143–160.

Davies, R. B. and Crouchley, R. (1984), 'Calibrating longitudinal models of residential-mobility and migration - an assessment of a non-parametric marginal likelihood approach', *Regional Science and Urban Economics* **14**(2), 231–247.

Dewanji, A. (1992), 'A note on a test for competing risks with missing failure type', *Biometrika* **79**(4), 855–857.

dos Santos, D. M., Davies, R. B. and Francis, B. (1995), 'Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer', *Journal of Statistical Planning and Inference* **47**, 111 – 127.

Duchateau, L. and Janssen, P. (2008), *The Frailty Model*, Springer.

Economou, P. and Caroni, C. (2005), 'Graphical tests for the assumption of gamma and inverse gaussian frailty distributions', *Lifetime Data Anal* **11**(4), 565–82.

Elbers, C. and Ridder, G. (1982), 'True and spurious duration dependence - the identifiability of the proportional hazard model', *Review of Economic Studies* **49**(3), 403–409.

Everitt, B. and Hand, D. (1981), *Finite Mixture Distributions*, Chapman and Hall.

Fahrmeir, L. and Tutz, G. (1994), *Multivariate Statistical Modelling based on Generalised Linear Models*, Springer-Verlag, New York.

Fieuws, S. and Verbeke, G. (2004), 'Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach', *Statistics in Medecine* **23**, 3093 – 3104.

Fieuws, S. and Verbeke, G. (2006a), 'Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles', *Biometrics* **62**, 424 – 431.

Fieuws, S., Verbeke, G., Boen, F. and Delecluse, C. (2006b), 'High dimensional multivariate mixed models for binary questionnaire data', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **55**(4), 449 – 460.

Fieuws, S., Verbeke, G. and Molenberghs, G. (2007), 'Random-effects models for multivariate repeated measures', *Statistical Methods in Medical Research* **16**, 387 – 397.

Fine, J. P. and Gray, R. J. (1999), 'A proportional hazards model for the subdistribution of a competing risk', *Journal of the American Statistical Association* **944**, 496 – 509.

Fine, J. P., Jiang, H. and Chappell, R. (2001), 'On semi-competing risks data', *Biometrika* **88**(4), 907 – 919.

Finkelstein, M. and Esaulovac, V. (2006), 'Asymptotic behavior of a general class of mixture failure rates', *The Advances in Applied Probability* **38**(1), 244 – 262.

Finkelstein, M. and Esaulovac, V. (2008), 'On asymptotic failure rates in bivariate frailty competing risks models', *Statistics and Probability Letters* **78**, 1174 – 1180.

Frhwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, Springer Science + Business Media, LLC.

Gelfand, A. E., Ghosh, S. K., Christiansen, C., Soumerai, S. B. and McLaughlin, T. J. (2000), 'Proportional hazards models: a latent competing risk approach', *Applied statistics* **49**(3), 385 – 397.

Giard, N., Lichtenstein, P. and Yashin, A. I. (2002), 'A multistate model for the genetic analysis of the ageing process', *Statistics in Medicine* **21**(17), 2511–26.

Gill, R. D. (1989), 'Non-parametric and semi-parametric maximum-likelihood estimators and the von mises method .1', *Scandinavian Journal of Statistics* **16**(2), 97–128.

Goetghebeur, E. and Ryan, L. (1995), 'Analysis of competing risks survival data when some failure types are missing', *Biometrika* **82**(4), 821–833.

Gueorguieva, R. (2001), 'A multivariate generalized linear mixed model for joint modelling of clustered outcomes in the exponential family', *Statistical Modelling* **1**(3), 177 – 193.

Gustafson, P. (1997), 'Large hierarchical baysian analysis of multivariate survival data', *Biometrics* **53**, 230 – 242.

Hall, D. B. and Wang, L. (2005), 'Two-component mixtures of generalized linear mixed effects models for cluster correlated data', *Statistical Modelling* **5**(1), 21–37.

Han, A. and Hausman, J. A. (1990), 'Flexible parametric estimation of duration and competing risk models', *Journal of Applied Econometrics* **5**, 1 – 28.

Hanagal, D. D. (2008), 'Frailty regression models in mixture distributions', *Journal of Statistical Planning and Inference* **138**(8), 2462–2468.

Heckman, J. J. and Singer, B. (1982a), *The identification problem in econometric models for duration data*, Cambridge University Press, Cambridge, pp. 39–77.

Heckman, J. J. and Singer, B. (1984), 'A method for minimising the impact of distributional assumptions in econometric models for duration data', *Econometrica* **52**, 271 – 320.

Heckman, J. J. and Singer, B. (1985), 'Social science duration analysis in longitudinal analysis of labor market data', *Cambridge University Press* pp. 39 – 110.

Henderson, R. and Oman, P. (1999), 'Effect of frailty on marginal regression estimates in survival analysis', *Journal of the Royal Statistical Society Series B-Statistical Methodology* **61**, 367–379. Part 2.

Hougaard, P. (1984), 'Life table methods for heterogeneous populations', *Biometrika* **71**, 75 – 83.

Hougaard, P. (1986a), 'Survival models for heterogeneous populations derived from stable distributions', *Biometrika* **73**, 387 – 396.

Hougaard, P. (1986b), 'A class of multivariate failure time distributions', *Biometrika* **73**, 671 – 678.

Hougaard, P. (1995), 'Frailty models for survival data', *Lifetime Data Analysis* **1**, 255 – 273.

Hougaard, P. (2000), *Analysis of Multivariate Survival Data*, Springer, New York.

Hsu, Y. W. (2000), 'On the bock-aitkin procedure - from an em algorithm perspective', *Psychometrika* **65**(4), 547–549.

Huang, X. and Wolfe, R. A. (2002), 'A frailty model for informative censoring', *Biometrics* **58**, 510 – 520.

Hudgens, M. G., Satten, G. A. and Longini, J. I. M. (2001), 'Nonparametric maximum likelihood estimation for competing risks survival data subject to interval censoring and truncation', *Biometrics* **57**, 74 – 80.

Jeong, J. H. and Oakes, D. (2005), 'Effects of different hazard ratios on asymptotic relative efficiency of estimates from cox's model', *Communications in Statistics-Theory and Methods* **34**(2), 429–448.

Jewell, N. P., Laan, M. V. D. and Henneman, T. (2003), 'Nonparametric estimation from current status data with competing risks', *Biometrika* **90**(1), 183 – 197.

Jiang, H., Fine, J. and Chappell, R. J. (2004), 'Semiparametric methods for semi-competing risks problem with censoring and truncation', *Harvard University Biostatistics Working Paper Series. Working Paper 15* .

Jonker, M. A., Bhulai, S., Boomsma, D. I., Ligthart, R. S. L., Posthuma, D. and Vaart, A. W. V. D. (2009), 'Gamma frailty model for linkage analysis with application to interval-censored migraine data', *Biostatistics* **10**(1), 187–200.

Kalbfleisch, J. D. and Prentice, R. L. (2002), *The statistical analysis of failure time data*, second edn, Wiley-Interscience.

Kaplan, E. L. and Meier, P. (1958), 'Nonparametric estimation from incomplete observations', *Journal of the American Statistical Association* **53**, 457 – 481.

Keiding, N., Andersen, P. and Klein, J. (1997), 'The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates', *Statistics in Medicine* **16**, 215 – 224.

Kheiri, S., Kimber, A. and Meshkani, M. R. (2007), 'Bayesian analysis of an inverse gaussian correlated frailty model', *Computational Statistics and Data Analysis* **51**(11), 5317–5326.

Kheiri, S., Meshkani, M. R. and Faghihzadeh, S. (2005), 'A correlated frailty model for analysing risk factors in bilateral corneal graft rejection for keratoconus: a bayesian approach', *Statistics in Medicine* **24**(17), 2681–2693.

Klein, J. P. (1992b), 'Semiparametric estimation of random effects using the cox model based on the em algorithm', *Biometrics* **48**, 795806.

Klein, J. P. and Moeschberger, M. L. (1997), *Survival analysis - techniques for censored and truncated data*, Springer, New York.

Klein, J. P., Moeschberger, M., Li, Y. H. and Wang, S. T. (1992a), *Estimating random effects in the Framingham Heart Study*, Kluwer Academic Publishers, pp. 99 – 120.

Klein, J. P., Pelz, C. and Zhang, M. J. (1999), 'Random effects for censored data by a multivariate normal regression model', *Biometrics* **55**, 497 – 506.

Korsgaard, I. R. and Andersen, A. H. (1998), 'The additive genetic gamma frailty model', *Scandinavian Journal of Statistics* **25**, 255 – 269.

Kundu, D. (2004), 'Parameter estimation for partially complete time and type of failure data', *Biometrical Journal* **46**(2), 165–179.

Laird, N. (1978), 'Nonparametric maximum likelihood estimation of a mixing distribution', *Journal of the American Statistical Association* **73**(364), 805–811.

Lam, K. F. and Kuk, A. Y. C. (1997), 'A marginal likelihood approach to estimation in frailty models', *Journal of the American Statistical Association* **92**(439), 985–990.

Lambert, P. and Collett, C., eds (2002), *Shared frailty accelerated failure time models for clustered survival data*, International Biometric Society, Freiburg, Germany.

Lambert, P., Collett, D., Kimber, A. and Johnson, R. (2004), 'Parametric accelerated failure time models with random effects and an application to kidney transplant survival', *Statistics in Medicine* **23**(20), 3177 – 3192.

Lancaster, T. and Nickell, S. (1980), 'The analysis of re-employment probabilities for the unemployed', *Journal of the Royal Statistical Society. Series A* **143**(2), 141–165.

Lawless, J. F. (1982), *Statistical models and methods for lifetime data*, Wiley and Sons, New York.

Lee, E. and Wang, J. (2003), *Statistical Methods for Survival Data Analysis*, J. Wiley.

Li, H. (2002), 'An additive genetic gamma frailty model for linkage analysis of diseases with variable age of onset using nuclear families', *Life time Data Analysi* **8**, 315 – 334.

Li, Y. and Lin, X. (2000), 'Covariate measurement errors in frailty model for clustered survival data', *Biometrics* **87**, 849866.

Lillard, L. A. (1993), 'Simultaneous equations for hazards: marriage duration and fertility timing', *Journal of Econometrics* **56**, 189 – 217.

Lillard, L. A., Brian, M. J. and Waite, M. J. (1995), 'Premarital cohabitation and subsequent marital dissolution: a matter of self-selection?', *Demography* **32**, 437 – 457.

Lin, D. Y. and Ying, Z. L. (1994), 'Semiparametric analysis of the additive risk model', *Biometrika* **81**(61 - 71).

Lin, D. Y. and Ying, Z. L. (1995), 'Semiparametric analysis of general additive-multiplicative hazard models for counting process', *Annals of Statistics* **23**, 1712 – 1734.

Locatelli, I., Lichtenstein, P. and Yashin, A. I. (2004), 'The heritability of breast cancer: A bayesian correlated frailty model applied to swedish twins data', *Twin Research* **7**(2), 182–191.

Locatelli, I., Rosina, A., Lichtenstein, P. and Yashin, A. I. (2007), 'A correlated frailty model with long-term survivors for estimating the heritability of breast cancer', *Statistics in Medicine* **26**(20), 3722–3734.

Lu, K. and Tsiatis, A. A. (2005), 'Comparison between two partial likelihood approaches for the competing risks model with missing cause of failure', *Lifetime Data Analysis* **11**, 29 – 40.

Lunn, M. and McNeil, D. (1995), 'Applying cox regression to competing risks', *Biometrics* **51**(2), 524 – 532.

Manda, S. O. M. (2011), 'A nonparametric frailty model for clustered survival data', *Communications in Statistics-Theory and Methods* **40**(5), 863–875.

Manton, K., Stallard, E. and Vaupel, J. (1986), 'Alternative models for heterogeneity of mortality risks among the aged', *Journal of the American Statistical Association* **81**, 635 – 644.

McCulloch, C. E. and Searle, S. R. (2001), *Generalized, Linear, and Mixed Models*, Wiley-Interscience, New York.

McGilchrist, C. A. (1993), 'Reml estimation for survival models with frailty', *Biometrics* **49**, 221 – 225.

McGilchrist, C. A. and Aisbett, C. W. (1991), 'Regression with frailty in survival analysis', *Biometrics* **47**, 461 – 466.

McLachlan, G. J. and Krishnan, T. (2008), *The EM Algorithm and Extensions*, Wiley series in probability and statistics, New Jersey.

Miller, R. G. (1981), *Survival analysis*, John Wiley and Sons, New York.

Minami, M. (2003), 'A multivariate extension of inverse gaussian distribution derived from inverse relationship', *Communications in Statistics-Theory and Methods* **32**(12), 2285–2304.

Molenberghs, G. and Verbeke, G. (2011), 'On the weibull-gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, cauchy, and extreme-value distributions', *Journal of Statistical Planning and Inference* **141**(2), 861–868.

Muller, K. E. and Stewart, P. W. (2006), *Linear Model Theory Univariate, Multivariate, and Mixed Models*, John Wiley and Sons, Hoboken, New Jersey.

Naskar, M. (2008), 'Semiparametric analysis of clustered survival data under nonparametric frailty', *Statistica Neerlandica* **62**(2), 155–172.

Naskar, M., Das, K. and Ibrahim, J. G. (2005), 'A semiparametric mixture model for analyzing clustered competing risks data', *Biometrics* **61**, 729 – 737.

Oakes, D. (1982), 'A concordance test for independence in the presence of censoring', *Biometrics* **38**, 451 – 455.

Olsen, M. K. and Schafer, J. L. (2001), 'A two-part random effects model for semicontinuous longitudinal data', *Journal of the American Statistical Association* **96**(454), 730–745.

Oskrochi, G. R. and Crouchley, R. (2004), Modelling breast cancer data with informative dropout, *in* 'Proceedings of the 19th International Workshop on Statistical Modelling', Firenze University Press, Florence, Italy.

Oskrochi, G. R. and Davies, R. B. (1997), 'An em-type algorithm for multivariate mixture models', *Statistics and Computing* **7**(2), 145 – 151.

Pan, W. (2001), 'Using frailties in the accelerated failure time model', *Lifetime Data Analysis* **7**, 55 – 64.

Pankratz, V. S., de Andrade, M. and Therneau, T. M. (2005), 'Random-effects cox proportional hazards model: General variance components methods for time-to-event data', *Genetic Epidemiology* **28**(2), 97–109.

Peng, Y. W. and Zhang, J. J. (2008), 'Estimation method of the semiparametric mixture cure gamma frailty model', *Statistics in Medicine* **27**(25), 5177–5194.

Pennell, M. L. and Dunson, D. B. (2006), 'Bayesian semiparametric dynamic frailty models for multiple event time data', *Biometrics* **62**(4), 1044–1052.

Perperoglou, A., van Houwelingen, H. C. and Henderson, R. (2006), 'A relaxation of the gamma frailty (burr) model', *Statistics in Medicine* **25**(24), 4253–4266.

Peterson, J. (1998), 'An additive frailty model for correlated life times', *Biometrics* **54**(646 - 661).

Pipper, C. B. and Martinussen, T. (2004), 'An estimating equation for parametric shared frailty models with marginal additive hazards', *Royal Statistical Society. Series B* **66**(1), 207 – 220.

Prentice, R. L., Kalbfleisch, J. D., Peterson, J. A. V., Flournoy, N., Farewell, V. T. and Breslow, N. E. (1978), 'The analysis of failure times in the presence of competing risks', *Biometrics* **34**(4), 541 – 554.

Price, D. L. and Manatunga, A. K. (2001), 'Modelling survival data with a cured fraction using frailty models', *Statistics in Medicine* **20**, 1515 – 1527.

Ravishanker, N. and Dey, D. (2000), 'Multivariate survival models with a mixture of positive stable frailties', *Methodology and Computing in Applied Probability* **2**(3), 293–308.

Richardson, S. and Green, P. J. (1997), 'On bayesian analysis of mixtures with an unknown number of components (with discussion)', *Journal of the Royal Statistical Society, Series B* **59**, 731 – 792.

Ripatti, S., Larsen, K. and Palmgren, J. (2002), 'Maximum likelihood inference for multivariate frailty models using an automated mcem algorithm', *Lifetime Data Analysis* **8**, 349 – 360.

Ripatti, S. and Palmgren, J. (2000), 'Estimation of multivariate frailty models using penalised partial likelihood', *Biometrics* **56**, 1016 – 1022.

Sahu, S. K., Dey, D. K., Aslanidou, H. and Sinha, D. (1997), 'A weibull regression model with gamma frailties for multivariate survival data', *Lifetime Data Anal* **3**(2), 123–37.

Sastry, N. (1997), 'A nested frailty model for survival data, with an application to the study of child survival in northeast brazil', *Journal of the American Statistical Association* **92**, 426 – 435.

Scallan, A. J. (1987), 'A glim model for repeated measurements', *GL1M Newsletter* **15**, 10 – 22.

Schwarz, G. E. (1978), 'Estimating the dimension of a model', *Annals of Statistics* **6**, 461464.

Shih, J. H. and Louis, T. A. (1995), 'Assessing gamma frailty models for clustered failure time data', *Lifetime Data Analysis* **1**, 205 – 220.

Sinha, D. (1993), 'Semiparametric bayesian-analysis of multiple event time data', *Journal of the American Statistical Association* **88**(423), 979–983.

Slud, E. V., Byar, D. P., Schatzkin, A., Prentice, R. and Kalbfleisch, J. (1988), 'Dependent competing risks and the latent-failure model', *Biometrics* **44**(4), 1203 – 1205.

Stefanescu, C. and Turnbull, B. W. (2006), 'Multivariate frailty models for exchangeable survival data with covariates', *Technometrics* **48**(3), 411–417.

Thum, Y. M. (1997), 'Hierarchical linear models for multivariate outcomes', *Journal of Educational and Behavioral Statistics* **22**(1), 77 – 108.

Vaida, F. and Xu, R. (2000), 'Proportional hazards models with random effects', *Statistics in Medicine* **19**(24), 3309 – 3324.

van Duijn, M. A. J. and Bockenholt, U. (1995), 'Mixture models for the analysis of repeated count data', *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **44**(4), 473 – 485.

Vaupel, J. W. (1990), 'Relative risks: frailty model of life history data', *Theoretical Population Biology* **37**, 220 – 234.

Vaupel, J. W., Manton, K. G. and Stallard, E. (1979), 'The impact of heterogeneity in individual frailty on the dynamics of mortality', *Demography* **16**, 439 – 454.

Vaupel, J. W. and Yashin, A. I. (1985), 'The deviant dynamics of death in heterogeneous populations', *Sociological Methodology* **15**, 179 – 211.

Verbeke, G. and Molenberghs, G. (2000), *Linear Mixed Models for longitudinal data*, Springer, New York.

Weibull, W. (1939), 'A statistical theory of the strength of material', *Proc. Roy. Swedish Inst. Eng. Res.* **151**(1).

West, B. T., Welch, K. B., Galecki, A. T. and Gillespie, B. W. (2007), *Linear Mixed Models: A Practical Guide Using Statistical Software*, Chapman and Hall/CRC, New York.

Whitmore, G. A. and Lee, M. L. T. (1991), 'A multivariate survival distribution generated by an inverse gaussian mixture of exponentials', *Technometrics* **33**, 39 – 50.

Wienke, A. (2007), Frailty Models in Survival Analysis, PhD thesis, Martin-Luther-Universit at Halle-Wittenberg.

Wienke, A. (2010a), *Frailty Models in Survival Analysis*, Chapman and Hall/CRC, New York.

Wienke, A., Arbeev, K. G., Locatelli, I. and Yashin, A. I. (2005), 'A comparison of different bivariate correlated frailty models and estimation strategies', *Mathematical Biosciences* **198**(1), 1–13.

Wienke, A., Christensen, K., Skytthe, A. and Yashin, A. (2002), 'Genetic analysis of cause of death in a mixture model of bivariate lifetime data', *Statistical Modelling* **2**(2), 89–102.

Wienke, A., Lichtenstein, P. and Yashin, A. I. (2003), 'A bivariate frailty model with a cure fraction for modeling familial correlations in diseases', *Biometrics* **59**(4), 1178–1183.

Wienke, A., Ripatti, S., Palmgren, J. and Yashin, A. (2010b), 'A bivariate survival model with compound poisson frailty', *Statistics in Medicine* **29**(2), 275–283.

Xu, L. and Zhang, J. (2009), 'An alternative estimation for the accelerated failure time mixture cure model', *Communication in Statistics* **38**, 1980 – 1990.

Xu, L. and Zhang, J. (2010), 'Em-like algorithm for the semiparametric accelerated failure time gamma frailty model', *Computational Statistics and Data Analysis* **54**, 1467 – 1474.

Xue, X. and Brookmeyer, R. (1996), 'Bivariate frailty model for the analysis of multivariate survival time', *Lifetime Data Analysis* **2**(3), 277 – 290.

Xue, X. and Ding, Y. (1999), 'Assessing heterogeneity and correlation of paired failure times with the bivariate frailty model', *Statistics in Medicine* **18**(8), 907–918.

Yashin, A. I. and Iachine, I. (1999), 'Dependent hazards in multivariate survival problems', *Journal of Multivariate Analysis* **71**, 241 – 261.

Yashin, A. I. and Iachine, I. A. (1995a), 'Genetic-analysis of durations - correlated frailty model applied to survival of danish twins', *Genetic Epidemiology* **12**(5), 529–538.

Yashin, A. I. and Iachine, I. A. (1995b), 'Survival of related individuals: an extension of some fundamental results of heterogeneity analysis', *Mathematical population studies* **5**(4), 321–377.

Yashin, A. I., Vaupel, J. W. and Iachine, I. A. (1995), 'Correlated individual frailty: An advantageous approach to survival analysis of bivariate data', *Mathematical Population Studies* **5**, 145 – 159.

Yin, G. (2007), 'Model checking for additive hazards model with multivariate survival data', *Journal of Multivariate Analysis* **98**(5), 1018–1032.

Yin, G. (2008), 'Bayesian transformation cure frailty models with multivariate failure time data', *Statistics in Medicine* **27**(28), 5929–5940.

Yin, G. and Ibrahim, J. G. (2005), 'A class of bayesian shared gamma frailty models with multivariate failure time data', *Biometrics* **61**, 208 – 216.

Yu, B. B. (2006), 'Estimation of shared gamma frailty models by a modified em algorithm', *Computational Statistics & Data Analysis* **50**(2), 463–474.

Zahl, P. H. (1997), 'Frailty modelling for the excess hazard', *Statistics in Medecine* **16**, 1573 – 1585.

Zhang, J. and Peng, Y. (2007), 'An alternative estimation method for the accelerated failure time frailty model', *Computational Statistics and Data Analysis* **51**(9), 4413–4423.

Zhong, X. and Li, H. (2004), 'Score tests of genetic association in the presence of linkage based on the additive genetic gamma frailty model', *Biostatistics* **5**(2), 307 – 327.