Modelling and Analysis of Bivariate Lifetime Data using Reversed Hazard Rates

Thesis submitted to the
Cochin University of Science and Technology
for the Award of Degree of
DOCTOR OF PHILOSOPHY
under the Faculty of Science

by

Gleeja V. L.

Department of Statistics
Cochin University of Science and Technology
Cochin- 682 022

May 2008

Certificate

Certified that the thesis entitled *'Modelling and Analysis of Bivariate Lifetime Data using Reversed Hazard Rates'* is a bonafide record of works done by Miss. Gleeja V. L. under my guidance in the Department of Statistics, Cochin University of Science and Technology, Cochin-22, Kerala, India and that no part of it has been included anywhere previously for the award of any degree or title.

Cochin-22
09 May 2008

Dr. P. G. Sankaran
(Supervising Guide)

Contents

Chapter One

Preliminaries

1.1 Introduction

Lifetime data is a term used for describing data that measures time to occurrence of some event. The event may be death, appearance of some disease, relapse from remission, equipment breakdown etc. The development of models and methods to deal with lifetimes took place in the second half of the twentieth century. The development proceeded into two main inter mingling streams; viz reliability theory and survival analysis. The reliability theory concerns with models for lifetimes of components and systems in the engineering and industrial fields and the survival analysis concerns with medical and similar biological phenomena.

Lifetime or time to event is usually considered as a positive real valued random variable having a continuous distribution function. The definition of lifetime includes a time scale and time origin, as well as specification of the event that determines lifetime. In some instances, time may represent age, with the time origin as the birth of the individual. In other instances, the natural time origin may be occurrence of some event such as entry into a study or diagnosis of a particular disease. In some situations, it is difficult to say precisely when the event occurs, for example, the case of appearance of tumour. The time scale is not always real or chronological time, especially where machines or equipments are considered. It could be the number of operations a component performs before it breaks down. The following examples illustrate various types of lifetime data that arise in practical situations.

1. Manufactured items with mechanical or electric components are often subjected to life tests in order to obtain information on their durability. This involves putting items in operation, often in laboratory setting and observing them until they fail. It is common here to refer to the lifetimes as 'failure times' since when an item ceases operating satisfactorily, it is said to have 'failed'.

2. In medical studies dealing with potential fatal diseases, one is interested in the survival time of individuals with disease, measured from the date of diagnosis or some other starting point. For example, it is common to compare treatments for a disease at least partly in terms of survival time distributions of patients receiving the different treatments.

3. A standard experiment in the investigation of carcinogenic substance is one in which laboratory animals are subjected to doses of the substance and then observed to see if they develop tumours. The main variable of interest is the time to appearance of a tumour, measured from when the dose is administered.

4. In remission period of leukemia patients, the patient, though not free of disease, is free of symptoms. The length of remission period is a variable of interest in this study. The patients in the state of remission are followed over time to see how long they stay in remission.

In survival studies, many subjects fail to continue to be in the study till the event of interest occurs. This leads to incomplete data due to censored observations. The analysis of lifetime data under censoring is a major issue in survival studies.

1.2 Censoring

Censoring is inevitable in survival and reliability studies because the experimenter is unable to obtain complete information on lifetime of individuals. For example, patients in a clinical trial may withdraw from the study, or the study

may have to be terminated at a prefixed time point. There are various categories of censoring such as right censoring, left censoring and interval censoring.

1.2.1 Right Censoring

In both engineering and medical applications, right censoring is the most common form of censoring with lifetime data. In right censoring only lower bounds on lifetime are available for some individuals. Right censoring arises in certain situations because some individuals are still surviving at the time that study is terminated. In other instances, individual may move away from the study area for reasons unconnected with the study, so contact is lost. In some other situations, individuals may be withdrawn or decide to withdraw from the study because of worsening or improving prognosis.

Two types of right censoring are built into the design of experiments to reduce the time taken for completing the study.

Type I censoring: Sometimes experiments run over a fixed time period in such a way that an individual lifetime will be known exactly only if it is less than a predetermined value. In such situations the data are said to be Type I or time censored. For example, in a life test experiment n items are simultaneously put into operation. The study is terminated at a predetermined time t_0. Suppose that r items are failed by this time and the remaining $n-r$ items are operative. Then there are $n-r$ censored items and the data consist of lifetimes of r failed items and the censoring time t_0 for the remaining $n-r$ items. Type I censoring occurs frequently in medical research when a decision is made to terminate a study at a date on which not all individual's lifetime will be known.

Type II censoring: The term type II censoring refers to the situation where n individuals start on study at the same time, and the study terminates once k lifetimes have been observed. Thus only the smallest k lifetimes, in a random sample of n, are observed where k is a specified integer between 1 and n. This type of censoring is also known as order censoring or failure censoring.

1.2.2 Left Censoring

Left censoring occurs in life test applications when a unit has failed at the time of its first inspection; we know only that the unit failed before the inspection time. In other situations, left censored observations arise when the exact value of a response has not been observed and we have, instead, an upper bound on that response. Consider, for example, a measuring instrument that lacks the sensitivity needed to measure the observations below a known threshold. When the measurement is taken, if the signal is below the instrument threshold, we know only that the measurement is less than the threshold.

The data set may contain both left and right censored observations and in that case lifetimes are known as doubly censored. A psychiatrist collected data to determine the age at which children have learned to perform a particular task. The lifetime was the time the child has taken to learn to perform the task from date of birth. Those children who already knew how to perform the task, when he started the study were left censored and those who didn't learn the task even by the time the study ends were right censored observations.

1.2.3 Interval Censoring

Interval censoring is still another type of censoring which occurs when the lifetime is only known to occur within an interval. Such pattern occurs when patients in a clinical trial have periodic follow up and the patient's event time is only known to fall in an interval. A comprehensive review on different type of censoring is available in Lawless (2003).

Another form of incomplete data that arises in survival and reliability studies is truncation.

1.3 Truncation

Truncation occurs when only those individuals whose lifetime lies within a particular range are observed. An individual whose lifetime is not in that certain

range is not observed and no information on this subject is available to the investigator. Truncation is different from censoring, because in censoring there is at least partial information on each subject.

Left truncation occurs when we observe only those individuals whose lifetime exceeds the truncation time. A common example of left truncation is the problem of estimating the distribution of the diameters of microscopic particles. The only particles big enough to be seen based on the resolution of the microscope are observed and smaller particles do not come to the attention of investigator. In survival studies, the truncation event may be exposure to some disease, entry into retirement home, occurrence of some intermediate event etc. As an example, consider the study on ages at death of residents of retirement home. The ages at which individuals enters to the retirement home is the truncation time. Since an individual must survive to a sufficient age to enter the retirement home, all individuals who died earlier have no chance to be in the study and are considered as left truncated. In this type of truncation, any subjects who experience the event prior to the truncation time are not observed. The truncation time is often called the delayed entry time since we only observe subjects from this time until they die or are censored.

Right truncation occurs when we observe the lifetime only when lifetime is less than or equal to the truncation time. Right truncation arises, for example, in estimating the distribution of stars from the earth, in that stars too far away are not visible and are right truncated. Another example is mortality study based on death records. Right truncated data is particularly relevant to the studies of AIDS. As an example, Kalbfleisch and Prentice (2002) gives data on transfusion related AIDS cases in the United States. The study contains those individuals who were diagnosed with AIDS prior to 1988 and for whom the mode of infection was determined to be by blood transfusion. The distribution of the time from infection to diagnosis of AIDS (incubation period) is of interest. In this study, the lifetimes included in the study are subject to very strong selection favouring the shorter incubation times and individuals whose diagnosis occurs after the end of the study

period are not included in the study (thus right truncated). For elaborate discussion on truncation, one could refer to Lawless (2003) and Klein and Moeschberger (2003).

1.4 Basic Concepts

Let T be a nonnegative random variable representing lifetimes of individuals having absolutely continuous distribution function $F(.)$ with respect to the Lebesgue measure. Let $f(.)$ denote the probability density function (p.d.f.) of T. All functions, unless stated otherwise, are defined over the interval $[0,\infty)$.

1.4.1 Survivor Function

A basic function that describes lifetime data is the survivor function, which is defined as

$$S(t) = P(T \geq t) = \int_t^\infty f(x)dx .$$

The survivor function $S(t)$, is the probability of an individual surviving beyond time t. In the context involving lifetimes of systems or manufactured items, $S(t)$ is referred to as the reliability function. $S(t)$ is a non-increasing continuous function with $S(0)=1$ and $\lim_{t\to\infty} S(t)=0$. The p.d.f. of T may be represented as

$$f(t) = -\frac{dS(t)}{dt}.$$

1.4.2 Hazard Rate

An important function that characterizes lifetime distributions is the hazard rate $h(t)$, defined as

$$h(t) = \lim_{\Delta t\to 0} \frac{P\left(t \leq T < t+\Delta t \,|\, T \geq t\right)}{\Delta t}.$$

The hazard rate specifies the instantaneous rate of death or failure at time t, given that the individual survives up to time t. Thus $h(t)\Delta t$ is the approximate probability of death in the interval $[t, t+\Delta t)$, given survival up to time t. The hazard rate is also known as conditional failure rate in reliability, the force of mortality in demography, the intensity function in stochastic processes, the age-specific failure rate in epidemiology, the inverse of Mill's ratio in economics or simply the hazard function. When the p.d.f. of T, $f(t)$ exists, then the hazard rate is expressed as

$$h(t) = \frac{f(t)}{S(t)} \qquad (1.1)$$

$$= -\frac{d \log S(t)}{dt}.$$

The hazard rate fully specifies the distribution of T and determines the survivor function. Integrating (1.1) with respect to t and using $S(0) = 1$, we obtain

$$S(t) = \exp\left\{-\int_0^t h(x)\,dx\right\}. \qquad (1.2)$$

The p.d.f of T can be obtained from (1.1) and (1.2) as

$$f(t) = h(t).\exp\left\{-\int_0^t h(x)\,dx\right\}.$$

A related function is cumulative hazard rate $H(t)$, defined as

$$H(t) = \int_0^t h(x)\,dx. \qquad (1.3)$$

$S(t)$ can be represented in terms of $H(t)$ as

$$S(t) = \exp\{-H(t)\}. \qquad (1.4)$$

1.4.3 Reversed Hazard Rate

Recently, another concept, that is useful in the survival studies, is developed which is referred as reversed hazard rate. The reversed hazard rate (RHR) of T is defined as

$$m(t)=\lim_{\Delta t\to 0}\frac{P(t-\Delta t<T\le t\,|\,T\le t)}{\Delta t}. \tag{1.5}$$

The reversed hazard rate specifies the instantaneous rate of death or failure at time t, given that it failed before time t. That is, in a small interval, $m(t)\Delta t$ is the approximate probability of failure in the interval $(t-\Delta t,t]$, given failure before the end of the interval. Reversed hazard rate was proposed as a dual to the hazard rate by Barlow et al. (1963). When the p.d.f. of T, $f(t)$ exists, (1.5) is expressed as

$$m(t)=\frac{f(t)}{F(t)}=\frac{d\log F(t)}{dt}, \tag{1.6}$$

where $F(t)$ is the distribution function of T. The RHR, $m(t)$ determines the distribution function uniquely by the relationship

$$F(t)=\exp\left\{-\int_t^{\infty}m(u)\,du\right\},$$

which was given by Keilson and Sumita (1982). The p.d.f of T can be obtained from $m(t)$ using the relationship

$$f(t)=m(t)\exp\left\{-\int_t^{\infty}m(u)\,du\right\}.$$

Ware and DeMets (1976) used RHR for the estimation of the distribution function in the presence of left censored observations. Shaked and Shanthikumar (1994) presented several results based on reversed hazard rate ordering and characterization of lifetime distributions based on RHR. Block et al. (1998) pointed out that there is no non-negative random variable having an increasing

RHR distribution and observed that increasing hazard rate distributions like Weibull, gamma and lognormal distributions are decreasing reversed hazard rate distributions. Block et al. (1998) characterized properties for k out of n systems in terms of RHR. Finkelstein (2002) expressed the relation between hazard rate $h(t)$ and RHR $m(t)$ as

$$m(t)=\frac{h(t)S(t)}{F(t)}=\frac{h(t)}{\exp\left\{\int_0^t h(u)\,du\right\}-1}.$$

Lawless (2003) developed nonparametric estimators of $S(t)$ for the right truncated observations using reversed hazard rates. Chandra and Roy (2005) defined classes of distributions based on RHR and studied their implicative relationship. RHR is useful in forensic science and in actuarial science, as the time elapsed since failure is a quantity of interest in order to predict the actual time of failure. For more properties and applications of RHR function, one could refer to Kalbfleisch and Lawless (1989), Gupta and Nanda (2001), Chandra and Roy (2001), Gupta and Wu (2001), Nair and Asha (2004), Nair et al. (2005), Bartoszewicz and Skolimowska (2006) and Sankaran and Gleeja (2007a).

The cumulative reversed hazard rate $M(t)$ is defined as

$$M(t)=\int_t^\infty m(x)\,dx.$$

$F(t)$ can be represented in terms of $M(t)$ as $F(t)=\exp\{-M(t)\}$.

1.4.4 Mean Waiting Time

RHR is closely related to another important concept known as the mean waiting time. The mean waiting time (MWT) of an item failed in an interval $[0,t]$ is defined as

$$\mu(t)=E(t-T\mid T\le t)=\frac{\int_0^t F(u)\,du}{F(t)}.$$

MWT is also known as expected inactivity time (EIT) or mean past lifetime (MPL). Assuming $\mu(t)$ as differentiable, MWT is related to $m(t)$ through the relationship

$$m(t)=\frac{1-\mu'(t)}{\mu(t)},$$

where $\mu'(t)=\frac{d\mu(t)}{dt}$. The distribution function can be uniquely determined from relation

$$F(t)=\exp\left\{-\int_t^{\infty}\frac{1-\mu'(u)}{\mu(u)}du\right\}.$$

Chandra and Roy (2001) studied various properties of mean waiting time with respect to RHR. Finkelstein (2002) focused the importance of MWT in defining RHR and studied its properties. Li and Lu (2003) established some stochastic comparisons on MWT and residual life of series and parallel systems and presented some applications based on these comparisons. Reliability properties of MWT are given in Nanda et al. (2003). Asadi (2006) studied properties of MWT for components of parallel system. One could refer to Kayid and Ahmad (2004) and Nanda et al. (2006) for further properties of MWT.

1.5 Estimation

One of the basic objectives in lifetime data analysis is to estimate the distribution function $F(t)$ or the survivor function $S(t)$. The two common approaches used in such contexts are parametric and non-parametric approaches. In parametric method, we assume that the random variable T follows some p.d.f. $f(t;\theta)$ where the functional form of $f(t;\theta)$ is known but the parameter θ is unknown. Continuous distributions such as exponential, Weibull, lognormal, loglogistic, Pareto and inverse Gaussian are commonly used for modelling lifetime data. For estimation of parameters, we can employ different procedures such as maximum likelihood, method of moments, Bayesian techniques etc. For

more details on parametric lifetime models and their estimation, one may refer to Martz and Waller (1982), Sinha (1986) and Lawless (2003).

In many practical situations, the functional form of $f(t)$ is seldom known. In such situations, the estimation of $F(t)$ or $S(t)$ is done using nonparametric methods. The non-parametric methods for the estimation of distribution function or survivor function have become popular within survival studies for several reasons. In many occasions, the sample size is not enough to determine the parametric model associated with given data. Further, lifetime data often have some features that can not be easily explained by parametric models. Censoring and truncation cause problems for the determination of appropriate parametric model for the given lifetime data. If there are no censored observations in a sample of size n, $S(t)$ can be estimated by the empirical survivor function, defined as

$$\hat{S}_{ESF}(t) = \frac{\text{Number of observations} \geq t}{n}, \; t \geq 0.$$

When there are right censored observations, some modifications are necessary. Accordingly, Kaplan and Meier (1958) defined the product-limit estimator for the survivor function $S(t)$.

1.5.1 Kaplan-Meier Estimator (Product-Limit Estimator)

Suppose that n individuals have lifetimes represented by random variables $T_1, T_2, ..., T_n$ which are subject to right censoring. Let $C_1, C_2, ..., C_n$ be the corresponding censoring times. Then the observed data consist of $(\tilde{t}_i, \delta_i)$, where $\tilde{t}_i = \min(T_i, C_i)$ and $\delta_i = I(T_i = \tilde{t}_i)$, $i = 1, 2, .., n$ with $I(.)$ as the usual indicator function. Suppose that there are k $(k \leq n)$ distinct times $t_1 < t_2 < ... < t_k$ at which

death occur and $d_j = \sum_{i=1}^{n} I\left(\tilde{t}_i = t_j, \delta_i = 1\right)$ represent the number of deaths at t_j. Then the product limit estimator of $S(t)$ is defined as

$$\hat{S}(t) = \prod_{j: t_j \le t} \frac{n_j - d_j}{n_j}, \tag{1.7}$$

where $n_j = \sum_{i=1}^{n} I\left(\tilde{t}_i \ge t_j\right)$ is the number of individuals at risk at t_j, which is the number of individuals alive and uncensored just prior to t_j. The product-limit estimator does not change at censoring time points. The product-limit estimator can be derived as a nonparametric maximum likelihood estimator. When there are no censored observations, it reduces to empirical survivor function.

Another approach is to develop nonparametric estimator of $S(t)$ using the estimator of cumulative hazard rate. Accordingly, a nonparametric estimator of $H(t)$ was proposed by Nelson (1969) and then independently by Aalen in his doctoral thesis in 1972.

1.5.2 Nelson-Aalen Estimator

The estimator of the cumulative hazard rate corresponding to (1.3) is given by Riemann-Stieltjes integral as

$$\hat{H}(t) = \int_0^t d\hat{H}(u).$$

Thus the estimator of $H(t)$ is given by

$$\hat{H}(t) = \sum_{j: t_j \le t} \frac{d_j}{n_j}. \tag{1.8}$$

This is called the empirical cumulative hazard rate but is more commonly known as the Nelson-Aalen estimator.

Thus from (1.4), $S(t)$ can be estimated by

$$\hat{S}(t) = \exp\left\{-\hat{H}(t)\right\}.$$

Both the Kaplan-Meier and Nelson-Aalen estimators possess desirable large sample properties like strong consistency and asymptotic normality. It is important to note that both $\hat{H}(t)$ and $\hat{S}(t)$ are nonparametric maximum likelihood estimators. For more properties of (1.7) and (1.8), one may refer to Lawless (2003).

When there are left censored observations, Ware and DeMets (1976) developed the estimator of distribution function by reversing time.

1.5.3 Estimation of Distribution Function for Left Censored Data

Suppose that events occur at times $t_1 < t_2 < ... < t_k$ and let r_i be the number of events occurring at t_i and n_i be the number of observations (censored or not) less than or equal to t_i, and define

$$p_i = \frac{n_i - r_i}{n_i},\ i = 1, 2, ..., k.$$

Then estimate of distribution function is given as

$$\hat{F}(t) = \begin{cases} \prod_{i:\, t_i > t} p_i, & t \le t_k \\ 1, & t > t_k. \end{cases}$$

For more properties of $\hat{F}(t)$, one could refer to Ware and DeMets (1976).

1.6 Regression Models

An important way to represent heterogeneity in a population is the use of explanatory variables or covariates or risk factors. The explanatory variables can be basic variables like gender and age; they can be factors of particular interest, like treatment in a drug trial; or they can be nuisance variables, which are helpful to include in order to describe the risk of events as precisely as possible, but whose effect we are not directly interested in.

Consider lifetime variable $T > 0$, and suppose that a vector of explanatory variables $\underline{x}' = (x_1, x_2, ..., x_p)$ is available on each individual, their measurements having been taken at the beginning of the study. Thus, $\underline{x}$ may include quantitative variables (such as blood pressure, temperature, age and weight) and qualitative variables (such as gender, race, treatment and disease status). Sometimes, it is of great interest to ascertain the relationship between the lifetime variable T and one or more explanatory variables. This would be the case if one were comparing survivor functions for two or more treatments, to determine the prognosis of a patient with various characteristics, or identifying pertinent risk factors for a particular disease, controlling for relevant confounders. The main aim in such contexts is to understand and exploit the relationship between lifetime and covariates. One way to achieve this is through regression models, in which the dependence of lifetime in covariates is explicitly recognized. The most commonly used regression model is Cox (1972)'s proportional hazards model.

1.6.1 Proportional Hazards Model

Cox (1972) defined proportional hazards model as

$$h(t) = \phi h_0(t) \tag{1.9}$$

where $h_0(t)$ is an arbitrary baseline hazard rate and ϕ is some real constant of proportionality and is a measure of relative risk. If $\underline{x}$ is a column vector of covariates and $\underline{\beta}$ is a column vector of parameters, then $\phi = e^{\underline{\beta}'\underline{x}}$, and the model can be represented as

$$h(t \mid \underline{x}) = e^{\underline{\beta}'\underline{x}} h_0(t). \tag{1.10}$$

The model (1.10) is called proportional hazards model, since the ratio of hazard rates

$$\frac{h(t \mid \underline{x})}{h(t \mid \underline{x}^*)} = e^{\underline{\beta}'\underline{x} - \underline{\beta}'\underline{x}^*}, \tag{1.11}$$

for two individuals with covariate vector $\underline{x}$ and $\underline{x}^*$, is independent of time. The identity (1.11) is called the relative risk (hazard ratio) of an individual with risk factor $\underline{x}$ having the event as compared to an individual with risk factor $\underline{x}^*$.

From (1.9), the survivor functions can be related as

$$S(t)=\left[S_0(t)\right]^{\phi},$$

where $S_0(t)$ is the baseline survivor function. The class of models provided by this process is sometimes referred to as the Lehmann class (Lehmann, 1953). For a comprehensive review on this topic, one can refer Kalbfleisch and Prentice (2002) and Lawless (2003).

1.6.2 Proportional Reversed Hazards Model

Gupta et al. (1998) proposed a dual model called proportional reversed hazards model, which is expressed as

$$m(t)=\theta m_0(t) \tag{1.12}$$

where $\theta>0$ and $m_0(t)$ is the baseline reversed hazard rate. Then the relation between distribution functions can be expressed as

$$F(t)=\left[F_0(t)\right]^{\theta}$$

where $F_0(t)$ is the baseline distribution function.

The proportional reversed hazards model has strong similarity with the proportional hazards model, but is applicable in situations where proportional hazards model becomes inappropriate. For example, the model (1.12) is useful in the analysis of left censored or right truncated data. Gupta et al. (1998) and Gupta and Gupta (2007) studied the monotonicity of hazard rate and reversed hazard rate of the model (1.12). The properties based on stochastic comparisons and results related to ageing notions of random lifetimes are given in Crescenzo (2000). Chen et al. (2004) employed the proportional RHR models to study the longitudinal

pattern of recurrent gap times. Further, Chen et al. (2004) introduced the concept of frailties in proportional RHR models. The applications and methods of inference of the model (1.12) are investigated in Sengupta et al. (1998) and Gupta and Gupta (2007).

1.7 Multivariate Lifetime Data

Multivariate lifetime data arise when each study subject experiences several events or when we study repeated occurrence of the same event or when there exists some grouping of subjects which induces some dependence among lifetimes of the same group. The occurrence of blindness in the left and right eye of a person, the sequence of tumour occurrence and the onset of genetic disease among family members are some examples of such situations in biomedical research.

Hougaard (2000) provides an extensive treatment on the analysis of multivariate lifetime data. Hazard rates and other related functions discussed in the case of univariate lifetime data can be defined in various ways for multivariate lifetime data. For simplicity, we confine our discussions to bivariate lifetime data.

1.8 Bivariate Lifetime Data

In many practical situations, one may have paired lifetime data. For example, times to death or times to initial contraction of disease may be of interest for littermate pairs of rats or for twin studies in humans. The time to deterioration level or time to reaction of a treatment may be of interest in pairs of lungs, kidneys, eyes or ears of humans. In reliability, one may be interested in the distribution of the life lengths of a particular pair of components in a system.

Let $T=(T_1,T_2)$ be a non-negative random vector admitting an absolute continuous distribution function $F(t_1,t_2)$ with respect to the Lebesgue measure. Then the survivor function of T, denoted by $S(t_1,t_2)$, is given by

$$S(t_1,t_2)=P(T_1\geq t_1,T_2\geq t_2),$$

which is related to $F(t_1,t_2)$ as

$$S(t_1,t_2)=1-F(t_1,\infty)-F(\infty,t_2)+F(t_1,t_2).$$

If the density function of T, $f(t_1,t_2)$, exists, then we have

$$f(t_1,t_2)=\frac{\partial^2 S(t_1,t_2)}{\partial t_1\partial t_2}=\frac{\partial^2 F(t_1,t_2)}{\partial t_1\partial t_2}.$$

1.8.1 Bivariate Hazard Rate

In the bivariate set up, we can define the hazard rate in more than one way. The first definition of bivariate hazard rate was given by Basu (1971).

Basu (1971) defined the bivariate hazard rate as a scalar quantity given by

$$h(t_1,t_2)=\frac{f(t_1,t_2)}{S(t_1,t_2)}.$$

Unlike the univariate case, $h(t_1,t_2)$, in general, does not determine the bivariate distribution uniquely.

A second approach in defining bivariate hazard rate is provided by Johnson and Kotz (1975). Johnson and Kotz defined bivariate hazard rate as a vector given by

$$r(t_1,t_2)=(r_1(t_1,t_2),r_2(t_1,t_2)), \tag{1.13}$$

where

$$r_i(t_1,t_2)=-\frac{\partial \log S(t_1,t_2)}{\partial t_i},\ i=1,2.$$

$r_1(t_1,t_2)$ is the instantaneous rate of failure of T_1 at time t_1 given that T_1 was alive at time $T_1=t_1-$ and that T_2 survived beyond time $T_2=t_2$. The meaning of $r_2(t_1,t_2)$ is similar.

When components $r_i(t_1,t_2)$, $i=1,2$ exists and are continuous in the open set containing $R_2^+ = \{(t_1,t_2) \mid t_i > 0, i=1,2\}$ by choosing a path orthogonal to the axis connecting $(0,0)$ and (t_1,t_2) in R_2^+, we have the representation from Galambos and Kotz (1978) as an extension of the one dimensional relationship (1.2). Accordingly, $S(t_1,t_2)$ can be determined from (1.13) as

$$S(t_1,t_2) = \exp\left\{-\int_0^{t_1} r_1(u,0)du - \int_0^{t_2} r_2(t_1,u)\,du\right\} \tag{1.14}$$

or

$$S(t_1,t_2) = \exp\left\{-\int_0^{t_1} r_1(u,t_2)du - \int_0^{t_2} r_2(0,u)\,du\right\}. \tag{1.15}$$

Thus, the vector $r(t_1,t_2)$ uniquely determines the distribution of T through (1.14) and (1.15).

Darbowska (1988) provided representation of bivariate survivor function in terms of cumulative hazard rate which is a vector of three components that correspond to double and single failures. The cumulative hazard rate vector is defined as

$$\Lambda^*(t_1,t_2) = \left(\Lambda_{10}(t_1,t_2), \Lambda_{01}(t_1,t_2), \Lambda_{11}(t_1,t_2)\right)$$

where $\Lambda_{10}(dt_1,t_2) = \dfrac{-S(dt_1,dt_2)}{S(t_1-,t_2)}$, $\Lambda_{01}(t_1,dt_2) = \dfrac{-S(dt_1,dt_2)}{S(t_1,t_2-)}$,

$\Lambda_{11}(dt_1,dt_2) = \dfrac{-S(dt_1,dt_2)}{S(t_1-,t_2-)}$ and $\Lambda_{10}(0,t_2) = \Lambda_{01}(t_1,0) = \Lambda_{11}(0,0) = 0$.

The bivariate survivor function is uniquely represented using $\Lambda(t_1,t_2)$ as

$$S(t_1,t_2) = \prod_{u\le t_1}\left(1-\Lambda_{10}(du,0)\right)\prod_{v\le t_2}\left(1-\Lambda_{01}(0,dv)\right)\prod_{\substack{u\le t_1\\ v\le t_2}}\left(1-L(du,dv)\right)$$

where $L(du,dv)=\dfrac{\Lambda_{10}(du,v-)\Lambda_{01}(u-,dv)-\Lambda_{11}(du,dv)}{(1-\Lambda_{10}(du,v-))(1-\Lambda_{01}(u-,dv))}$.

Darbowska (1988) also provided an extension of the above representation to the right censored set up. Cox (1972), Marshall (1975), Shanbag and Kotz (1987) Basu and Sun (1997) and Finkelstein (2003) have also discussed different versions of hazard rate in the bivariate (multivariate) set up. The various definitions of reversed hazard rate in bivariate set up will be discussed in Chapter 2.

1.9 Frailty Models

Consider lifetimes of several individuals, which can not be assumed to be independent, because they are related in some way. A prototype bivariate example is twins where monozygotic (identical) twins have all their genetic material in common and the dizygotic (fraternal) twins have a part of their genetic material in common, like other siblings. However, both types of twins share the same childhood environment, including particular pre-birth period, and therefore are more like each other than other siblings. Another example is ordinary siblings where siblings share some genes and a part of childhood environment. Other examples are married couples who share a common unmeasured environment, the group life insurance of employees sharing a common environment at their workplace and the times to occurrence of different nonlethal diseases within the same individual. In each of these situations, there is some association within the groups of lifetimes.

A model that is increasingly popular for modelling association between individual lifetimes within subgroups is the frailty model. The notion of frailty was introduced by Vaupel et al. (1979). It provides a convenient way to introduce random effects, dependence and unobserved heterogeneity into the models for lifetime data. For example, in clinical studies, the effect of a drug or a treatment may differ substantially between subgroups of patients. In such contexts, individuals posses different frailties and that those patients who are most frail will

die earlier than others. The most common model for the frailty is a common random effect that acts multiplicatively on the hazard rates of all subgroup members. The model assumes that the hazard rate for lifetime T given an unobservable random variable W is

$$h(t)=Wh_0(t), \tag{1.16}$$

where $h_0(t)$ is the baseline hazard rate. The group variation is described by the random variable W and is common to individuals in a group and constant over time. The survivor function corresponding to (1.16) can be represented as

$$S(t)=\left[S_0(t)\right]^W, \tag{1.17}$$

where $S_0(t)$ is the baseline survivor function. From (1.17), it is clear that frailty models are extensions of proportional hazards model of Cox (1972).

The most widely used frailty model is shared frailty model in which one assumes that the hazard rate for lifetime T_j given W is given as $W\,h_j(t)$ for $j=1,2$. Then bivariate survivor function for (T_1,T_2) given W is

$$S\left(t_1,t_2\,|W\right)=\exp\left\{-W\left[H_1(t_1)+H_2(t_2)\right]\right\},$$

where $H_j(t_j)=\int_0^{t_j}h_j(u)\,du$, is the cumulative hazard rate T_j given W for $j=1,2$. The bivariate survival function is then given by integrating out W as

$$S(t_1,t_2)=E\left(\exp\left\{-W\left[H_1(t_1)+H_2(t_2)\right]\right\}\right)=L\left(H_1(t_1)+H_2(t_2)\right), \tag{1.18}$$

where $E(.)$ denotes the mathematical expectation and $L(s)=\int e^{-sw}dG(w)$ is the Laplace transform of the distribution of W, with $G(w)$ as the distribution function of W. Klein (1992) pointed out that if the realization of W is less than one, then all the members of the group tend to experience the event of interest at an earlier time, where as opposite occurs if W is greater than one. Therefore positive association between group members is induced by frailty.

One important problem in the area of frailty models is the choice of the frailty distribution. The most common choice of distribution for W is gamma distribution with mean one and unknown variance $\theta > 0$. Then, for the shared frailty model, the survival function (1.18) is obtained as

$$S(t_1, t_2) = \left[S_1(t_1)^{-\theta} + S_2(t_2)^{-\theta} - 1 \right]^{-\frac{1}{\theta}}.$$

When $\theta = 0$, lifetimes T_1 and T_2 are independent and high value of θ indicates a high correlation between lifetimes. Frailty model with gamma distribution has been studied in Vaupel et al. (1979), Oakes (1982), Clayton and Cuzick (1985), Klein (1992) and Andersen et al. (1993). Balakrishnan and Peng (2006) discussed frailty model with generalised gamma distribution. Other choices for distribution of frailty like positive stable, Weibull, lognormal etc. are discussed in Hougaard (2000).

For the analysis of shared frailty models, one could refer to Clayton (1978), Clayton and Cuzick (1985) and Hougaard (2000). Maximum likelihood estimation of the parameters of the model via the EM algorithm was studied by Nielsen et al. (1992). Murphy (1994, 1995) proved the consistency and asymptotic normality of the maximum likelihood estimators.

The shared frailty model explains association within groups (family, litter or clinic) or for recurrent events facing the same individual. However, this approach has some limitations. First, it forces unobserved factors to be same within group, which is not acceptable. For example, sometimes it may be inappropriate to assume that both partners in a twin share all of their unobserved risks. Second, the dependence between lifetimes within group is based on marginal distributions. However, when covariates are present in proportional hazards model with gamma distributed frailty, the dependence parameter and the population heterogeneity are confounded (Clayton and Cuzick, 1985). This implies that the joint distribution can be identified from the marginal distributions (Hougaard, 1986). Third, in most cases, shared frailty will only induce a positive

association within the group. However, there are some situations in which lifetimes for subjects within the same group are negatively associated. For example, this applies to growth rates for animals in the same litter that have limited food supply.

To avoid these limitations, correlated frailty model is developed for the analysis of bivariate lifetime data, in which two associated random variables are used to characterize the frailty effect for each pair. Assuming gamma distribution as the distribution of frailty random variables and by splitting, the frailty for an individual in a group, into two components, Yashin and Iachine (1995) introduced correlated gamma frailty model. For example, for the twin data, the frailty of the first twin in a pair consist of $Y_0 + Y_1$ and the frailty of the second twin consist of $Y_0 + Y_2$, where Y_0, Y_1 and Y_2 are independent gamma distributed random variables. The frailties of the two twins are then correlated by the shared part of the frailty Y_0, which describes the common genes and environment. Y_1 and Y_2 describes the possible heterogeneity between the individuals after having accounted for common genes and environment. The identifiability of the model is described in Yashin and Iachine (1999 b) and the analysis of the model is discussed in Parner (1998). Yashin and Iachine (1997), Yashin and Iachine (1999 a), Iachine (2001), Wienke (2003) and Wienke et al. (2005) also investigated various properties of the correlated frailty model and a comparison of shared frailty and correlated frailty model is discussed in Zdravkovic et al. (2004).

1.10 Association Measures

Time dependent association measures play a vital role in the analysis of bivariate survival data. For example, in a genetic study researchers wish to know the degree to which genetic factors influence life span. To measure the degree of association, one can examine monozygotic (MZ) and dizygotic (DZ) twin data. The difference in strength of association between MZ and DZ twin measures the genetic impact. Many researchers believe that important genetic influences may exist only in old age. Thus association measures indexed by age or time provide a

means of detecting such a difference. Accordingly, several such measures indexed by time have been proposed in literature to assist modelling and analysis of bivariate survival data.

Let $T=(T_1,T_2)$ be a nonnegative random vector representing dependent life spans, having an absolutely continuous survival function $S(t_1,t_2)=P(T_1>t_1,T_2>t_2)$ with respect to the Lebesgue measure. For example, T could be considered as cohort study ages at diagnosis of breast cancer of mother and daughter. Clayton (1978) introduced a measure of association as

$$\theta(t_1,t_2)=\frac{S(t_1,t_2)\left[\partial^2 S(t_1,t_2)/\partial t_1\partial t_2\right]}{\left[\partial S(t_1,t_2)/\partial t_1\right]\left[\partial S(t_1,t_2)/\partial t_2\right]}. \tag{1.19}$$

The measure (1.19) can be interpreted as the ratio of the hazard rate of conditional distribution of T_1 given $T_2=t_2$ to that of T_1 given $T_2>t_2$. Associated with (1.19), Oakes (1989) defined a measure of association, $g(t_1,t_2)=\dfrac{\theta(t_1,t_2)-1}{\theta(t_1,t_2)+1}$, which is a conditional version of Kendall's (1938) concordance coefficient. It is to be noted that $\theta(t_1,t_2)=1\,(g(t_1,t_2)=0)$ if and only if the variables T_1 and T_2 are independent. Later, Holland and Wang (1987) proposed a measure of association,

$$p(t_1,t_2)=\frac{\partial^2 \log f(t_1,t_2)}{\partial t_1\partial t_2} \tag{1.20}$$

where $f(t_1,t_2)$ is the joint density function of T, to quantify the dependence between T_1 and T_2 at different time points. Note that measure (1.20) is a natural continuous analogue of the local cross product ratios for bivariate discrete data. When $p(t_1,t_2)=0$, T_1 and T_2 are independent and conversely. Anderson et al. (1992) defined two association measures, one based on the joint and marginal survival functions of T_1 and T_2, which is given by

$$\psi(t_1,t_2)=\frac{S(t_1,t_2)}{S_1(t_1)S_2(t_2)} \tag{1.21}$$

where $S_i(t_i)$ is the marginal survival function of T_i, $i=1,2$. Obviously, $\psi(t_1,t_2)=1$ if and only if T_1 and T_2 are independent. Instead of using survival function in (1.21), we can use the mean residual life in studying association as proposed by Anderson et al. (1992) through

$$\phi(t_1,t_2)=\frac{r_1^*(t_1,t_2)}{r_1^*(t_1,0)} \tag{1.22}$$

or

$$\phi(t_1,t_2)=\frac{r_2^*(t_1,t_2)}{r_2^*(0,t_2)} \tag{1.23}$$

where $r_i^*(t_1,t_2)=E\left(T_i-t_i \mid T_1>t_1,T_2>t_2\right)$, $i=1,2$ is the i^{th} component of vector mean residual function given in Arnold and Zahedi (1988). Recently, Fan et al. (2000) have proposed a general class of measures which include as special cases the two non-parametric measures given in Fan et al. (1998). Based on the deviations from the conditional expected values, Bairamov et al. (2003) has suggested a new measure of linear dependence which is closely related to the correlation curve of Bjerve and Doksum (1992). Gupta (2003) provides a detailed analysis of association measures (1.19), (1.21), (1.22) and (1.23) by establishing their interrelationships, characterizations and some interesting identities useful for the reliability modelling. Recently, Nair and Sankaran (2008) introduced a new measure of association for bivariate survival data using product moment of residual life given in Nair et al. (2004) and the mean residual life functions. The measure (1.20) is defined only when the joint density function $f(t_1,t_2)$ exists. The measures described above except (1.20) are useful for right censored data as they are based on residual lifetimes of T_1 and T_2.

Though there has been much research on analyzing bivariate (multivariate) lifetime data under right censoring, very little has done for the analysis of left censored or right truncated lifetime data. As pointed out in Lawless (2003), reversed hazard rate, instead of hazard rate, would be more appropriate for the

analysis of such data. This brings in the relevance and need for development of new stochastic models based on reversed hazard rates, which is the focal theme of the present study. The results obtained here are of interest in their own right in survival analysis and also in various applied studies where concepts in survival analysis are used with differing interpretations.

1.11 Present Study

In studies involving lifetime, there are many situations where data is left censored. For example, in the univariate set up, Baboons in the Amboseli Reserve, Kenya, sleep in the trees and descend for foraging at some time of the day. Observers often arrive later in the day than this descent and for such days they can only ascertain that the descent took place before a particular time, so that the descent times are left censored (Andersen et al., 1993). In the multivariate set up, consider the lifetime data on patients of Fibrodysplasia Ossificans Progressiva (FOP), discussed in Jones and Rocke (2002). In 1992, 44 patient members of the International FOP Association responded to a postal survey of the age at onset of heterotopic ossification at each of 15 anatomic sites. For each patient in the survey, and for each anatomic site, the patient was asked to record the date at onset of heterotopic ossification. The data is multivariate, with 15 observations on each subject corresponding to the status of each of 15 anatomical sites. Left censoring occurred when the patient replied that a joint was already involved but they could not provide the date of onset. From the 660 onset times in the survey, 41 were left censored. As mentioned earlier, reversed hazard rate is more appropriate for the analysis of left censored or right truncated data. The analysis of bivariate right truncated data, using reversed hazard rate, is discussed in Gurler (1996).

Reversed hazard rate plays a vital role in the analysis of parallel systems, in reliability and survival analysis. For example, in certain systems or situations, sometimes the failure is prevented through numerous safety measures. Then actual failure can only occur when all lines of defence have been breached. This describes a parallel system and the parallel structure with n components works if

at least one of the components works. Examples of parallel system include automobile head lights, overhead projectors with backup bulb etc. For a system with n independent parallel components, let T denotes lifetime of the system and $T_1, T_2, ..., T_n$ denote the lifetimes of its components. Then, the distribution function of the system is

$$F_T(t) = \prod_{i=1}^{n} F_i(t),$$

where $F_i(t)$ is the distribution function of the i^{th} component. Then, the system reversed hazard rate is the sum of the component reversed hazard rates, expressed as

$$m_T(t) = \sum_{i=1}^{n} m_i(t),$$

where $m_T(t)$ denotes the system reversed hazard rate and $m_i(t)$ denotes the reversed hazard rate of the i^{th} component. When the components are i.i.d., reversed hazard rate of the system is the product of the reversed hazard rate of any component and the number of components. Sengupta et al. (1998) applied the proportional reversed hazards model for the analysis of data on parallel systems and Asadi (2006) discussed mean past lifetime of parallel system.

Motivated by the growing importance of reversed hazard rate in the analysis of left censored or right truncated lifetime data, we introduce new statistical models using reversed hazard rates for the analysis multivariate lifetime data. The thesis is organized into six chapters. After this introductory chapter, where we have pointed out the relevance and scope of the study along with a review of literature, the remaining chapters are addressed to some new results. In Chapter 2, we discuss various definitions of bivariate reversed hazard rate, existing in literature, which are useful for the analysis of dependent data. A unique representation for bivariate distribution function, using bivariate reversed hazard rates, is introduced. Based on this unique representation a class of bivariate distributions is proposed. Applications of the proposed class of distributions are also discussed.

Frailty models using reversed hazard rates are extensively studied in Chapter 3. Chen et al. (2004) used the concept of frailty with RHR to account for heterogeneity of gap times, in the analysis of longitudinal pattern of recurrent gap times. We used the concept of frailty to incorporate the unobserved heterogeneity in left censored situations. After discussing the univariate reversed hazards frailty model, we extend the reversed hazards frailty models to bivariate situation. The shared gamma frailty reversed hazards model is studied and the estimation of the parameters of the model using maximum likelihood method via EM algorithm is explored. The shared gamma frailty reversed hazards model with covariates is also discussed. Finally, we illustrate the applicability of the model with Australian twin data given in Duffy et al. (1990).

Bivariate correlated reversed hazards frailty model becomes useful if the individual frailties are correlated. In Chapter 4, we developed bivariate correlated gamma frailty reversed hazards model, to consider the left censored situations where the individual frailties are correlated. The model is extended to the multivariate set up also. The estimation of the parameters of the model via EM algorithm is investigated. Finally, we illustrate the importance of the model with the DZ Australian twin data in Duffy et al. (1990).

As mentioned in Section 1.10, time dependence measures are useful for the analysis of bivariate lifetime models. Accordingly, in Chapter 5, association measures using distribution functions and reversed hazard rates are discussed. We introduce four association measures to examine dependence among variables and study their properties. The association measures are also investigated in terms of frailty. Non-parametric estimation of association measures is discussed and the asymptotic properties of the estimators are studied. The importance of these association measures is illustrated using real life examples.

Finally, Chapter 6 summarises major conclusions of the present study. Certain open problems and suggestions for future study are also presented.

Chapter Two

Bivariate Reversed Hazard Rates*

2.1 Introduction

The reversed hazard rate plays a pivotal role in modelling and analysis of left censored lifetime data. Unlike the univariate set up discussed in Chapter one, there are more than one definition for reversed hazard rate in the multivariate setup. Gurler (1996) introduced reversed hazard rate as a three component vector in the bivariate set up, which is analogous to bivariate hazard vector given in Darbowska (1988). The reversed hazard rate vector is used for the estimation of distribution function when the data is right truncated (see Gurler (1996, 1997)). Roy (2002) defined bivariate reversed hazard rate as a two component vector and studied its properties. Further, Roy (2002) introduced a class of bivariate distributions using reversed hazard rate vector. It is proved that if both the marginal distributions belong to decreasing reversed hazard rate class, then the bivariate distribution belong to a class of distributions having bivariate decreasing reversed hazard rate property. Later, Bismi (2005) introduced a scalar definition of bivariate reversed hazard rate and used it to characterize bivariate Burr distributions. However, a systematic study on various definitions and properties of bivariate reversed hazard rate has not been carried out so far. Motivated by this, we discuss various definitions of bivariate reversed hazard rates, study their properties and propose a general class of bivariate distributions that extend the model given in Roy (2002).

* Published some results in the Journal of the Japan Statistical Society (see Sankaran and Gleeja (2006)) and some other results in the journal Metrika (see Sankaran and Gleeja (2007b))

The chapter is organized as follows. In Section 2.2, we give various definitions of bivariate reversed hazard rates and discuss their properties. An exponential representation of bivariate distributions using bivariate reversed hazard rates is given in Section 2.3. In Section 2.4, we develop a new family of bivariate distributions and study its properties. Various special cases of the family are pointed out. The parametric analysis of a model which belongs to the proposed family is given in Section 2.5. In Section 2.6, a bivariate proportional reversed hazards model is derived from the proposed class of distributions. Various applications of the new class of distributions are pointed out in Section 2.7. Finally, Section 2.8 gives conclusion of the chapter.

2.2 Bivariate Reversed Hazard Rates

Let $T=(T_1,T_2)$ be a nonnegative random vector representing lifetimes of two components of a system with an absolutely continuous distribution function $F(t_1,t_2)$ in the support of $D=[0,b_1]\times[0,b_2]$ where (b_1,b_2) is such that $b_j<\infty$ and $b_j=\inf\{t \mid F_j(t)=1\}$, $j=1,2$. Suppose that the probability density function (p.d.f.) of T, $f(t_1,t_2)$ exists. As mentioned in Section 2.1, in the bivariate set up, reversed hazard rate can be defined in more than one way.

Gurler (1996) defined bivariate reversed hazard rate as a three component vector given by

$$\Lambda(t_1,t_2)=\left(\Lambda_{12}(dt_1,dt_2),\Lambda_1(dt_1,t_2),\Lambda_2(t_1,dt_2)\right)$$

where $\Lambda_{12}(dt_1,dt_2)=\dfrac{F(dt_1,dt_2)}{F(t_1,t_2)}$, $\Lambda_1(dt_1,t_2)=\dfrac{F(dt_1,t_2)}{F(t_1,t_2)}$ and

$\Lambda_2(t_1,dt_2)=\dfrac{F(t_1,dt_2)}{F(t_1,t_2)}$.

The vector $\Lambda(t_1,t_2)$ is used for the estimation of $F(t_1,t_2)$ when the lifetime data is right truncated.

Roy (2002) defined reversed hazard rate as a two component vector given by

$$k(t_1,t_2)=\left(k_1(t_1,t_2),k_2(t_1,t_2)\right)$$

where

$$k_i(t_1,t_2)=\lim_{\Delta t_i\to 0}\frac{P\left(t_i-\Delta t_i<T_i\le t_i \mid T_1\le t_1,T_2\le t_2\right)}{\Delta t_i}=\frac{\partial \log F(t_1,t_2)}{\partial t_i},\ i=1,2. \quad (2.1)$$

For $i=1$, $k_1(t_1,t_2)\Delta t_1$, is the probability of failure of the first component in the interval $(t_1-\Delta t_1,t_1]$ given that it has failed before t_1 and the second has failed before t_2. The interpretation for $k_2(t_1,t_2)$ is similar. From Roy (2002), it follows that $k_i(t_1,t_2)$ determine $F(t_1,t_2)$ uniquely by the relationships

$$F(t_1,t_2)=\exp\left\{-\int_{t_1}^{b_1}k_1(u,b_2)\,du-\int_{t_2}^{b_2}k_2(t_1,u)\,du\right\} \quad (2.2)$$

or

$$F(t_1,t_2)=\exp\left\{-\int_{t_1}^{b_1}k_1(u,t_2)\,du-\int_{t_2}^{b_2}k_2(b_1,u)\,du\right\}, \quad (2.3)$$

where $k_1(t_1,b_2)=m_1(t_1)$ and $k_2(b_1,t_2)=m_2(t_2)$ are the marginal reversed hazard rate of T_1 and T_2 respectively.

Later, Bismi (2005) defined bivariate scalar reversed hazard rate as

$$m(t_1,t_2)=\frac{f(t_1,t_2)}{F(t_1,t_2)}. \quad (2.4)$$

It can be easily seen that (2.4) is a natural extension of the univariate reversed hazard rate given in (1.6). $m(t_1,t_2)\Delta t_1\Delta t_2+o(\Delta t_1,\Delta t_2)$ can be interpreted as the probability of failure of components 1 and 2 in intervals $(t_1-\Delta t_1,t_1]$ and $(t_2-\Delta t_2,t_2]$ respectively, given that they failed before (t_1,t_2).

It can be seen that $\Lambda_{12}(dt_1,dt_2)=m(t_1,t_2)dt_1dt_2$, $\Lambda_1(dt_1,t_2)=k_1(t_1,t_2)dt_1$ and $\Lambda_2(t_1,dt_2)=k_2(t_1,t_2)dt_2$.

A third definition of reversed hazard rate that play vital role in the analysis of dependent data, is defined as

$$k^*(t_1,t_2)=\left(k_1^*(t_1,t_2),k_2^*(t_1,t_2)\right),$$

where

$$k_i^*(t_1,t_2)=\lim_{\Delta t_i\to 0}\frac{P\left(t_i-\Delta t_i<T_i\le t_i\,|\,T_i\le t_i,T_j=t_j\right)}{\Delta t_i}=\frac{f\left(t_i\,|\,T_j=t_j\right)}{F\left(t_i\,|\,T_j=t_j\right)}, \qquad (2.5)$$

with $f\left(t_i|T_j=t_j\right)$ as the conditional density function of T_i given $T_j=t_j$ and $F\left(t_i|T_j=t_j\right)$ as the conditional distribution function of T_i given $T_j=t_j$, i,j=1,2, $i\neq j$. Thus the definition (2.5) is nothing but univariate RHR of conditional variable T_i given $T_j=t_j$. Since conditional distributions, in general, does not uniquely determine the joint density, (2.5) does not provide $F(t_1,t_2)$ uniquely.

2.3 Exponential Representation

From (2.2) and (2.3), it follows that $F(t_1,t_2)$ can be represented by $k_i(t_1,t_2)$, $i=1,2$ in two different ways. In the following, we give a unique representation for $F(t_1,t_2)$ in terms of bivariate reversed hazard rate given in (2.1) and (2.4).

Theorem 2.1

The distribution function of $T=(T_1,T_2)$ can be represented in terms of reversed hazard rates as

$$F(t_1,t_2)=\exp\left\{-\int_{t_1}^{b_1} m_1(u)\,du\right\}\exp\left\{-\int_{t_2}^{b_2} m_2(v)\,dv\right\}$$

$$\exp\left\{\int_{t_1}^{b_1}\int_{t_2}^{b_2}\left(m(u,v)-k_1(u,v)k_2(u,v)\right)dvdu\right\}. \quad (2.6)$$

Proof

The bivariate distribution function $F(t_1,t_2)$ of $T=(T_1,T_2)$ can be written as

$$F(t_1,t_2)=F_1(t_1)F_2(t_2)\exp\{A(t_1,t_2)\}, \quad (2.7)$$

where $A(t_1,t_2)=\log\dfrac{F(t_1,t_2)}{F_1(t_1)F_2(t_2)}$ and $F_i(t_i)=\exp\left\{-\int_{t_i}^{b_i} m_i(u)\,du\right\}$, $i=1,2$.

The function $A(t_1,t_2)$ can be viewed as a measure of dependence between T_1 and T_2 and we can write $A(t_1,t_2)=\int_{t_1}^{b_1}\int_{t_2}^{b_2}\phi(u,v)\,dvdu$, where $\phi(u,v)$ is some bivariate function.

Consider the representation

$$F(t_1,t_2)=\exp\{-H(t_1,t_2)\}, \quad (2.8)$$

where

$$H(t_1,t_2)=\int_{t_1}^{b_1} m_1(u)\,du+\int_{t_2}^{b_2} m_2(v)\,dv-\int_{t_1}^{b_1}\int_{t_2}^{b_2}\phi(u,v)\,dvdu. \quad (2.9)$$

The representation (2.8) can be viewed as a generalization of the univariate exponential representation to the bivariate case.

Now consider

$$k_i(t_1,t_2)=-\frac{\partial H(t_1,t_2)}{\partial t_i}=m_i(t_i)+\frac{\partial A(t_1,t_2)}{\partial t_i},\ i=1,2. \quad (2.10)$$

Differentiating both sides of (2.10) we get

$$\frac{\partial^2 A(t_1,t_2)}{\partial t_1\partial t_2}=\frac{f(t_1,t_2)}{F(t_1,t_2)}-\frac{\partial\log F(t_1,t_2)}{\partial t_1}\frac{\partial\log F(t_1,t_2)}{\partial t_2},$$

which gives

$$\phi(u,v) = m(u,v) - k_1(u,v)k_2(u,v). \tag{2.11}$$

Thus from (2.8), (2.9) and (2.11), we obtain (2.6).

Remark 2.1 It may be noted that (2.6) can be written as

$$F(t_1,t_2) = \exp\left\{-\int_{t_1}^{b_1} k_1(u,b_2)\,du\right\}\exp\left\{-\int_{t_2}^{b_2} k_2(b_1,v)\,dv\right\}$$
$$\exp\left\{\int_{t_1}^{b_1}\int_{t_2}^{b_2}\left(m(u,v)-k_1(u,v)k_2(u,v)\right)dvdu\right\}.$$

Remark 2.2 If $m(u,v)-k_1(u,v)k_2(u,v) = -\gamma m_1(u)m_2(v)$, $0 \le \gamma \le 1$, then (2.6) reduces to the model given in Roy (2002),

$$F(t_1,t_2) = F_1(t_1)F_2(t_2)\exp\{-\gamma \log F_1(t_1).\log F_2(t_2)\}.$$

Remark 2.3 If $m(u,v)-k_1(u,v)k_2(u,v) = \dfrac{\theta f_1(u) f_2(v)}{\left[1+\theta(1-F_1(u))(1-F_2(v))\right]^2}$, where $-1 \le \theta \le 1$, then (2.6) reduces to Morgenstern (1956)'s family given by

$$F(t_1,t_2) = F_1(t_1)F_2(t_2)\left[1+\theta(1-F_1(t_1))(1-F_2(t_2))\right].$$

Example 2.1 When $F_1(t_1) = 1-e^{-t_1}$, $F_2(t_2) = 1-e^{-t_2}$, where $t_1, t_2 > 0$, and

$m(u,v)-k_1(u,v)k_2(u,v) = \dfrac{\theta e^{-(u+v)}}{\left[1+\theta e^{-(u+v)}\right]^2}$ with $-1 \le \theta \le 1$, then (2.6) reduces to model II of bivariate exponential distribution,

$$F(t_1,t_2) = \left(1-e^{-t_1}\right)\left(1-e^{-t_2}\right)\left(1+\theta e^{-t_1-t_2}\right),\ t_1,t_2 > 0,$$

by Gumbel (1960) as a special application of Morgenstern (1956).

Example 2.2 When $F_1(t_1) = \left(1+e^{-t_1}\right)^{-1}$, $F_2(t_2) = \left(1+e^{-t_2}\right)^{-1}$, $t_1, t_2 > 0$ and

$m(u,v)-k_1(u,v)k_2(u,v)=e^{-(u+v)}\left(1+e^{-u}+e^{-v}\right)^{-2}$, then (2.6) reduces to standard bivariate logistic distribution by Gumbel (1961),

$$F(t_1,t_2)=\left(1+e^{-t_1}+e^{-t_2}\right)^{-1},\ t_1,t_2>0.$$

Example 2.3 When $F_1(t_1)=1-e^{-t_1}$, $F_2(t_2)=1-e^{-t_2}$, $t_1,t_2>0$ and

$$m(u,v)-k_1(u,v)k_2(u,v)=-\frac{\left[\theta+e^{u+\theta uv}\theta(u+uv\theta-1)+e^{v+\theta uv}\theta(v+uv\theta-1)\right.}{\left(e^{u+\theta uv}+e^{v+\theta uv}-e^{u+v+\theta uv}-1\right)^2},$$
$$\left.-e^{u+v+\theta uv}\left(1+(u+v-1)\theta+uv\theta^2\right)+e^{u+v+2\theta uv}\right]$$

where $0\le\theta\le1$, then (2.6) reduces to model I of bivariate exponential distribution by Gumbel (1960) with distribution function,

$$F(t_1,t_2)=\left(1-e^{-t_1}-e^{-t_2}+e^{-t_1-t_2-\theta t_1t_2}\right),\ t_1,t_2>0.$$

Example 2.4 When $F_1(t_1)=e^{-1/t_1}$, $F_2(t_2)=e^{-1/t_2}$, $t_1,t_2>0$ and

$m(u,v)-k_1(u,v)k_2(u,v)=\dfrac{-\theta}{(uv)^2}$, where $0\le\theta\le1$, then (2.6) reduces to bivariate inverse exponential distribution given by

$$F(t_1,t_2)=\exp\left\{-\frac{1}{t_1}-\frac{1}{t_2}-\frac{\theta}{t_1t_2}\right\},\ t_1,t_2>0.$$

Now we study properties of bivariate reversed hazard rates (2.1), (2.4) and (2.5). We can prove that $m(t_1,t_2)=m_1(t_1)m_2(t_2)$ implies the independence, where $m_i(t_i)$ is marginal reversed hazard rate of T_i, $i=1,2$.

Theorem 2.2

The variables T_1 and T_2 are independent if and only if

$$m(t_1,t_2)=m_1(t_1)m_2(t_2),\ \text{for all } t_1,t_2>0. \qquad (2.12)$$

Proof

When T_1 and T_2 are independent, we have

$$f(t_1,t_2)=f_1(t_1)f_2(t_2) \text{ and } F(t_1,t_2)=F_1(t_1)F_2(t_2),$$

where $f_i(t_i)$ and $F_i(t_i)$ respectively denote the marginal density and distribution functions of T_i, $i=1,2$. Thus from (2.4), we can easily obtain (2.12).

To prove the converse, from (2.4) and (2.7), we have

$$\frac{F(t_1,t_2)}{F_1(t_1)F_2(t_2)}=\frac{f(t_1,t_2)}{f_1(t_1)f_2(t_2)}=\exp\{A(t_1,t_2)\},$$

which gives

$$f(t_1,t_2)=\exp\{A(t_1,t_2)\}f_1(t_1)f_2(t_2). \tag{2.13}$$

Now differentiating (2.7) with respect to t_1, we get

$$\frac{\partial F(t_1,t_2)}{\partial t_1}=m_1(t_1)F(t_1,t_2)+F(t_1,t_2)\frac{\partial A(t_1,t_2)}{\partial t_1}. \tag{2.14}$$

Also we have

$$\frac{\partial F(t_1,t_2)}{\partial t_1}=\int_0^{t_2} f(t_1,v)dv. \tag{2.15}$$

Now using (2.13) in (2.15), we get

$$\frac{\partial F(t_1,t_2)}{\partial t_1}=m_1(t_1)F(t_1,t_2)-\int_0^{t_2} m_1(t_1)F(t_1,v)\frac{\partial A(t_1,v)}{\partial v}dv. \tag{2.16}$$

Equating (2.14) and (2.16), we get

$$F(t_1,t_2)\frac{\partial A(t_1,t_2)}{\partial t_1}=-m_1(t_1)\int_0^{t_2} F(t_1,v)\frac{\partial A(t_1,v)}{\partial v}dv. \tag{2.17}$$

Similarly, we get

$$F(t_1,t_2)\frac{\partial A(t_1,t_2)}{\partial t_2}=-m_2(t_2)\int_0^{t_1} F(u,t_2)\frac{\partial A(u,t_2)}{\partial u}du. \tag{2.18}$$

Substituting (2.10) in (2.17) and (2.18), we get

$$F(t_1,t_2)k_1(t_1,t_2)=m_1(t_1)\int_0^{t_2}F(t_1,v)m_2(v)dv \tag{2.19}$$

and

$$F(t_1,t_2)k_2(t_1,t_2)=m_2(t_2)\int_0^{t_1}F(u,t_2)m_1(u)du. \tag{2.20}$$

Dividing (2.19) by (2.20), we get

$$m_1(t_1)k_2(t_1,t_2)\int_0^{t_2}F(t_1,v)m_2(v)dv=m_2(t_2)k_1(t_1,t_2)\int_0^{t_1}F(u,t_2)m_1(u)du. \tag{2.21}$$

Now, using (2.10) in (2.21), we get

$$m_1(t_1)k_2(t_1,t_2)F(t_1,t_2)-m_1(t_1)k_2(t_1,t_2)\int_0^{t_2}F(t_1,v)\frac{\partial A(t_1,v)}{\partial v}dv$$

$$=m_2(t_2)k_1(t_1,t_2)F(t_1,t_2)-m_2(t_2)k_1(t_1,t_2)\int_0^{t_1}F(u,t_2)\frac{\partial A(u,t_2)}{\partial u}du$$

or

$$F(t_1,t_2)\left[m_1(t_1)k_2(t_1,t_2)-m_2(t_2)k_1(t_1,t_2)\right]$$
$$=m_1(t_1)k_2(t_1,t_2)\int_0^{t_2}F(t_1,v)\frac{\partial A(t_1,v)}{\partial v}dv-m_2(t_2)k_1(t_1,t_2)\int_0^{t_1}F(u,t_2)\frac{\partial A(u,t_2)}{\partial u}du. \tag{2.22}$$

The equation (2.22) is satisfied for all $(t_1,t_2)\in D$ if either

$$F(t_1,t_2)=\int_0^{t_2}F(t_1,v)\frac{\partial A(t_1,v)}{\partial v}dv=\int_0^{t_1}F(u,t_2)\frac{\partial A(u,t_2)}{\partial u}du \tag{2.23}$$

or

$$\frac{\partial A(t_1,v)}{\partial v}=\frac{\partial A(u,t_2)}{\partial u}=0. \tag{2.24}$$

If possible, suppose that (2.23) is true. Then, differentiating (2.23) with respect to t_2, we get $k_2(t_1,t_2)=\frac{\partial A(t_1,t_2)}{\partial t_2}=k_2(t_1,t_2)-m_2(t_2)$, which gives $m_2(t_2)=0$ and

similarly, if we differentiate (2.23) with respect to t_1, we get $m_1(t_1)=0$. But $m_1(t_1)=m_2(t_2)=0$ is not valid. So the identity (2.23) is not true.

Now, if (2.24) is true, then $k_2(t_1,t_2)=m_2(t_2)$ and $k_1(t_1,t_2)=m_1(t_1)$. Thus from (2.12), we obtain

$$m(t_1,t_2)=k_1(t_1,t_2)k_2(t_1,t_2). \tag{2.25}$$

Substituting (2.25) in (2.6), we get $F(t_1,t_2)=F_1(t_1)F_2(t_2)$, which completes the proof.

Theorem 2.3

T_1 and T_2 are independent if and only if

$$k_i^*(t_1,t_2)=m_i(t_i),\ i=1,2. \tag{2.26}$$

Proof

When T_1 and T_2 are independent,

$$f\left(t_i \mid T_j=t_j\right)=f_i(t_i) \text{ and } F\left(t_i \mid T_j=t_j\right)=F_i(t_i),\ i,j=1,2,\ i\neq j,$$

where $f_i(t_i)$ and $F_i(t_i)$ respectively denotes the marginal density and distribution functions of T_i, $i=1,2$. Thus, we obtain (2.26) from (2.5).

To prove the converse, note that (2.26) can be written as

$$\frac{f(t_1,t_2)}{\dfrac{\partial F(t_1,t_2)}{\partial t_j}}=\frac{\partial \log F_i(t_i)}{\partial t_i},\ i,j=1,2,\ i\neq j,$$

which gives

$$\frac{\partial}{\partial t_i}\left(\log\frac{\partial F(t_1,t_2)}{\partial t_j}\right)=\frac{\partial \log F_i(t_i)}{\partial t_i},\ i,j=1,2,\ i\neq j. \tag{2.27}$$

Integrating (2.27) with respect to t_i, we get

$$\frac{\partial F(t_1,t_2)}{\partial t_j} = F_i(t_i).B(t_j),\ i,j=1,2,\ i\neq j, \tag{2.28}$$

where $B(t_j)$ is a function independent of t_i. Since $F_i(b_i)=1$ and $F(b_i,t_j)=F_j(t_j)$, when $t_i \to b_i$ in (2.28), we get

$$B(t_j)=\frac{\partial F_j(t_j)}{\partial t_j}.$$

Now, integrating (2.28) with respect to t_j, we obtain

$$F(t_1,t_2)=F_i(t_i).F_j(t_j)+C(t_i),\ i,j=1,2,\ i\neq j, \tag{2.29}$$

where $C(t_i)$ is a function independent of t_j. When $t_j \to b_j$ in (2.29), we get $C(t_i)=0$. It follows that T_1 and T_2 are independent.

Theorem 2.4

T_1 and T_2 are independent if and only if

$$k_i(t_1,t_2)=m_i(t_i),\ i=1,2. \tag{2.30}$$

Proof

When T_1 and T_2 are independent, $F(t_1,t_2)=F_1(t_1)F_2(t_2)$. Then, from (2.1) we get (2.30). The converse part can be proved by substituting (2.30) in (2.2) and (2.3).

Theorem 2.5

T_1 and T_2 are independent if and only if

$$k_i^*(t_1,t_2)=k_i(t_1,t_2),\ i=1,2. \tag{2.31}$$

Proof

When T_1 and T_2 are independent, equating (2.26) and (2.30), we get (2.31). To prove the converse, note that (2.5) can be written as

$$k_i^*(t_1,t_2)=\frac{f(t_1,t_2)}{\frac{\partial F(t_1,t_2)}{\partial t_j}},\ i,j=1,2,\ i\neq j. \tag{2.32}$$

Then using (2.1) and (2.32) in (2.31), we get

$$f(t_1,t_2)=\frac{\partial \log F(t_1,t_2)}{\partial t_i}\frac{\partial F(t_1,t_2)}{\partial t_j},\ i,j=1,2,\ i\neq j. \tag{2.33}$$

Dividing (2.33) by $F(t_1,t_2)$, we obtain (2.25). Then by substituting (2.25) in (2.6), we get $F(t_1,t_2)=F_1(t_1)F_2(t_2)$, which completes the proof.

Theorem 2.6

T_1 and T_2 are independent if and only if

$$m(t_1,t_2)=k_1(t_1,t_2)k_2(t_1,t_2).$$

Proof

When T_1 and T_2 are independent, from (2.12) and (2.30), we obtain (2.25). The proof of the converse part follows from (2.6) and (2.25).

2.4 A New Family of Bivariate Distributions

On the basis of (2.6), we construct a new class of bivariate distributions.

Theorem 2.7

Let $F(t_1,t_2)$ be a bivariate distribution function defined by exponential representation (2.6). Assume that

(I) $\alpha_i>0,\beta_i\geq 0, i=1,2$ (II) $\beta_2\geq\beta_1$

(III) $\alpha_i-\beta_2\geq 0, i=1,2$ and (IV) $\beta_1 m(u,v)\geq\beta_2 k_1(u,v)k_2(u,v),\ u,v\geq 0.$

Then

$$F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,t_2)=(F_1(t_1))^{\alpha_1}(F_2(t_2))^{\alpha_2}\exp\left\{\int_{t_1}^{\phi_1}\int_{t_2}^{\phi_2}(\beta_1 m(u,v)-\beta_2 k_1(u,v)k_2(u,v))\,dvdu\right\} \tag{2.34}$$

defines a class of bivariate distribution function for some lifetimes (T_1,T_2) with marginals $(F_1(t_1))^{\alpha_1}$ and $(F_2(t_2))^{\alpha_2}$ respectively.

Proof

Condition (IV) is just a stronger version of $\phi(u,v)\geq 0$ on a parental distribution $F(t_1,t_2)$. When $\beta_1=\beta_2$, condition (IV) reduces to the condition $\phi(u,v)\geq 0$. Obviously, due to conditions (I) to (IV) of the theorem, the corresponding boundary conditions trivially hold. Thus, $F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(0,t_2)=F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,0)=F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(0,0)=0$.

To check the non-negativity of the joint probability density function, we differentiate (2.34) twice. Differentiating (2.34) with respect to t_2, we get

$$\frac{\partial F(t_1,t_2)}{\partial t_2}=\left[F_1(t_1)\right]^{\alpha_1}\left[F_2(t_2)\right]^{\alpha_2}\exp\left\{\int_{t_1}^{b_1}\int_{t_2}^{b_2}\left(\beta_1 m(u,v)-\beta_2 k_1(u,v)k_2(u,v)\right)dvdu\right\}$$
$$\left[\alpha_2 m_2(t_2)-\int_{t_1}^{b_1}\left(\beta_1 m(u,t_2)-\beta_2 k_1(u,t_2)k_2(u,t_2)\right)du\right]. \tag{2.35}$$

Now, differentiating (2.35) with respect to t_1, we obtain the joint p.d.f. as

$$\begin{aligned}
&f_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,t_2)\\
&=\left[F_1(t_1)\right]^{\alpha_1}\left[F_2(t_2)\right]^{\alpha_2}\exp\left\{\int_{t_1}^{b_1}\int_{t_2}^{b_2}\left(\beta_1 m(u,v)-\beta_2 k_1(u,v)k_2(u,v)\right)dvdu\right\}\\
&\left\{\alpha_1\alpha_2 m_1(t_1)m_2(t_2)+\alpha_2 m_2(t_2)\left[-\int_{t_2}^{b_2}\left(\beta_1 m(t_1,v)-\beta_2 k_1(t_1,v)k_2(t_1,v)\right)dv\right]\right.\\
&\quad+\alpha_1 m_1(t_1)\left[-\int_{t_1}^{b_1}\left(\beta_1 m(u,t_2)-\beta_2 k_1(u,t_2)k_2(u,t_2)\right)du\right]\\
&\quad+\left[-\int_{t_2}^{b_2}\left(\beta_1 m(t_1,v)-\beta_2 k_1(t_1,v)k_2(t_1,v)\right)dv\right]\\
&\qquad\left[-\int_{t_1}^{b_1}\left(\beta_1 m(u,t_2)-\beta_2 k_1(u,t_2)k_2(u,t_2)\right)du\right]\\
&\quad\left.+\left[\beta_1 m(t_1,t_2)-\beta_2 k_1(t_1,t_2)k_2(t_1,t_2)\right]\right\}
\end{aligned}$$

$$= F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,t_2)\left\{\left[\alpha_1 m_1(t_1) - \int_{t_2}^{b_2}\left(\beta_1 m(t_1,v) - \beta_2 k_1(t_1,v)k_2(t_1,v)\right)dv\right]\right.$$
$$\left.\left[\alpha_2 m_2(t_2) - \int_{t_1}^{b_1}\left(\beta_1 m(u,t_2) - \beta_2 k_1(u,t_2)k_2(u,t_2)\right)du\right]\right\}$$
$$+ F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,t_2)\left[\beta_1 m(t_1,t_2) - \beta_2 k_1(t_1,t_2)k_2(t_1,t_2)\right]. \tag{2.36}$$

Now using assumptions (I) to (IV), we get

$$\left[\alpha_1 m_1(t_1) - \int_{t_2}^{b_2}\left(\beta_1 m(t_1,v) - \beta_2 k_1(t_1,v)k_2(t_1,v)\right)dv\right]$$
$$\geq \left[\alpha_1 m_1(t_1) - \beta_2 \int_{t_2}^{b_2}\left(m(t_1,v) - k_1(t_1,v)k_2(t_1,v)\right)dv\right]$$
$$= \left[\alpha_1 m_1(t_1) - \beta_2 \int_{t_2}^{b_2}\phi(t_1,v)\,dv\right]$$
$$= \left[\alpha_1 m_1(t_1) - \beta_2 \frac{\partial A(t_1,t_2)}{\partial t_1}\right]$$
$$= \left[\alpha_1 m_1(t_1) + \beta_2\left(k_1(t_1,t_2) - m_1(t_1)\right)\right]$$
$$= \left[(\alpha_1 - \beta_2)m_1(t_1) + \beta_2 k_1(t_1,t_2)\right] \geq 0. \tag{2.37}$$

Similarly,

$$\left[\alpha_2 m_2(t_2) - \int_{t_1}^{b_1}\left(\beta_1 m(u,t_2) - \beta_2 k_1(u,t_2)k_2(u,t_2)\right)du\right]$$
$$\geq \left[(\alpha_2 - \beta_2)m_2(t_2) + \beta_2 k_2(t_1,t_2)\right] \geq 0. \tag{2.38}$$

Now substituting (2.37), (2.38) and the assumption (IV) in equation (2.36), we get $f_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,t_2) \geq 0$, which completes the proof.

Remark 2.4 For $i = 1,2$, $\alpha_i = \alpha$ and $\beta_i = \beta$, the model (2.34) reduces to

$$F_{\alpha,\alpha,\beta,\beta}(t_1,t_2) = \left(F_1(t_1)\right)^{\alpha}\left(F_2(t_2)\right)^{\alpha}\exp\left\{\beta.A(t_1,t_2)\right\}.$$

Remark 2.5 If $\alpha_i = \beta_i = \alpha$, $i = 1,2$, conditions (I) to (IV) reduce to $\alpha > 0$ and $\phi(u,v) > 0$.

Remark 2.6 Let the dependence structure of the parental distribution function $F(t_1,t_2)$ be such that $k_i^*(t_1,t_2)=(1+\theta)k_i(t_1,t_2)$, for $\theta>0$. Then, the bivariate distribution function is

$$F(t_1,t_2)=\left[\left(F_1(t_1)\right)^{-\theta}+\left(F_2(t_2)\right)^{-\theta}-1\right]^{-(1/\theta)}. \tag{2.39}$$

Now, $\beta_1 m(u,v)-\beta_2 k_1(u,v)k_2(u,v)=m(u,v)\left(\beta_1-\frac{\beta_2}{(1+\theta)}\right)$

$$=\frac{(1+\theta)}{\theta}\left(\beta_1-\frac{\beta_2}{(1+\theta)}\right)\phi(u,v)=\beta\phi(u,v), \tag{2.40}$$

where $\beta=\frac{(1+\theta)}{\theta}\left(\beta_1-\frac{\beta_2}{(1+\theta)}\right)$. Using condition (IV), $\left(\beta_1(1+\theta)-\beta_2\right)\geq 0$.

Now (2.39) can be generalized using (2.40) as

$$F_{\alpha_1,\alpha_2,\beta_1,\beta_2}(t_1,t_2)=\left(F_1(t_1)\right)^{\alpha_1-\beta}\left(F_2(t_2)\right)^{\alpha_2-\beta}\left[\left(F_1(t_1)\right)^{-\theta}+\left(F_2(t_2)\right)^{-\theta}-1\right]^{-(\beta/\theta)}. \tag{2.41}$$

The bivariate model (2.39) is analogous to the Clayton copula model (Clayton and Cuzick, 1985) based on cross ratio function.

Remark 2.7 When $\alpha_i=\beta_i=1$, $i=1,2$, from Examples 2.1 and 2.2, it is obvious that model II of bivariate exponential distribution by Gumbel (1960) and standard bivariate logistic distribution of Gumbel (1961) belong to the class of distributions given in (2.34).

Example 2.5 When $\beta_i=\theta$, $\alpha_i=1$, $i=1,2$, $F_1(t_1)=\exp\left\{-e^{-t_1}\right\}$,

$F_2(t_2)=\exp\left\{-e^{-t_2}\right\}$, $t_1,t_2>0$ and $m(u,v)-k_1(u,v)k_2(u,v)=\frac{2e^{u+v}}{\left(e^u+e^v\right)^3}$, then

(2.34) reduces to the bivariate extreme value distribution

$$F(t_1,t_2)=\exp\left\{-e^{-t_1}-e^{-t_2}+\theta\left(e^{t_1}+e^{t_2}\right)^{-1}\right\},\ 0\leq\theta\leq 1,\ t_1,t_2>0.$$

Remark 2.8 The assumption (IV) is not a necessary condition of the model (2.34). If $\alpha_i = \beta_i = k$, $i = 1,2$ and $m(u,v) - k_1(u,v)k_2(u,v) = -\gamma m_1(u)m_2(v)$ for $0 \le \gamma \le 1$, then (2.34) reduces to the bivariate model of Roy (2002) given by

$$F_{k,k,k,k}(t_1,t_2) = (F_1(t_1))^k (F_2(t_2))^k \exp\left\{-\left(\frac{\gamma}{k}\right)\log(F_1(t_1))^k .\log(F_2(t_2))^k\right\}.$$

The bivariate inverse exponential distribution with distribution function

$$F(t_1,t_2) = \exp\left(\frac{-\alpha_1}{t_1} - \frac{\alpha_2}{t_2} - \frac{\beta}{t_1 t_2}\right)$$

is also a member of the family (2.34) with $\beta_1 = \beta_2 = \beta$.

The model (2.39) can be derived analogous to Clayton and Cuzick (1985).

Theorem 2.8

Let $H(t_1,t_2) = -\log F(t_1,t_2)$, $H_i = \dfrac{\partial H(t_1,t_2)}{\partial t_i}$, $i = 1,2$ and $H_{12} = \dfrac{\partial^2 H(t_1,t_2)}{\partial t_1 \partial t_2}$.

Assume that $H(t_1,t_2)$ is twice continuously differentiable with $H(b_1,b_2) = 0$, $H(t_1,b_2) = M_1(t_1)$ and $H(b_1,t_2) = M_2(t_2)$, where $M_1(t_1)$ and $M_2(t_2)$ are marginal cumulative reversed hazard rates of T_1 and T_2 respectively. Then the unique solution of

$$H_{12} = -\theta H_1 H_2 \tag{2.42}$$

is given by

$$F(t_1,t_2) = \begin{cases} \left[(F_1(t_1))^{-\theta} + (F_2(t_2))^{-\theta} - 1\right]^{-\left(\frac{1}{\theta}\right)}, & \theta > 0 \\ F_1(t_1).F_2(t_2), & \theta = 0. \end{cases}$$

Proof

From (2.1), we note that

$$H_i = \frac{\partial H(t_1,t_2)}{\partial t_i} = -k_i(t_1,t_2),\ i = 1,2.$$

For $\theta = 0$, T_1 and T_2 are independent and $F(t_1, t_2) = F_1(t_1).F_2(t_2)$.

For $\theta > 0$, from (2.42) we can write,

$$H_1 = -\theta^{-1} \frac{H_{12}}{H_2} = -\theta^{-1} \frac{\partial \log H_2}{\partial t_1}. \tag{2.43}$$

Integrating (2.43) with respect to t_1, we get

$$H(t_1, t_2) = -\theta^{-1} \log H_2 + C(t_2)$$

for some function $C(.)$. As $t_1 \to b_1$, we obtain

$$M_2(t_2) = -\theta^{-1} \log M_2'(t_2) + C(t_2),$$

where $M_2'(t_2) = \frac{\partial M_2(t_2)}{\partial t_2}$.

Therefore,

$$H(t_1, t_2) = -\theta^{-1} \log H_2 + M_2(t_2) + \theta^{-1} \log M_2'(t_2),$$

which gives

$$H_2 = M_2'(t_2) e^{-\theta H(t_1, t_2)} . e^{\theta M_2(t_2)}. \tag{2.44}$$

Substituting $V = e^{-\theta H(t_1, t_2)}$ in (2.44), we get

$$\frac{\partial V}{\partial t_2} = -V^2 \theta M_2'(t_2) e^{\theta M_2(t_2)},$$

which provides,

$$e^{\theta H(t_1, t_2)} = e^{\theta M_2(t_2)} - B(t_1), \tag{2.45}$$

for some function $B(.)$. For $t_2 \to b_2$, we get

$$B(t_1) = 1 - e^{\theta M_1(t_1)}.$$

Therefore, we can write (2.45) as

$$F(t_1, t_2) = \left[e^{\theta M_1(t_1)} + e^{\theta M_2(t_2)} - 1 \right]^{-\left(\frac{1}{\theta}\right)}, \ \theta > 0,$$

or

$$F(t_1, t_2) = \left[\left(F_1(t_1)\right)^{-\theta} + \left(F_2(t_2)\right)^{-\theta} - 1 \right]^{-\left(\frac{1}{\theta}\right)}, \ \theta > 0.$$

2.5 Parametric Analysis

Let us consider the parametric analysis of the model (2.39) where marginal distribution functions $F_1(t_1)$ and $F_2(t_2)$ are specified up to Lehmann alternatives. Suppose that $F_1(t_1)=\left[F_{01}(t_1)\right]^{\gamma_1}$ and $F_2(t_2)=\left[F_{02}(t_2)\right]^{\gamma_2}$, where $F_{01}(t_1)$ and $F_{02}(t_2)$ are known distribution functions and $\gamma_1>0$ and $\gamma_2>0$ are known parameters. Then we have the joint p.d.f. of (T_1,T_2) as

$$f(t_1,t_2)=\frac{\gamma_1\gamma_2(1+\theta)\left[F_{01}(t_1)\right]^{-\gamma_1\theta-1}\left[F_{02}(t_2)\right]^{-\gamma_2\theta-1}f_{01}(t_1)f_{02}(t_2)}{\left[\left[F_{01}(t_1)\right]^{-\gamma_1\theta}+\left[F_{02}(t_2)\right]^{-\gamma_2\theta}-1\right]^{\left(\frac{1}{\theta}+2\right)}},\ \theta>0,$$

where $f_{01}(t_1)$ and $f_{02}(t_2)$ are probability density functions corresponding to the distribution functions $F_{01}(t_1)$ and $F_{02}(t_2)$.

Let us consider the transformation $V_1=-\log F_{01}(T_1)$ and $V_2=-\log F_{02}(T_2)$. Then the joint density of (V_1,V_2) becomes

$$f(v_1,v_2)=\frac{\gamma_1\gamma_2(1+\theta)\exp\{v_1\gamma_1\theta+v_2\gamma_2\theta\}}{\left[\exp\{v_1\gamma_1\theta\}+\exp\{v_2\gamma_2\theta\}-1\right]^{\left(\frac{1}{\theta}\right)+2}}.$$

Consider n independent and identically distributed (i.i.d.) samples (v_{1i},v_{2i}), $i=1,2,...,n$ from (V_1,V_2). Then the likelihood function is

$$L(\theta,\gamma_1,\gamma_2)=\frac{\left[\gamma_1\gamma_2(1+\theta)\right]^n\exp\left\{\sum_{i=1}^{n}(v_{1i}\gamma_1+v_{2i}\gamma_2)\theta\right\}}{\prod_{i=1}^{n}\left[\exp\{v_{1i}\gamma_1\theta\}+\exp\{v_{2i}\gamma_2\theta\}-1\right]^{\left(\frac{1}{\theta}\right)+2}}.$$

To obtain the maximum likelihood estimates of θ, γ_1 and γ_2, we take the first order derivatives of the log likelihood function with respect to θ, γ_1 and γ_2, which are obtained as

$$\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\theta}=\frac{n}{1+\theta}+\sum_{i=1}^{n}(v_{1i}\gamma_1+v_{2i}\gamma_2)+\sum_{i=1}^{n}\left(\frac{1}{\theta^2}\right)\log\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]$$
$$-\sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\frac{\left[e^{v_{1i}\gamma_1\theta}v_{1i}\gamma_1+e^{v_{2i}\gamma_2\theta}v_{2i}\gamma_2\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]},$$

(2.46)

$$\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_1}=\frac{n}{\gamma_1}+\sum_{i=1}^{n}v_{1i}\theta-\sum_{i=1}^{n}(1+2\theta)\frac{e^{v_{1i}\gamma_1\theta}v_{1i}}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]},\tag{2.47}$$

and

$$\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_2}=\frac{n}{\gamma_2}+\sum_{i=1}^{n}v_{2i}\theta-\sum_{i=1}^{n}(1+2\theta)\frac{e^{v_{2i}\gamma_2\theta}v_{2i}}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}.\tag{2.48}$$

Equating (2.46), (2.47) and (2.48) to zero and solving the equations using numerical techniques, we can obtain maximum likelihood estimates of the parameters θ, γ_1 and γ_2.

To obtain information matrix of the estimates of θ, γ_1 and γ_2, we find the second order derivatives of log likelihood function as

$$\frac{\partial^2 \log L(\theta,\gamma_1,\gamma_2)}{\partial\theta^2}=\frac{-n}{(1+\theta)^2}+\sum_{i=1}^{n}(v_{1i}\gamma_1+v_{2i}\gamma_2)+\sum_{i=1}^{n}\left(\frac{2}{\theta^2}\right)\frac{\left[e^{v_{1i}\gamma_1\theta}v_{1i}\gamma_1+e^{v_{2i}\gamma_2\theta}v_{2i}\gamma_2\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}$$
$$-\sum_{i=1}^{n}\left(\frac{2}{\theta^3}\right)\log\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]-\sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\frac{\left[e^{v_{1i}\gamma_1\theta}(v_{1i}\gamma_1)^2+e^{v_{2i}\gamma_2\theta}(v_{2i}\gamma_2)^2\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}$$
$$+\sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\left[\frac{e^{v_{1i}\gamma_1\theta}(v_{1i}\gamma_1)+e^{v_{2i}\gamma_2\theta}(v_{2i}\gamma_2)}{e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1}\right]^2,\tag{2.49}$$

$$\frac{\partial^2 \log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_1^2}=\frac{-n}{\gamma_1^2}+\sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\frac{e^{2v_{1i}\gamma_1\theta}(v_{1i}\theta)^2}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2}$$
$$-\sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\frac{e^{v_{1i}\gamma_1\theta}(v_{1i}\theta)^2}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}$$
$$=\frac{-n}{\gamma_1^2}+\sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\frac{e^{v_{1i}\gamma_1\theta}\left[1-e^{v_{2i}\gamma_2\theta}\right](v_{1i}\theta)^2}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2},\tag{2.50}$$

$$\frac{\partial^2 \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_2^{\,2}} = \frac{-n}{\gamma_2^{\,2}} + \sum_{i=1}^{n}\left(\frac{1}{\theta}+2\right)\frac{e^{v_{2i}\gamma_2\theta}\left[1-e^{v_{1i}\gamma_1\theta}\right](v_{2i}\theta)^2}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2}, \tag{2.51}$$

$$\frac{\partial^2 \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_1 \partial \theta} = \sum_{i=1}^{n} v_{1i} + \sum_{i=1}^{n}\left(\frac{1}{\theta}\right)\frac{e^{v_{1i}\gamma_1\theta}v_{1i}}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}$$

$$+\sum_{i=1}^{n}\frac{\left(\frac{1}{\theta}+2\right)e^{v_{1i}\gamma_1\theta}v_{1i}\theta\left[e^{v_{1i}\gamma_1\theta}v_{1i}\gamma_1+e^{v_{2i}\gamma_2\theta}v_{2i}\gamma_2\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2} - \sum_{i=1}^{n}\frac{\left(\frac{1}{\theta}+2\right)\left[e^{v_{1i}\gamma_1\theta}v_{1i}^{\,2}\gamma_1\theta+e^{v_{1i}\gamma_1\theta}v_{1i}\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}$$

$$=\sum_{i=1}^{n} v_{1i} + \sum_{i=1}^{n}\frac{2e^{v_{1i}\gamma_1\theta}v_{1i}}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]} + \sum_{i=1}^{n}\frac{(1+2\theta)e^{v_{1i}\gamma_1\theta}v_{1i}\left[e^{v_{2i}\gamma_2\theta}v_{2i}\gamma_2+v_{1i}\gamma_1-e^{v_{2i}\gamma_2\theta}v_{1i}\gamma_1\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2},$$

(2.52)

$$\frac{\partial^2 \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_2 \partial \theta} = \sum_{i=1}^{n} v_{2i} + \sum_{i=1}^{n}\frac{2e^{v_{2i}\gamma_2\theta}v_{2i}}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}$$

$$+\sum_{i=1}^{n}\frac{(1+2\theta)e^{v_{2i}\gamma_2\theta}v_{2i}\left[e^{v_{1i}\gamma_1\theta}v_{1i}\gamma_1+v_{2i}\gamma_2-e^{v_{1i}\gamma_1\theta}v_{2i}\gamma_2\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2}, \tag{2.53}$$

and

$$\frac{\partial^2 \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_2 \partial \gamma_1} = \sum_{i=1}^{n}\frac{(1+2\theta)\theta e^{v_{1i}\gamma_1\theta}e^{v_{2i}\gamma_2\theta}v_{1i}v_{2i}}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]^2}. \tag{2.54}$$

Then we find out $E(V_1)$, $E(V_2)$, $E(V_1^2)$ and $E(V_2^2)$.

$$E(V_1) = \int_0^\infty\int_0^\infty v_1\gamma_1\gamma_2(1+\theta)e^{v_1\gamma_1\theta+v_2\gamma_2\theta}\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]^{-\left(\frac{1}{\theta}\right)-2}dv_2dv_1$$

$$= \int_0^\infty v_1\gamma_1\gamma_2(1+\theta)e^{v_1\gamma_1\theta}\int_0^\infty e^{v_2\gamma_2\theta}\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]^{-\left(\frac{1}{\theta}\right)-2}dv_2dv_1.$$

Consider the transformation $Y=\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]$, then

$$E(V_1) = \int_0^\infty v_1\gamma_1\frac{(1+\theta)}{\theta}e^{v_1\gamma_1\theta}\left(\int_{e^{v_1\gamma_1\theta}}^\infty y^{-\left(\frac{1}{\theta}\right)-2}dy\right)dv_1 = \int_0^\infty v_1\gamma_1 e^{-v_1\gamma_1}dv_1 = \frac{1}{\gamma_1}.$$

Similarly, we get $E(V_2)=\frac{1}{\gamma_2}$, $E(V_1^2)=\frac{2}{\gamma_1^2}$ and $E(V_2^2)=\frac{2}{\gamma_2^2}$.

Now, we can obtain

$$E\left(\log\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]\right)$$

$$=\int_0^\infty\int_0^\infty \log\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]\gamma_1\gamma_2(1+\theta)e^{v_1\gamma_1\theta+v_2\gamma_2\theta}\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]^{-\left(\frac{1}{\theta}\right)-2}dv_2dv_1.$$

Considering the transformation, $U_1=e^{V_1\gamma_1\theta}$ and $U_2=e^{V_2\gamma_2\theta}$, we get

$$E\left(\log\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]\right)=\int_1^\infty\int_1^\infty \log\left[u_1+u_2-1\right]\frac{(1+\theta)}{\theta^2}\left[u_1+u_2-1\right]^{-\left(\frac{1}{\theta}\right)-2}du_2du_1$$

$$=\frac{2\theta+\theta^2}{1+\theta}. \qquad (2.55)$$

$$E\left(\frac{V_1\gamma_1e^{V_1\gamma_1\theta}}{\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]}\right)$$

$$=\int_0^\infty\int_0^\infty \frac{v_1\gamma_1e^{v_1\gamma_1\theta}}{\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]}\gamma_1\gamma_2(1+\theta)e^{v_1\gamma_1\theta+v_2\gamma_2\theta}\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]^{-\left(\frac{1}{\theta}\right)-2}dv_2dv_1.$$

Considering the transformation, $U_1=e^{V_1\gamma_1\theta}$ and $U_2=e^{V_2\gamma_2\theta}$, we get

$$E\left(\frac{V_1\gamma_1e^{V_1\gamma_1\theta}}{\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]}\right)=\int_1^\infty\int_1^\infty\frac{(1+\theta)}{\theta^3}u_1\left[u_1+u_2-1\right]^{-\left(\frac{1}{\theta}\right)-3}\log(u_1)du_2du_1$$

$$=\frac{1+\theta}{1+2\theta}. \qquad (2.56)$$

Similarly, we obtain

$$E\left(\frac{(V_1\gamma_1)^2e^{V_1\gamma_1\theta}}{\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]}\right)=\frac{2(1+\theta)}{1+2\theta},\ E\left(\frac{(V_1\gamma_1)^2e^{2V_1\gamma_1\theta}}{\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]^2}\right)=\frac{2(1+\theta)}{1+3\theta}, \qquad (2.57)$$

and

$$E\left(\frac{V_1V_2\gamma_1\gamma_2e^{V_1\gamma_1\theta+V_2\gamma_2\theta}}{\left[e^{V_1\gamma_1\theta}+e^{V_2\gamma_2\theta}-1\right]^2}\right)=\frac{1}{1+3\theta}+\frac{1}{(1+2\theta)(1+3\theta)}\int_1^\infty\int_1^\infty\frac{1}{u_1u_2}\left[u_1+u_2-1\right]^{-\frac{1}{\theta}}du_1du_2$$

$$=\rho(\theta)\ \text{(say)}. \tag{2.58}$$

Now, using (2.49) - (2.54) and (2.55)-(2.58), we get

$$\frac{-1}{n}E\left(\frac{\partial^2\log L(\theta,\gamma_1,\gamma_2)}{\partial\theta^2}\right)=\frac{1}{(1+\theta)^2}+\frac{2}{\theta(1+\theta)(1+2\theta)}-\frac{4(1+\theta)}{(1+3\theta)}-\frac{2(1+2\theta)\rho(\theta)}{\theta},$$

$$\frac{-1}{n}E\left(\frac{\partial^2\log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_i^2}\right)=\frac{1}{\gamma_i^2}\left[1+\frac{2(1+\theta)\theta^2}{(1+3\theta)}\right],\ i=1,2,$$

$$\frac{-1}{n}E\left(\frac{\partial^2\log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_i\partial\theta}\right)=\frac{1}{\gamma_i}\left[1-(1+2\theta)\rho(\theta)+2+\frac{4\theta^3}{(1+2\theta)(1+3\theta)}\right],\ i=1,2$$

and

$$\frac{-1}{n}E\left(\frac{\partial^2\log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_1\partial\gamma_2}\right)=\frac{-(1+2\theta)\theta\rho(\theta)}{\gamma_1\gamma_2},$$

which are the elements of the information matrix. The variance of the estimate of θ can be obtained from the information matrix.

Now we consider the case when $\theta\to 0$. First we consider the first order derivatives when $\theta\to 0$. Using L-Hospital's rule, we have

$$\lim_{\theta\to0}\left(\frac{1}{\theta^2}\right)\log\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]=-v_{1i}v_{2i}\gamma_1\gamma_2\ \text{and}$$

$$\lim_{\theta\to0}\left(\frac{1}{\theta}\right)\frac{\left[e^{v_{1i}\gamma_1\theta}v_{1i}\gamma_1+e^{v_{2i}\gamma_2\theta}v_{2i}\gamma_2\right]}{\left[e^{v_{1i}\gamma_1\theta}+e^{v_{2i}\gamma_2\theta}-1\right]}=-2v_{1i}v_{2i}\gamma_1\gamma_2.$$

Therefore as $\theta\to 0$, (2.46), (2.47) and (2.48) reduces to

$$\frac{\partial\log L(\theta,\gamma_1,\gamma_2)}{\partial\theta}=n-\sum_{i=1}^{n}(v_{1i}\gamma_1+v_{2i}\gamma_2)-\sum_{i=1}^{n}(v_{1i}\gamma_1v_{2i}\gamma_2)+2\sum_{i=1}^{n}(v_{1i}\gamma_1v_{2i}\gamma_2)$$
$$-2\sum_{i=1}^{n}(v_{1i}\gamma_1+v_{2i}\gamma_2),$$

$$=\sum_{i=1}^{n}(v_{1i}\gamma_1-1)(v_{2i}\gamma_2-1), \tag{2.59}$$

$$\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_1}=\frac{n}{\gamma_1}-\sum_{i=1}^{n} v_{1i} \tag{2.60}$$

and

$$\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_2}=\frac{n}{\gamma_2}-\sum_{i=1}^{n} v_{2i}\,. \tag{2.61}$$

The estimates of γ_1 and γ_2 are obtained from (2.60) and (2.61) by equating to zero.

To obtain the variance covariance matrix as $\theta \to 0$, we need to find $E(V_1V_2)$, $E(V_1^2V_2)$ and $E(V_1^2V_2^2)$.

$$E(V_1V_2)=\int_0^\infty\int_0^\infty v_1v_2\gamma_1\gamma_2(1+\theta)e^{v_1\gamma_1\theta+v_2\gamma_2\theta}\left[e^{v_1\gamma_1\theta}+e^{v_2\gamma_2\theta}-1\right]^{-\left(\frac{1}{\theta}\right)-2}dv_2dv_1\,.$$

Considering the transformation, $U_1=e^{V_1\gamma_1\theta}$ and $U_2=e^{V_2\gamma_2\theta}$, we get

$$E(V_1V_2)=\int_1^\infty\int_1^\infty \log u_1\log u_2\frac{(1+\theta)}{\gamma_1\gamma_2\theta^4}\left[u_1+u_2-1\right]^{-\left(\frac{1}{\theta}\right)-2}du_2du_1$$

$$=\frac{1}{\gamma_1\gamma_2}\left[\frac{1}{(\theta-1)^2}-\int_1^\infty\frac{\left(-u_2^{-\frac{1}{\theta}+2}\right)}{u_2^{\,2}(\theta-1)(2\theta-1)}du_2-\int_1^\infty\int_1^\infty\frac{\left[u_1+u_2-1\right]^{-\left(\frac{1}{\theta}\right)+2}}{u_1^{\,2}u_2^{\,2}(\theta-1)(2\theta-1)}du_2du_1\right].$$

As $\theta\to 0$, $E(V_1V_2)=\dfrac{1}{\gamma_1\gamma_2}$. In similar steps, as $\theta\to 0$, we obtain

$E(V_1^2V_2)=\dfrac{2}{\gamma_1^{\,2}\gamma_2}$, $E(V_1V_2^2)=\dfrac{2}{\gamma_1\gamma_2^{\,2}}$ and $E(V_1^2V_2^2)=\dfrac{4}{\gamma_1\gamma_2^{\,2}}$.

Now we can obtain the elements of variance covariance matrix. From (2.59), we get

$$E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\theta}\right)^2=E\left(\sum_{i=1}^{n}(V_{1i}\gamma_1-1)(V_{2i}\gamma_2-1)\right)^2$$

$$= E\left(\gamma_1^2\gamma_2^2\left[\sum_{i=1}^{n} V_{1i}^2 V_{2i}^2 + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{2i}V_{1j}V_{2j}\right] + \gamma_1^2\left[\sum_{i=1}^{n} V_{1i}^2 + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{1j}\right]\right.$$

$$+ \gamma_2^2\left[\sum_{i=1}^{n} V_{2i}^2 + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{2i}V_{2j}\right] + n^2 - 2\gamma_1^2\gamma_2\left[\sum_{i=1}^{n} V_{1i}^2 V_{2i} + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{2i}V_{1j}\right]$$

$$- 2\gamma_1\gamma_2^2\left[\sum_{i=1}^{n} V_{1i}V_{2i}^2 + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{2i}V_{2j}\right] + 2n\gamma_1\gamma_2\sum_{i=1}^{n} V_{1i}V_{2i}$$

$$\left. +2\gamma_1\gamma_2\left[\sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{2j} + \sum_{i=1}^{n} V_{1i}V_{2i}\right] - 2n\gamma_1\sum_{i=1}^{n} V_{1i} - 2n\gamma_2\sum_{i=1}^{n} V_{2i}\right).$$

Substituting the values of $E(V_1)$, $E(V_2)$, $E(V_1^2)$, $E(V_2^2)$, $E(V_1V_2)$, $E(V_1^2V_2)$ and $E(V_1^2V_2^2)$, we obtain $E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\theta}\right)^2 = n$. Similarly, from (2.59), (2.60) and (2.61), we get

$$E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_1}\right)^2 = E\left(\frac{n}{\gamma_1} - \sum_{i=1}^{n} V_{1i}\right)^2$$

$$= E\left(\frac{n^2}{\gamma_1^2} - \left(\sum_{i=1}^{n} V_{1i}^2 + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{1j}\right) - \frac{2n}{\gamma_1}\sum_{i=1}^{n} V_{1i}\right) = \frac{n}{\gamma_1^2},$$

$$E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_2}\right)^2 = E\left(\frac{n}{\gamma_2} - \sum_{i=1}^{n} V_{2i}\right)^2 = \frac{n}{\gamma_2^2},$$

$$E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\theta}\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial\gamma_1}\right) = E\left(\left(\sum_{i=1}^{n}(V_{1i}\gamma_1 - 1)(V_{2i}\gamma_2 - 1)\right)\left(\frac{n}{\gamma_2} - \sum_{i=1}^{n} V_{2i}\right)\right)$$

$$= E\left(n\gamma_2\sum_{i=1}^{n} V_{1i}V_{2i} - n\sum_{i=1}^{n} V_{1i} - n\frac{\gamma_2}{\gamma_1}\sum_{i=1}^{n} V_{2i} + \frac{n^2}{\gamma_1} - \gamma_1\gamma_2\left[\sum_{i=1}^{n} V_{1i}^2 V_{2i} + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{2i}V_{1j}\right]\right.$$

$$\left. +\gamma_1\left[\sum_{i=1}^{n} V_{1i}^2 + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{1j}\right] + \gamma_2\left[\sum_{i=1}^{n} V_{1i}V_{2i} + \sum_{i=1}^{n}\sum_{i\neq j}^{n}\sum_{j=1}^{n} V_{1i}V_{2j}\right] - n\sum_{i=1}^{n} V_{1i}\right)$$

$$= 0$$

and

$$E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_1}\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_2}\right)=E\left(\left(\frac{n}{\gamma_1}-\sum_{i=1}^{n}V_{1i}\right)\left(\frac{n}{\gamma_2}-\sum_{i=1}^{n}V_{2i}\right)\right)$$

$$=E\left(\frac{n^2}{\gamma_1\gamma_2}-\frac{n\sum_{i=1}^{n}V_{2i}}{\gamma_1}-\frac{n\sum_{i=1}^{n}V_{1i}}{\gamma_2}+\sum_{i=1}^{n}V_{1i}V_{2i}+\sum_{i=1}^{n}\sum_{i\neq j\ j=1}^{n}V_{1i}V_{2j}\right)=0.$$

Similarly $E\left(\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial \theta}\frac{\partial \log L(\theta,\gamma_1,\gamma_2)}{\partial \gamma_2}\right)=0$. Thus we get variance covariance matrix of the log likelihood, when $\theta\to 0$, is diagonal with elements $n,\ \frac{n}{\gamma_1^2},\ \frac{n}{\gamma_2^2}$.

2.6 Bivariate Marginal Proportional Reversed Hazards Model

Let (T_1,T_2) be a random vector with the support $0<t_1<t_2<\infty$. Let marginal reversed hazards satisfy proportional reversed hazards property for some baseline reversed hazards $m_{01}(t_1)$ and $m_{02}(t_2)$. Then, we can obtain families of distributions for the random vector (T_1,T_2) as a special case of (2.34).

1. When $\alpha_1=\theta_1$, $\alpha_2=\theta_2$ and $\beta_1=\beta_2=\theta_2$ the conditions (I) to (IV) becomes $\theta_1>0$, $\theta_2>0$, $\theta_1\geq\theta_2$ and $m(u,v)\geq k_1(u,v)k_2(u,v)$, $u,v\geq 0$ and we get

$$F(t_1,t_2,\theta_1,\theta_2)=\left[F_1(t_1)\right]^{\theta_1-\theta_2}\left[F(t_1,t_2)\right]^{\theta_2}, \qquad (2.62)$$

where the marginal distribution functions are $\left[F_1(t_1)\right]^{\theta_1}$ and $\left[F_2(t_2)\right]^{\theta_2}$.

2. For $\alpha_1=\theta_1$, $\alpha_2=\theta_2$ and $\beta_1=\beta_2=\theta_1$ the conditions (I) to (IV) becomes $\theta_1>0$, $\theta_2>0$, $\theta_1\leq\theta_2$ and $m(u,v)\geq k_1(u,v)k_2(u,v)$, $u,v\geq 0$ and we get

$$F(t_1,t_2,\theta_1,\theta_2)=\left[F_2(t_2)\right]^{\theta_2-\theta_1}\left[F(t_1,t_2)\right]^{\theta_1}, \qquad (2.63)$$

where the marginal distribution functions are $\left[F_1(t_1)\right]^{\theta_1}$ and $\left[F_2(t_2)\right]^{\theta_2}$.

3. When $\alpha_1 = \theta$, $\alpha_2 = \theta$ and $\beta_1 = \beta_2 = \theta$ the conditions (I) to (IV) becomes $\theta > 0$ and $m(u,v) \geq k_1(u,v)k_2(u,v)$, $u,v \geq 0$ and hence we obtain

$$F(t_1,t_2,\theta) = \left[F(t_1,t_2)\right]^{\theta}, \tag{2.64}$$

with the marginal distribution functions as $\left[F_1(t_1)\right]^{\theta}$ and $\left[F_2(t_2)\right]^{\theta}$.

The distribution functions (2.62), (2.63) and (2.64) satisfies

$$m_i(t_i) = \theta_i m_{0i}(t_i),\ i = 1,2, \tag{2.65}$$

where $m_i(t_i)$ represents the RHR of marginal distribution function of T_i and $m_{0i}(t_i)$ represents the baseline RHR of marginal distribution function of T_i for $i = 1,2$.

In the following, we give two examples of families of distributions $F(t_1,t_2,\theta_1,\theta_2)$.

Example 2.6. Suppose that Y and W are two independent nonnegative random variables such that Y has distribution function $Q(.)$ and density function $q(.)$ and W has distribution function $G(.)$ and density function $g(.)$. Let $T_1 = Y$ and $T_2 = Y + W$. Then $0 < t_1 < t_2 < \infty$ and $F(t_1,t_2) = \int_0^{t_1} q(u)G(t_2 - u)du$, $0 < t_1 < t_2 < \infty$. From (2.62), (2.63) and (2.64), we get

$$F(t_1,t_2,\theta_1,\theta_2) = \begin{cases} \left[Q(t_1)\right]^{\theta_1-\theta_2}\left[\int_0^{t_1} q(u)G(t_2-u)du\right]^{\theta_2}, & \theta_1 \geq \theta_2 \\ \left[\int_0^{t_2} q(u)G(t_2-u)du\right]^{\theta_2-\theta_1}\left[\int_0^{t_1} q(u)G(t_2-u)du\right]^{\theta_1}, & \theta_1 \leq \theta_2 \\ \left[\int_0^{t_1} q(u)G(t_2-u)du\right]^{\theta}, & \theta_1 = \theta_2 = \theta. \end{cases} \tag{2.66}$$

If $q(t)=g(t)=e^{-t}$, (2.66) can be written as,

$$F(t_1,t_2,\theta_1,\theta_2)=\begin{cases}\left[1-e^{-t_1}\right]^{\theta_1-\theta_2}\left[1-e^{-t_1}-t_1e^{-t_2}\right]^{\theta_2}, & \theta_1\geq\theta_2\\ \left[1-(1+t_2)e^{-t_2}\right]^{\theta_2-\theta_1}\left[1-e^{-t_1}-t_1e^{-t_2}\right]^{\theta_1}, & \theta_1\leq\theta_2\\ \left[1-e^{-t_1}-t_1e^{-t_2}\right]^{\theta}, & \theta_1=\theta_2=\theta.\end{cases}$$

When $q(t)=e^{-t}$ and $G(t)=1-\{1-H'(t)\}e^{-H(t)}$, where H is twice continuously differentiable function on $[0,\infty]$ satisfying $H(\infty)=\infty$, $H(0)=H'(0)=0$, $0\leq H'(t)\leq 1$ and $H''(t)\geq 0$ with $H'(t)=\dfrac{dH(t)}{dt}$ and $H''(t)=\dfrac{d^2H(t)}{dt^2}$. Then (2.66) reduces to

$$F(t_1,t_2,\theta_1,\theta_2)=\begin{cases}\left[1-e^{-t_1}\right]^{\theta_1-\theta_2}\left[1-e^{-t_1}\left(1+e^{-H(t_2-t_1)}\right)-e^{-H(t_2)}\right]^{\theta_2}, & \theta_1\geq\theta_2\\ \left[1-e^{-H(t_2)}\right]^{\theta_2-\theta_1}\left[1-e^{-t_1}\left(1+e^{-H(t_2-t_1)}\right)-e^{-H(t_2)}\right]^{\theta_1}, & \theta_1\leq\theta_2\\ \left[1-e^{-t_1}\left(1+e^{-H(t_2-t_1)}\right)-e^{-H(t_2)}\right]^{\theta}, & \theta_1=\theta_2=\theta.\end{cases}$$

If we make monotone transformations $L(T_1)$ and $L(T_2)$ of T_1 and T_2 with $L(0)=0$, $L(\infty)=\infty$, and $\dfrac{dL(t)}{dt}>0$, then equation (2.66) becomes

$$F(t_1,t_2,\theta_1,\theta_2)=\begin{cases}\left[1-e^{-L(t_1)}\right]^{\theta_1-\theta_2}\left[1-e^{-L(t_1)}\left(1+e^{-H(L(t_2)-L(t_1))}\right)-e^{-H(L(t_2))}\right]^{\theta_2}, & \theta_1\geq\theta_2\\ \left[1-e^{-H(L(t_2))}\right]^{\theta_2-\theta_1}\left[1-e^{-L(t_1)}\left(1+e^{-H(L(t_2)-L(t_1))}\right)-e^{-H(L(t_2))}\right]^{\theta_1}, & \theta_1\leq\theta_2\\ \left[1-e^{-L(t_1)}\left(1+e^{-H(L(t_2)-L(t_1))}\right)-e^{-H(L(t_2))}\right]^{\theta}, & \theta_1=\theta_2=\theta.\end{cases}$$

(2.67)

When $H(t)=t-\log(1+t)$ and $L(t)=t^{\gamma}$, with $\gamma>0$, (2.67) reduces to

$$F(t_1,t_2,\theta_1,\theta_2)=\begin{cases}\left[1-e^{-t_1^{\gamma}}\right]^{\theta_1-\theta_2}\left[1-e^{-t_1^{\gamma}}+t_1^{\gamma}e^{-t_2^{\gamma}}-2e^{-t_2^{\gamma}}\left(1+t_2^{\gamma}\right)\right]^{\theta_2}, & \theta_1\geq\theta_2\\ \left[1-e^{-t_2^{\gamma}}\left(1+t_2^{\gamma}\right)\right]^{\theta_2-\theta_1}\left[1-e^{-t_1^{\gamma}}+t_1^{\gamma}e^{-t_2^{\gamma}}-2e^{-t_2^{\gamma}}\left(1+t_2^{\gamma}\right)\right]^{\theta_1}, & \theta_1\leq\theta_2\\ \left[1-e^{-t_1^{\gamma}}+t_1^{\gamma}e^{-t_2^{\gamma}}-2e^{-t_2^{\gamma}}\left(1+t_2^{\gamma}\right)\right]^{\theta}, & \theta_1=\theta_2=\theta.\end{cases}$$

Finally, if J is another function that satisfies $J(0)=0$, $J(\infty)=\infty$ and $\frac{dJ(t)}{dt}\geq\frac{dL(t)}{dt}$ where for all $t\geq 0$, then we get a family of distributions as

$$F(t_1,t_2,\theta_1,\theta_2)=\begin{cases}\left[1-e^{-J(t_1)}\right]^{\theta_1-\theta_2}\left[1-e^{-J(t_1)}\left(1+e^{-H(L(t_2)-L(t_1))}\right)-e^{-H(L(t_2))}\right]^{\theta_2}, & \theta_1\geq\theta_2\\ \left[1-e^{-H(L(t_2))}\right]^{\theta_2-\theta_1}\left[1-e^{-J(t_1)}\left(1+e^{-H(L(t_2)-L(t_1))}\right)-e^{-H(L(t_2))}\right]^{\theta_1}, & \theta_1\leq\theta_2\\ \left[1-e^{-J(t_1)}\left(1+e^{-H(L(t_2)-L(t_1))}\right)-e^{-H(L(t_2))}\right]^{\theta}, & \theta_1=\theta_2=\theta.\end{cases}$$

Example 2.7 Let Z be a nonnegative random variable with continuous density function $f(z)$. Suppose that Z and (Y_1,Y_2) are independent. Let $F(y_1,y_2)=\int_0^{y_1}q(u)G(y_2-u)du$. Then $T_1=Y_1+Z$ and $T_2=Y_2+Z$ has joint distribution function,

$$F(t_1,t_2)=\int_0^{t_1}\int_0^{t_1-z}q(u)G(t_2-z-u)f(z)dudz. \qquad (2.68)$$

Suppose that $q(t)=g(t)=e^{-t}$ and $f(z)=\frac{e^{-z}z^{\alpha-1}}{\left[\Gamma(\alpha)\right]}$, $0<z<\infty$, $\alpha>0$, then we can obtain (2.68) as

$$F(t_1,t_2)=\int_0^{t_1}\frac{e^{-z}z^{\alpha-1}}{\{\Gamma(\alpha)\}}dz-\frac{e^{-t_2}t_1^{\alpha+1}}{\{\Gamma(\alpha+2)\}}-\frac{e^{-t_1}t_1^{\alpha}}{\{\Gamma(\alpha+1)\}}, \qquad (2.69)$$

with marginals

$$F_1(t_1)=\int_0^{t_1}\frac{e^{-z}z^{\alpha-1}}{\{\Gamma(\alpha)\}}dz-\frac{e^{-t_1}t_1^{\alpha}}{\{\Gamma(\alpha+1)\}}, \tag{2.70}$$

and

$$F_2(t_2)=\int_0^{t_1}\frac{e^{-z}z^{\alpha-1}}{\{\Gamma(\alpha)\}}dz-\frac{e^{-t_2}t_2^{\alpha}}{\{\Gamma(\alpha+2)\}}\left((\alpha+1)+t_2\right). \tag{2.71}$$

Then the joint density function of (T_1,T_2) is

$$f(t_1,t_2)=\frac{e^{-t_2}t_1^{\alpha}}{\Gamma(\alpha+1)},$$

and thus (T_1,T_2) has a bivariate gamma distribution proposed by McKay (1934). Substituting (2.69), (2.70) and (2.71) in (2.62), (2.63) and (2.64), we can obtain $F(t_1,t_2,\theta_1,\theta_2)$ which satisfies (2.65).

2.7 Applications

The class of models (2.34) can be used to represent the lifetime of a parallel system in reliability analysis. Suppose that there are k - identical systems, each has two components. Let $T_i=(T_{1i},T_{2i})$ be the lifetime vector of the i^{th} system, $i=1,2,...,k$. Consider a parallel combination, as collection of two parallel connections, the first one with all the first components and the second one with all the second components. Thus the bivariate lifetime vector of the parallel combination is given by $U=(U_1,U_2)$, where $U_1=\max(T_{1i})$ and $U_2=\max(T_{2i})$, $i=1,2,...,k$. When the distribution of $T=(T_1,T_2)$ is of the form (2.34), the distribution of $U=(U_1,U_2)$ is obtained as

$$F^*(t_1,t_2)=F_1(t_1)^{\alpha_1 k}F_2(t_2)^{\alpha_2 k}\exp\left\{\int_{t_1}^{\infty}\int_{t_2}^{\infty}\left(k\beta_1 m(u,v)-k\beta_2 k_1(u,v)k_2(u,v)\right)dvdu\right\}.$$

Thus, we have a closure property of the model (2.34) under a bivariate parallel combination.

The bivariate lifetime model (2.39) can be used as a lifetime model induced by frailties in the following way.

Suppose that $T=(T_1,T_2)$ represents the lifetimes of a two component system. Suppose that there exists a positive random variable W such that the conditional distribution function of T_i given $W=w$ is

$$P(T_i<t_i\,|\,W=w)=F_i(t_i)^w,\ i=1,2, \tag{2.72}$$

and that given $W=w$, T_1 and T_2 are conditionally independent. Then (2.72) can be considered as a frailty model in the univariate setup. Then a bivariate frailty model is given by $F(t_1,t_2)=\int(F_1(t_1)F_2(t_2))^w dG(w)$, where $F_1(t_1)$ and $F_2(t_2)$ are some baseline distribution function of T_1 and T_2 respectively and $G(w)$ is the distribution function of W. The model (2.72) is also equivalent to proportional reversed hazards model of Gupta et al. (1998). If $F_i(t_i)$ is a distribution function, so also is

$$F_i^*(t_i)=\exp\left[-\left\{\frac{1}{F_i(t_i)}\right\}^{\theta}+1\right],\ i=1,2.$$

A random effects interpretation of the model (2.39) can be given in terms of $F_i^*(t_i)$, $i=1,2$. Let W have a gamma density with p.d.f. $g(w)\ \propto\ e^{-w}w^{\frac{1}{\theta}-1}$ and suppose that conditionally on $W=w$, T_1 and T_2 are independent with distribution functions $F_1^*(t_1)^w$ and $F_2^*(t_2)^w$ respectively. Then it is easy to see that, unconditionally, T_1 and T_2 have joint distribution function (2.39). This gives another interpretation of the model (2.39). Further, this representation gives a convenient method for simulating T_1 and T_2. As $\theta\to 0$, $F(t_1,t_2)\to F_1(t_1)F_2(t_2)$ corresponding to independence between T_1 and T_2.

The joint p.d.f. of T_1 and T_2 is

$$f(t_1,t_2)=\frac{\left[(1+\theta)f_1(t_1)f_2(t_2)K(t_1,t_2)^{-(\frac{1}{\theta}+2)}\right]}{\left[(F_1(t_1)F_2(t_2))^{(1+\theta)}\right]}$$

where $K(t_1,t_2)=(F_1(t_1))^{-\theta}+(F_2(t_2))^{-\theta}-1$.

We can generalize (2.39) by considering the distribution function of T_i, given $W=w$ as $F_i(t_i)^{\gamma_i}F_i^*(t_i)^w$, $i=1,2$ where γ_i is an additional parameter and the distribution of W is gamma with shape parameter β/θ. In this case, the joint distribution of T_1 and T_2 will be of the form (2.41) with $\alpha_i-\beta=\gamma_i$, $i=1,2$.

Thus, when two observed lifetimes T_1 and T_2 each depend on the same unobserved frailty via a proportional reversed hazards model, then this common dependence induces an association between the observed times.

2.8 Conclusion

In this chapter, various definitions of bivariate reversed hazard rates were studied and then a unique representation of bivariate distribution function in terms of bivariate reversed hazard rates was introduced. Based on the exponential representation of bivariate distribution function, a new class of bivariate distribution functions was developed. It was also proved that class of distributions given in Roy (2002) is a special case of the proposed class of bivariate distributions. A bivariate proportional reversed hazards model was also derived from the proposed class of models and some examples were given. The application of the model in the analysis of parallel system was discussed. Frailty models could be derived from the proposed class of models.

Chapter Three

Proportional Reversed Hazards Frailty Models[†]

3.1 Introduction

Frailty models are widely employed in bivariate survival data as they allow us to model the dependence through common random effect. For example, in survival studies of two related individuals, the lifetimes or ages at onset of disease depend on common unobserved risk factors such as genotypes, susceptibility to some disease or condition and genetic transmission in a family tree etc. Frailty models for such bivariate data are derived under conditional independence assumption by specifying latent variables that acts multiplicatively on the baseline hazard rates. But in many practical situations, reversed hazard rate is more appropriate to analyze the survival data. Andersen et al. (1993), Gurler (1996) and Lawless (2003) discussed the use of reversed hazard rates for the analysis of left censored or right truncated data. Duffy et al. (1990) considered Australian Twin data which consist of information on the age at appendectomy of monozygotic (MZ) and dizygotic (DZ) twins. There were 21 pairs with missing age at onset and therefore the data contains left censored observations. However, Duffy et al. (1990) excluded these left censored observations in the analysis. It is therefore, appropriate to model common random effect by including those left censored observations, which can be done by developing frailty models using reversed hazard rates. Accordingly, in this chapter, we introduce a class of frailty models, which will be useful for the analysis of left censored data.

[†] Results are summarized in Sankaran and Gleeja (2008b) and communicated.

The chapter is organized as follows. Section 3.2 introduces the proportional reversed hazards frailty model. The univariate gamma frailty reversed hazards model and shared gamma frailty reversed hazards model are also discussed. The estimation of the parameters of shared gamma frailty reversed hazards model, by maximum likelihood method, using EM algorithm is presented in Section 3.3. The properties of the estimators are also discussed. The shared gamma frailty reversed hazards model with covariates is discussed in Section 3.4. In Section 3.5, we apply the models to a real data set. A brief conclusion of the chapter is given in Section 3.6.

3.2 Proportional Reversed Hazards Frailty Models

Proportional reversed hazards frailty models for the lifetime data are derived under the conditional independence assumption by specifying the frailty random variables that acts multiplicatively on the baseline reversed hazard rates.

3.2.1 Univariate Frailty Reversed Hazards Model

Let T be a nonnegative random variable representing lifetimes of individuals having absolutely continuous distribution function $F(.)$ with respect to the Lebesgue measure over the interval $(0,b)$ where $b=\sup\{t \mid F(t)<1\}$. Let Z be the frailty random variable. Then RHR of T given frailty Z be

$$m(Z,t)=Zm_0(t),$$

where $m_0(t)$ is the baseline RHR of T. The conditional distribution function of T given Z is

$$F(t \mid Z)=\exp\{-ZM_0(t)\},$$

where $M_0(t)=\int_t^b m_0(u)du$ is the cumulative baseline RHR of T.

Thus unconditional distribution function of T is

$$F(t) = E\Big(F\big(t \mid Z\big)\Big) = E\big(\exp\{-ZM_0(t)\}\big)$$

$$= \int_z \exp\{-zM_0(t)\}\, dG(z) = L\big(M_0(t)\big),$$

where $E(.)$ denotes the mathematical expectation and $L(s) = \int e^{-sz} dG(z)$ is the Laplace transform of the distribution of Z, with $G(z)$ as the distribution function of Z.

3.2.1.1 Univariate Gamma Frailty Reversed Hazards Model

When the frailty random variable Z follows $Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$, we get

$$F(t) = \left[1 + \theta M_0(t)\right]^{-\frac{1}{\theta}}.$$

Univariate frailty reversed hazards models without observed covariates and without any parametric assumptions on $m_0(t)$ are not identifiable from lifetime data, as it is not possible to divide the variation into that within and that between individuals, if there is only one observation per individual. Consequently, univariate frailty models are not much helpful for the analysis of lifetime data.

3.2.2 Bivariate Frailty Reversed Hazards Model

Let $T = (T_1, T_2)$ be a nonnegative random vector representing lifetimes of two related individuals with an absolutely continuous distribution function $F(t_1, t_2)$ in the support of $D = [0, b_1] \times [0, b_2]$ where (b_1, b_2) is such that $b_j < \infty$, and $b_j = \inf\{t \mid F_j(t) = 1\}$, $j = 1, 2$. Let $m_j\big(Z_j, t_j\big) = Z_j m_{0j}\big(t_j\big)$, $j = 1, 2$ be their individual RHRs given frailties Z_1 and Z_2, where $m_{0j}\big(t_j\big)$ are the baseline reversed hazard rate of T_j, $j = 1, 2$. We assume that lifetimes (T_1, T_2) are conditionally independent given frailties Z_1 and Z_2. Then the distribution function of (T_1, T_2) given frailties Z_1 and Z_2 is

$$F(t_1,t_2 \mid Z_1,Z_2) = \exp\{-Z_1 M_{01}(t_1) - Z_2 M_{02}(t_2)\}, \tag{3.1}$$

where $M_{0j}(t_j) = \int_{t_j}^{b_j} m_{0j}(u)\,du$ is the cumulative baseline RHR of T_j, $j=1,2$.

If $g(z_1,z_2)$ is the joint density function of (Z_1,Z_2), then the bivariate distribution function of (T_1,T_2) is

$$\begin{aligned} F(t_1,t_2) &= \int_{z_2}\int_{z_1} F(t_1,t_2 \mid z_1,z_2)\, g(z_1,z_2)\,dz_1 dz_2 \\ &= E\left(F(t_1,t_2 \mid Z_1,Z_2)\right) = E\left(\exp\{-Z_1 M_{01}(t_1) - Z_2 M_{02}(t_2)\}\right) \\ &= L\left(M_{01}(t_1), M_{02}(t_2)\right), \end{aligned} \tag{3.2}$$

where $L(s_1,s_2)$ is the Laplace transform of (Z_1,Z_2).

The marginal distribution function of T_i is

$$F_i(t_i) = \int_{z_i} F_i(t_i \mid z_i)\,dG_i(z_i), \quad i=1,2, \tag{3.3}$$

where $G_i(z_i)$ is the distribution function of Z_i, $i=1,2$.

Retaining the notations in Chapter 2, we can obtain marginal RHR and bivariate RHR in terms of frailties as

$$m_i(t_i) = E\left(m_i(Z_i,t_i) \mid T_i \le t_i\right),\ i=1,2,$$

$$k_i(t_1,t_2) = E\left(m_i(Z_i,t_i) \mid T_1 \le t_1, T_2 \le t_2\right),\ i=1,2, \tag{3.4}$$

$$m(t_1,t_2) = E\left(m_1(Z_1,t_1)\, m_2(Z_2,t_2) \mid T_1 \le t_1, T_2 \le t_2\right), \tag{3.5}$$

and

$$k_i^*(t_1,t_2) = \frac{m(t_1,t_2)}{k_j(t_1,t_2)} = \frac{E\left(m_1(Z_1,t_1)\, m_2(Z_2,t_2) \mid T_1 \le t_1, T_2 \le t_2\right)}{E\left(m_j(Z_j,t_j) \mid T_1 \le t_1, T_2 \le t_2\right)},\ i,j=1,2,\ i \ne j.$$

We can also see that

$$m(t_1,t_2) - k_1(t_1,t_2)\,k_2(t_1,t_2) = Cov\left(m_1(Z_1,t_1)\, m_2(Z_2,t_2) \mid T_1 \le t_1, T_2 \le t_2\right).$$

The model corresponding to $Z_1 = Z_2 = Z$ is called shared frailty reversed hazards model.

3.2.2.1 Shared Gamma Frailty Reversed Hazards Model

For the shared frailty reversed hazards model, we have $m_j(Z,t_j) = Zm_{0j}(t_j)$ as individual RHRs given frailty Z where $m_{0j}(t_j)$, $j=1,2$ are the baseline reversed hazards. We assume that lifetimes (T_1,T_2) are conditionally independent given frailty Z. Then distribution function of (T_1,T_2) given frailty Z is

$$F(t_1,t_2 \mid Z) = \exp\left\{-\int_{t_1}^{b_1} m_1(Z,u)\,du - \int_{t_2}^{b_2} m_2(Z,v)\,dv\right\}.$$

If $G(z)$ is the distribution function of Z, then the bivariate distribution function of (T_1,T_2) is

$$F(t_1,t_2) = \int_z F(t_1,t_2 \mid z)\,dG(z) \;= E\big(F(t_1,t_2 \mid Z)\big)$$

$$= E\Big(\exp\{-Z\big(M_{01}(t_1)+M_{02}(t_2)\big)\}\Big) = L\big(M_{01}(t_1)+M_{02}(t_2)\big). \quad (3.6)$$

The marginal distribution function of T_j is

$$F_j(t_j) = \int_z F_j(t_j \mid z)\,dG(z), \quad j=1,2. \quad (3.7)$$

Assume that Z is an i.i.d. random variable following gamma density

$$g(z) = \frac{z^{(1/\theta)-1}\exp\{-z/\theta\}}{\theta^{(1/\theta)}\Gamma(1/\theta)}, \qquad \theta>0, \quad (3.8)$$

with mean 1 and variance θ. Then, using (3.8) in (3.6), we get

$$F(t_1,t_2) = \left[\theta M_{01}(t_1)+\theta M_{02}(t_2)+1\right]^{-(1/\theta)}, \quad (3.9)$$

as the distribution function of (T_1,T_2), where $M_{01}(t_1)$ and $M_{02}(t_2)$ are the cumulative baseline reversed hazard rates. The marginal distribution functions (3.7) are obtained as

$$F_j(t_j)=\left[\theta M_{0j}(t_j)+1\right]^{(-1/\theta)},\ j=1,2.$$

We can also represent (3.9) as

$$F(t_1,t_2)=\left[F_1(t_1)^{-\theta}+F_2(t_2)^{-\theta}-1\right]^{-(1/\theta)}. \tag{3.10}$$

The model (3.10) is gamma frailty reversed hazards model using marginal distribution functions, which is analogous to the model given by Clayton (1978).

Remark 3.1 The condition $k_i^*(t_1,t_2)=(1+\theta)k_i(t_1,t_2)$, uniquely determines the joint distribution function (3.10).

Identifiability of the model is established by the following theorem.

Theorem 3.1

Let $F(t_1,t_2)$ be a known distribution function given by (3.10) and $\theta>0$, and let $m_i(b_i)\neq 0$, $i=1,2$. Then the model is identifiable.

Proof

From (3.10), we obtain the joint density function as

$$f(t_1,t_2)=\frac{(1+\theta)m_1(t_1)m_2(t_2)F_1(t_1)^{-\theta}F_2(t_2)^{-\theta}}{\left[F_1(t_1)^{-\theta}+F_2(t_2)^{-\theta}-1\right]^{\left(\frac{1}{\theta}\right)+2}}.$$

Since $m_i(b_i)\neq 0$ and $F_i(b_i)=1$, $i=1,2$, we have

$$\theta=\frac{f(b_1,b_2)}{m_1(b_1)m_2(b_2)}-1.$$

From the above expression it is clear that the identified value is unique.

3.3 Estimation of Shared Gamma Frailty Reversed Hazards Model

As mentioned earlier, shared gamma frailty reversed hazards model is useful for the analysis of left censored data. So we consider the estimation of shared gamma frailty reversed hazards model, when the lifetime data is left censored. Let the lifetimes $T=(T_1,T_2)$ and censoring times $U=(U_1,U_2)$ be defined on a common probability space $(\Omega,\mathcal{F},P)$. Assume that T and U are independent. Then the observable random vectors are given by $\tilde{T}=(\tilde{T}_1,\tilde{T}_2)$ and $\delta=(\delta_1,\delta_2)$ where $\tilde{T}_j=max(T_j,U_j)$ and $\delta_j=I(T_j=\tilde{T}_j)$, $j=1,2$. Suppose now that $\tilde{T}_i=(\tilde{T}_{1i},\tilde{T}_{2i})$ and $\delta_i=(\delta_{1i},\delta_{2i})$, $i=1,2,...,n$ is an i.i.d. sample, each $(\tilde{T}_i,\delta_i)$ having the same distribution as $(\tilde{T},\delta)$.

Let $Y_{ij}(t)=I\{\tilde{T}_{ij}\le t\}$ and $N_{ij}(t)=I\{\tilde{T}_{ij}\le t,\delta_{ij}=1\}$, $i=1,2,...,n$ and $j=1,2$. We define the counting process $\mathbf{Y}(t)$ as $\mathbf{Y}(t)=(Y_i(t),\ i=1,2,...,n)$, where $Y_i(t)=\sum_{j=1,2}Y_{ij}(t)$, the total number failed or censored before time t and in the time interval $(t-dt,t]$ in the i[th] sample. Consider $\mathbf{N}$ as a multivariate counting process with components N_i, where components with the same value of the first index i share the same frailty variable Z_i. The counting process $\mathbf{N}(t)=(N_i(t),\ i=1,2,...,n\ )$, counts the observed failures $N_i(t)=\sum_{j=1,2}N_{ij}(t)$ by the time t in the i[th] sample. Thus $\mathbf{N}(t)=(N_i(t),\ i=1,2,...,n)$ is a multivariate counting process with intensity process $m_i(t)$ satisfying

$$m_i(t)=Z_iY_i(t)m_0(t)$$

where $m_0(t)$ represents unknown baseline reversed hazard rate and Z_i is unobservable i.i.d. random variable with Gamma($1/\theta$, $1/\theta$) distribution. We

consider maximum likelihood estimation of the parameter θ and the cumulative baseline reversed hazard rate $M_0(t) = \int_t^b m_0(s)\,ds$, where $b = Max(b_1, b_2)$.

Since censoring is independent of lifetime, the partial conditional likelihood based on $\mathbf{N}(t)$ is given by the product integral

$$dP = \prod_t \left\{ \prod_i \left[m_i(t) \right]^{\Delta N_i(t)} \left[1 - m_\cdot(t)\,dt \right]^{1-\Delta N_\cdot(t)} \right\}, \tag{3.11}$$

where $N_\cdot(t) = \sum_i N_i(t)$ and $m_\cdot(t) = \sum_{i=1}^{n} m_i(t)$. Considered as a function of Z, (3.11) is proportional to conditional density of the data $(\mathbf{N}(t), \mathbf{Y}(t))$ given $Z = z$. Substituting the specification of $m_i(t)$ and evaluating the product integral, we get

$$L(\theta) = \prod_i \left\{ \frac{z_i^{(1/\theta)-1} \exp\{-z_i/\theta\}}{\theta^{(1/\theta)}\Gamma(1/\theta)} \prod_t \left(z_i Y_i(t)\,dM_0(t) \right)^{\Delta N_i(t)} \exp\left[-z_i \int_0^b Y_i(s)\,dM_0(s) \right] \right\}. \tag{3.12}$$

Conditional on the data, Z_i are still independent and gamma distributed with parameters $(1/\theta) + N_i(b)$ and $(1/\theta) + \int_0^b Y_i(s)\,dM_0(s)$. Integrating out Z in (3.12), we get the marginal partial likelihood as

$$L(\theta) = \prod_i \left\{ \frac{\left(\Gamma\left((1/\theta) + N_i(b) \right) \prod_t \left(Y_i(t)\,dM_0(t) \right)^{\Delta N_i(t)} \right)}{\left(\theta^{(1/\theta)}\Gamma(1/\theta) \left[(1/\theta) + \int_0^b Y_i(s)\,dM_0(s) \right]^{(1/\theta)+N_i(b)} \right)} \right\}. \tag{3.13}$$

We take (3.13) as the likelihood function in which cumulative baseline RHR $M_0(t)$ is not necessarily absolutely continuous. The maximum likelihood

estimator for $M_0(t)$ will be discrete with jumps at jump times of $\mathbf{N}$ only. So its computation comes down to maximizing (3.13) over such $M_0(t)$, replacing $dM_0(t)$ by the jump of $M_0(t)$ at that time point. The estimates of the parameters which maximises (3.13), maximises (3.12) also. We use the EM algorithm to maximize (3.13). The E step is to estimate z_i,

$$\hat{z}_i = \frac{(1/\theta)+N_i(b)}{(1/\theta)+\int_0^b Y_i(s)\,dM_0(s)}.$$

The M step is then to calculate $\hat{\theta}$, the maximum likelihood estimator for θ from (3.13), and

$$\hat{M}_0(t) = \int_t^b \frac{dN_\cdot(s)}{\sum_i \hat{z}_i Y_i(s)}.$$

By general theory of EM algorithm, if this algorithm converges, it converges to a stationary point of $\log L(\theta)$.

Testing of independence between individuals within a pair can be done based on likelihood ratio test, where the null hypothesis is $H_0 : \theta = 0$. This null hypothesis is tested using $\chi^2_{(1)}$ approximation to the likelihood ratio test statistic, -2log Q = 2 (log L($\hat{\theta}$)-log L(0)).

Now we discuss the asymptotic properties of the estimators. The consistency of the estimators is established in Theorem 3.2.

Let T_1 be the first jump of $\mathbf{N}$, θ_0 lies in a known interval [0, S] and true cumulative baseline reversed hazard M_{00} be strictly decreasing and continuous on [0, b] for b<∞.

Theorem 3.2

Assume that

i. Y is a non-decreasing step function and $P(Y(t) \geq 1)$ has at most finite number of discontinuities in $t \in (0,b)$,

ii. $\underset{u \in (0,b)}{Inf} E(Y(u)) > 0$,

iii. $P(Y(T_1) \geq 1) < 1$,

then $\underset{t \in (0,b)}{Sup} \left| \hat{M}_0(t) - M_{00}(t) \right| \to 0$ almost surely (a.s.) and $\left| \hat{\theta} - \theta_0 \right| \to 0$ a. s.

Proof

The assumption (i) is used to prove that $\hat{M}_0(t)$ does not diverge to infinity, (ii) is used to ensure that counting process $\mathbf{N}(t)$ has sufficient activity on the entire interval so as to estimate the parameters, and (iii) excludes the possibility of $\mathbf{N}(t)$ having at most only one jump. The model becomes unidentifiable if all N_i have only one jump. The rest of proof of the theorem is similar to the one given in Murphy (1994).

The asymptotic normality of the estimators can be established in the following way.

Set $M_{0t}(0) = \int_0^b 1 + th_1(u) d\hat{M}_0(u)$ and $\theta_t = th_2 + \hat{\theta}$ for h_1 a function and h_2 a scalar, and differentiate at $t = 0$ to get $F_n\left(\hat{M}_0, \hat{\theta}\right)(h_1, h_2)$. Then, if $\left(\hat{M}_0, \hat{\theta}\right)$ maximizes $\log L(\theta)$, $F_n\left(\hat{M}_0, \hat{\theta}\right)(h_1, h_2) = 0$ for all (h_1, h_2). The form of F_n is given by

$$F_n = F_{n1} + F_{n2},$$

where $$F_{n1}\left(\hat{M}_0,\hat{\theta}\right)(h_1)=n^{-1}\sum_{i=1}^{n}\int_0^b h_1 dN_i-\frac{(1/\theta)+N_i(b)}{(1/\theta)+\int_0^b Y_i dM_0}\int_0^b h_1 Y_i dM_0$$

and

$$F_{n2}(M_0,\theta)(h_2)=h_2 n^{-1}\sum_{i=1}^{n}\int_0^b \frac{N_i(u)}{1+\theta N_i(u)}dN_i(u)$$

$$+\theta^{-2}\left(\log\left(1+\theta\int_0^b Y_i dM_0\right)-\frac{(1/\theta)+N_i(b)}{(1/\theta)+\int_0^b Y_i dM_0}\theta\int_0^b h_1 Y_i dM_0\right).$$

For $\theta=0$, the last term is taken as its limit as θ approaches zero to get $\left(\left(\int_0^b Y_i dM_0\right)^2/2\right)+N_i(b)\int_0^b Y_i dM_0$. The class of h is taken to be the space of bounded variation cross the reals. Define the norm to be $\|h\|_H=\|h_1\|_v+|h_2|$, where $\|h_1\|_v$ is absolute value of $h_1(0)$ plus the total variation of h_1 on the interval $[0,b]$. Define H_p to be the product space of bounded variation functions on $[0,b]$ and real valued scalars with norm $\|h\|_H=\|h_1\|_v+|h_2|\le p$. If $p=\infty$, then the inequality is strict. In the following p is assumed to be finite unless stated otherwise. Define $(M_0,\theta)(h)=\int_0^b h_1 dM_0+h_2\theta$. Then the parameter space Ψ can be considered to be a subset of $l^\infty(H_p)$, which is the space bounded by real valued functions on H_p under the supremum norm $\|U\|=\sup_{h\in H_p}|U(h)|$. The score function F_n is a random map from Ψ to $l^\infty(H_p)$ for all finite p.

Theorem 3.3

Assume that

i. $\underset{t\in(0,b)}{Sup}\left|\hat{M}_0(t)-M_{00}(t)\right|\to 0$ a.s. and $\left|\hat{\theta}-\theta_0\right|\to 0$ a. s.,

ii. There exist a constant K for which $\|Y\| \le K$ and $N(b) \le K$ a.s.,

iii. $\inf_{u \in (0,b)} E(Y(u)) > 0$,

iv. $P(Y(T_1) \ge 1) < 1$,

then $\sqrt{n}\left|\hat{M}_0(t) - M_{00}(t)\right| . \sqrt{n}\left|\hat{\theta} - \theta_0\right| \Rightarrow \mathcal{G}$

on $l^\infty(H_p)$; $\mathcal{G}$ is a tight Gaussian process on $l^\infty(H_p)$ with mean zero and covariance process

$$Cov(\mathcal{G}(h), \mathcal{G}(h')) = \int_0^b h_1 \sigma_{(1)}^{-1}(h') dM_{00} + h_2 \sigma_{(2)}^{-1}(h'),$$

where $\sigma = (\sigma_1, \sigma_2)$ is a continuously invertible linear operator from H_∞ onto H_∞ with inverse $\sigma^{-1} = \left({\sigma_{(1)}}^{-1}, {\sigma_{(2)}}^{-1}\right)$. The form of σ is as follows:

$$\sigma_1(h)(u) = h_1(u) E(ZY(u)) - E\left(\frac{\theta_0 \int_0^b Y h_1 dM_{00}}{1 + \theta_0 \int_0^b Y dM_{00}} ZY(u) \right)$$

$$- h_2 E\left(\frac{Y(u)}{1 + \theta_0 \int_0^b Y dM_{00}} \left(\int_0^b ZY dM_{00} - N(b) \right) \right)$$

and

$$\sigma_2(h) = h_2 E\left(-\frac{\partial^2 \log L(\theta, M_0)}{\partial \theta^2} |_{(\theta_0, A_0)} \right) - E\left(\frac{\int_0^b Y h_1 dM_{00}}{1 + \theta_0 \int_0^b Y dM_{00}} \left(\int_0^b ZY dM_{00} - N(b) \right) \right),$$

where

$$-\frac{\partial^2 \log L(\theta, M_0)}{\partial \theta^2}\Big|_{(\theta_0, A_0)} = n^{-1}\sum_{i=1}^{n}\int_0^b \left(\frac{N_i(u)}{1+\theta_0 N_i(u)}\right)^2 dN_i(u) - N_i(b)\left(\frac{\int_0^b Y_i dM_{00}}{1+\theta_0\int_0^b Y dM_{00}}\right)^2$$

$$+2\theta_0^{-3}\left(\ln\left(1+\theta_0\int_0^b Y_i dM_{00}\right) - \frac{\theta_0\int_0^b dM_{00}}{1+\theta_0\int_0^b Y_i dM_{00}} - \frac{1}{2}\left(\frac{\theta_0\int_0^b dM_{00}}{1+\theta_0\int_0^b Y_i dM_{00}}\right)^2\right).$$

When $\theta_0 = 0$, the last term above is defined by its limit, which is $\frac{2}{3}\left(\int_0^b Y_i dM_{00}\right)^3$.

Proof

Proof of the theorem follows from Murphy (1995).

3.4 Shared Gamma Frailty Reversed Hazards Model with Covariates

In survival studies, covariates are usually used to represent observable heterogeneity in a population. For example, Crouchley and Pickles (1995) discussed data on age at first marriage of UK volunteer sample of twins. Some of the observations were left censored, because some of the individuals were married but time of first marriage was not known. The covariate included was the gender of the twin pair, since woman tend to marry earlier than men. The proportional reversed hazards frailty models with covariates can be employed to account such variability.

Let $m_j(Z, t_j) = Z m_{0j}(t_j)\exp\{\underline{\beta}'\underline{x}_j\}$ be individual RHRs given frailty Z where $m_{0j}(t_j)$, $j = 1,2$ are the baseline reversed hazards, $\underline{\beta} = (\beta_1, \beta_2, ..., \beta_p)'$ is a px1 vector of regression parameters and $\underline{x}_j = (x_{j1}, x_{j2}, ..., x_{jp})'$, $j = 1,2$ is a px1 vector of covariates. Then proceeding as similar to shared gamma frailty reversed hazards frailty model, we get (3.6) as

$$F(t_1,t_2)=L\left(e^{\underline{\beta}'\underline{x}_1}M_{01}(t_1)+e^{\underline{\beta}'\underline{x}_2}M_{02}(t_2)\right).$$

Assuming that Z follows $Gamma\left(\frac{1}{\theta},\frac{1}{\theta}\right)$, we obtain bivariate distribution function of (T_1,T_2) as

$$F(t_1,t_2)=\left[\theta e^{\underline{\beta}'\underline{x}_1}M_{01}(t_1)+\theta e^{\underline{\beta}'\underline{x}_2}M_{02}(t_2)+1\right]^{-(1/\theta)} \tag{3.14}$$

and the marginal distribution functions as

$$F_j(t_j)=\left[\theta e^{\underline{\beta}'\underline{x}_j}M_{0j}(t_j)+1\right]^{-(1/\theta)},\ j=1,2.$$

which is in terms of cumulative baseline reversed hazard rates $M_{0j}(t_j)$, $j=1,2.$ Then distribution function of (T_1,T_2), (3.14) can be expressed in terms of marginal distribution functions as (3.10). The estimation of the parameters of the model is done via EM algorithm.

Retaining the notations and following the steps in Section 3.3, we obtain the log likelihood function for the parameters θ and $\underline{\beta}$, which can be expressed as

$$\log L(\theta,\beta)=L_1(\theta)+L_2(\underline{\beta}),$$

where $L_1(\theta)=-\frac{n}{\theta}\log\theta-n\ \log\Gamma(1/\theta)+\sum_{i=1}^{n}\left(\frac{1}{\theta}+N_i(b)-1\right)\left(\log\hat{z}_i-\frac{\hat{z}_i}{\theta}\right)$ and

$$L_2(\beta)=\sum_{i=1}^{n}\sum_{j=1}^{2}N_{ij}(b)\left(\underline{\beta}'\underline{x}_{ij}+\log\left(Y_i(t)\,dM_0(t)\right)\right)-\hat{z}_i e^{\underline{\beta}'\underline{x}_{ij}}\int_0^b Y_i(s)\,dM_0(s).$$

The E-step is to estimate z_i,

$$\hat{z}_i=\frac{(1/\theta)+N_i(b)}{(1/\theta)+\int_0^b\left(\sum_{j=1}^{2}Y_{ij}(s)e^{\underline{\beta}'\underline{x}_{ij}}\right)dM_0(s)}.$$

The M step is to obtain the estimates of θ and $\underline{\beta}$. The estimate of θ is obtained by maximising the log likelihood function $L_1(\theta)$ numerically and the estimate of $\underline{\beta}$ is obtained by maximising $L_2(\underline{\beta})$. Following the steps similar to Cox's partial likelihood we can express the partial likelihood for $\underline{\beta}$ as

$$L_2^*(\underline{\beta}) = \prod_{i=1}^{n}\prod_{j=1}^{2} \frac{e^{\underline{\beta}' \underline{x}_{ij}}}{\sum_{l=1}^{n} Y_{ij}(t_{(i)}) e^{\underline{\beta}' \underline{x}_{ij}}}. \tag{3.15}$$

The likelihood (3.15) is used to obtain the initial estimate of $\underline{\beta}$. The likelihood function (3.15) is considered by Sengupta et al. (1998) in a different way.

The estimate of the cumulative baseline reversed hazard rate is obtained from $L_2(\underline{\beta})$ as

$$\hat{M}_0(t) = \int_t^b \frac{dN_{\cdot}(s)}{\sum_i \hat{z}_i \left(\sum_{j=1}^{2} Y_{ij}(s) e^{\underline{\beta}' \underline{x}_{ij}} \right)}.$$

The steps are repeated until the convergence. By making suitable changes in the conditions of the theorems given in Parner (1998), the asymptotic properties of the estimates of θ and $\underline{\beta}$ can be established.

3.5 Data Analysis

Now we apply the model (3.10) to the Australian Twin data given in Duffy et al. (1990) which consists of information on the age at appendectomy of monozygotic (MZ) and dizygotic (DZ) twins as explained in Section 3.1. The genetic effect involved in the risk of appendectomy is the frailty random variable. The individuals with missing age at onset are the left censored observations. We consider pair of twins with uncensored age at onset and individuals having age at onset less than 11 are considered as left censored observations. The data consists of 203 pairs of MZ twins and 167 pairs of DZ twins. Of these MZ twin pairs, there

are 37 female pairs and 166 male pairs. Among the DZ twins, there are 82 same-sex female pairs, 19 same-sex male pairs and rest of them are opposite-sex pairs. The maximum likelihood estimate of θ for MZ and DZ twins is given in Table 3.1. The value of the frailty variable is estimated and those for MZ female pairs, MZ male pairs, DZ same-sex female pairs and DZ same-sex male pairs are given in Tables 3.2 - 3.5. It can be noted, from Table 3.2 – Table 3.5, that if the realization of Z is less than one, then all the members of the group tend to experience the event of interest at an earlier time, where as opposite occurs if Z is greater than one. Figures 3.1 - 3.4 show that the values of the frailty variable increase with increase in age at onset, for MZ and DZ twins. Tables 3.6-3.9 depicts values of cumulative baseline reversed hazard rate for MZ and DZ twins at different times and is decreasing as shown in Figures 3.5-3.8. The null hypothesis $H_0 : \theta = 0$ of independence between individuals within a pair is tested at 1% level of significance. The value of the test statistic for MZ twins and DZ twins is given in Table 3.10, which implies that individuals within a pair are dependent.

Table 3.1 Estimate of θ for twins

	MZ	DZ
Female	1.0036	0.47926
Male	0.47663	0.40887

Table 3.2 Estimate of frailty variable, $\hat{z}_i$, for MZ female pairs

Family ID	Twin 1		Twin 2		$\hat{z}_i$	Family ID	Twin 1		Twin 2		$\hat{z}_i$
	t_{i1}	δ_{i1}	t_{i2}	δ_{i2}			t_{i1}	δ_{i1}	t_{i2}	δ_{i2}	
337	26	1	21	1	1.7281	4257	27	1	31	1	2.1232
682	38	1	16	1	1.2489	4319	21	1	19	1	1.3848
759	11	0	11	0	0.1225	4628	11	1	18	1	0.54233
799	21	1	12	1	0.77413	5133	11	0	11	0	0.1225
1180	12	1	15	1	0.61453	9309	16	1	14	1	0.72309
1275	39	1	43	1	2.5118	9331	11	0	11	0	0.1225
1349	29	1	26	1	2.0318	10252	17	1	25	1	1.2377
1457	18	1	21	1	1.2424	11052	11	0	11	0	0.1225
1490	14	1	14	1	0.62411	12742	24	1	15	1	1.075
1957	20	1	14	1	0.86301	15196	14	1	17	1	0.74578
2113	20	1	19	1	1.3134	15420	19	1	11	0	0.37832
2135	18	1	15	1	0.87687	15438	20	1	18	1	1.1846
2176	40	1	52	1	2.6162	15737	11	0	28	1	0.41743
2944	18	1	11	1	0.54233	15879	12	1	42	1	0.85902
3200	11	1	11	0	0.24543	16000	19	1	40	1	1.6379
3264	11	1	21	1	0.59677	16234	12	1	11	0	0.28586
3872	45	1	56	1	2.848	16325	40	1	20	1	1.7719
3956	11	1	11	1	0.36837	20309	17	1	18	1	0.97438
3976	11	0	11	0	0.1225						

Table 3.3 Estimate of frailty variable, $\hat{z}_i$, for some MZ male pairs

Family ID	Twin 1		Twin 2		$\hat{z}_i$	Family ID	Twin 1		Twin 2		$\hat{z}_i$
	t_{i1}	δ_{i1}	t_{i2}	δ_{i2}			t_{i1}	δ_{i1}	t_{i2}	δ_{i2}	
42	17	1	17	1	0.94868	2034	22	1	11	0	0.55541
154	11	1	13	1	0.56621	2089	20	1	21	1	1.2242
265	15	1	16	1	0.81624	2114	18	1	18	1	1.0335
414	15	1	18	1	0.88742	2707	25	1	18	1	1.2057
524	14	1	17	1	0.78985	2867	15	1	25	1	1.0115
560	20	1	53	1	1.4804	2995	30	1	32	1	1.6589
574	45	1	18	1	1.3368	3111	24	1	17	1	1.1234
580	12	1	21	1	0.77163	3591	15	1	24	1	0.99386
621	15	1	19	1	0.91854	3620	34	1	35	1	1.7743
630	11	0	11	0	0.26241	3727	11	0	11	0	0.26241
660	22	1	17	1	1.0957	3840	13	1	16	1	0.7285
668	21	1	11	1	0.72645	3988	31	1	26	1	1.564
688	19	1	19	1	1.1221	4107	11	0	12	1	0.40418
744	21	1	12	1	0.77163	4148	16	1	15	1	0.81624
752	20	1	13	1	0.82889	4285	38	1	17	1	1.2512
776	11	0	12	1	0.40418	4305	38	1	19	1	1.3931
874	11	0	20	1	0.54335	4378	11	0	11	0	0.26241
895	19	1	18	1	1.076	4395	28	1	20	1	1.3523
898	12	1	11	0	0.40418	4446	16	1	22	1	1.0334
909	19	1	16	1	0.97309	4457	22	1	38	1	1.5202
910	11	0	25	1	0.57224	4486	24	1	13	1	0.86675
1062	13	1	11	0	0.42804	4505	40	1	11	1	0.80404
1066	30	1	25	1	1.5298	4520	36	1	16	1	1.1661
1091	22	1	27	1	1.4033	4680	11	0	11	0	0.26241
1142	26	1	18	1	1.2181	4682	34	1	11	0	0.60025
1814	12	1	18	1	0.72532	4745	11	0	16	1	0.48536
1818	14	1	17	1	0.78985	4765	11	0	21	1	0.54918
1826	20	1	18	1	1.1116	4840	17	1	12	1	0.70326
1828	24	1	12	1	0.7949	4995	11	0	14	1	0.44093
1859	13	1	11	1	0.56621	5089	48	1	48	1	1.8976
1899	11	0	11	0	0.26241	5097	16	1	15	1	0.81624
1911	11	0	11	0	0.26241	5107	23	1	11	0	0.56131

Table 3.4 Estimate of frailty variable, $\hat{z}_i$, for DZ female pairs

Family ID	Twin 1		Twin 2		$\hat{z}_i$	Family ID	Twin 1		Twin 2		$\hat{z}_i$
	t_{i1}	δ_{i1}	t_{i2}	δ_{i2}			t_{i1}	δ_{i1}	t_{i2}	δ_{i2}	
212	24	1	25	1	1.3694	4710	20	1	58	1	1.3696
298	34	1	22	1	1.3824	5290	29	1	16	1	1.0712
385	11	0	11	0	0.27151	5350	27	1	43	1	1.6126
589	14	1	13	1	0.67907	5394	20	1	29	1	1.2547
661	11	0	49	1	0.62446	5395	32	1	43	1	1.6788
725	21	1	11	0	0.547	5418	26	1	14	1	0.93885
741	23	1	19	1	1.1489	5433	45	1	23	1	1.5039
823	12	1	30	1	0.84402	5458	25	1	25	1	1.3898
843	11	0	48	1	0.62362	5491	13	1	11	0	0.44561
862	16	1	17	1	0.86368	5501	11	1	11	0	0.40163
892	11	0	11	0	0.27151	5503	16	1	15	1	0.80457
1063	11	0	11	0	0.27151	5587	24	1	30	1	1.445
1618	16	1	11	0	0.48867	5670	23	1	16	1	1.0072
1825	12	1	15	1	0.66631	5699	14	1	59	1	1.0228
1830	26	1	11	0	0.58664	5754	14	1	14	1	0.69646
1947	17	1	21	1	1.007	5778	22	1	38	1	1.4121
2104	11	0	17	1	0.50597	5951	17	1	17	1	0.90498
2208	22	1	30	1	1.3514	8853	18	1	17	1	0.9431
2706	15	1	11	0	0.47869	9261	23	1	70	1	1.5522
2977	28	1	20	1	1.2475	10247	54	1	52	1	1.8963
3256	11	1	11	0	0.40163	11713	11	0	21	1	0.547
3619	11	0	12	1	0.41879	11789	33	1	14	1	0.97189
3686	42	1	24	1	1.523	11886	18	1	14	1	0.81582
4248	22	1	47	1	1.4439	11909	35	1	21	1	1.352
4360	16	1	22	1	0.97679	12010	20	1	27	1	1.2402
4533	12	1	26	1	0.82614	12055	33	1	21	1	1.3306
4655	33	1	35	1	1.639	13874	25	1	58	1	0.81767

Table 3.5 Estimate of frailty variable $\hat{z}_i$ for DZ male pairs

Family ID	Twin 1		Twin 2		$\hat{z}_i$	Family ID	Twin 1		Twin 2		$\hat{z}_i$
	t_{i1}	δ_{i1}	t_{i2}	δ_{i2}			t_{i1}	δ_{i1}	t_{i2}	δ_{i2}	
134	11	0	11	0	0.35308	9363	21	1	25	1	1.4675
206	18	1	25	1	1.4261	13247	16	1	18	1	1.1572
541	16	1	22	1	1.1975	14674	14	1	26	1	1.1709
2673	17	1	17	1	1.2061	15043	24	1	15	1	1.1537
3310	12	1	12	1	0.69586	15092	35	1	13	1	1.0898
4061	16	1	16	1	1.035	15249	11	0	26	1	0.71412
4598	11	1	11	0	0.49744	16347	16	1	13	1	0.89801
4657	24	1	52	1	1.6024	20451	26	1	13	1	1.0675
5003	11	0	11	0	0.35308	20616	15	1	11	0	0.59494
5446	17	1	24	1	1.3199						

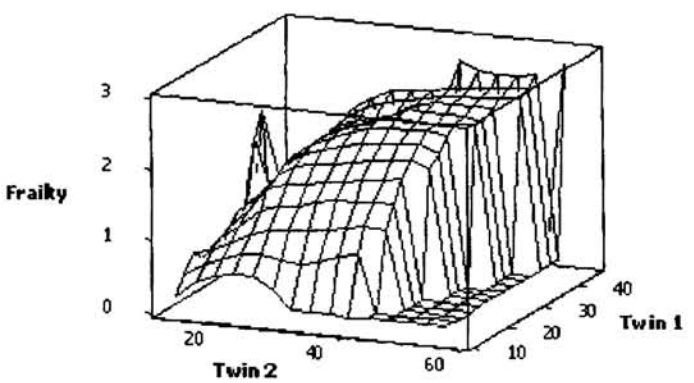

Figure 3.1 Plot of the estimate of frailty variable, $\hat{z}_i$, for MZ female pairs

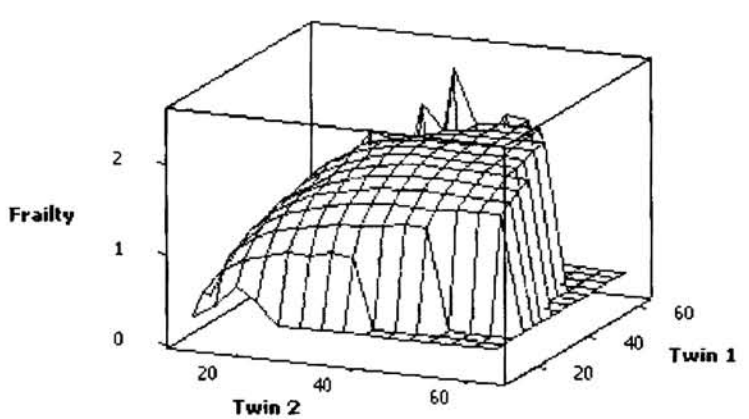

Figure 3.2 Plot of the estimate of frailty variable, $\hat{z}_i$, for MZ male pairs

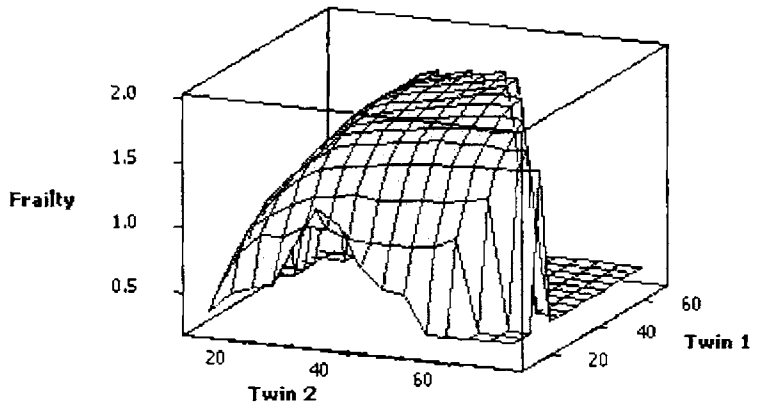

Figure 3.3 Plot of the estimate of frailty variable, $\hat{z}_i$, for DZ female pairs

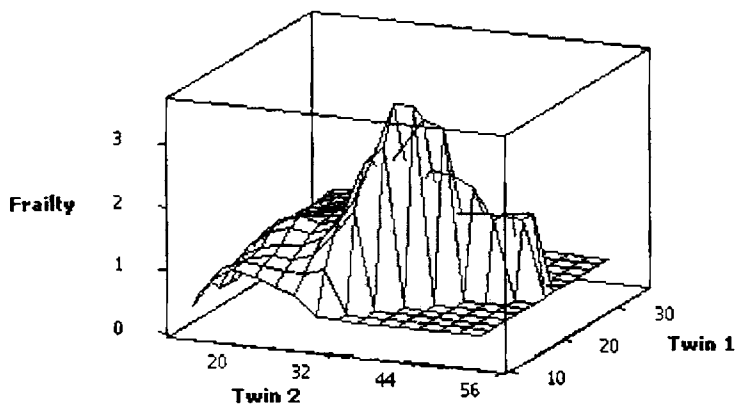

Figure 3.4 Plot of the estimate of frailty variable, $\hat{z}_i$, for DZ male pairs

Table 3.6: Estimates of the cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of MZ female twins

Age t	$\hat{M}_0(t)$	Age t	$\hat{M}_0(t)$	Age t	$\hat{M}_0(t)$	Age t	$\hat{M}_0(t)$
11	3.5689	19	0.71167	27	0.23811	40	0.12135
12	2.4185	20	0.57327	28	0.21726	42	0.073211
14	1.9024	21	0.45575	29	0.19659	43	0.057382
15	1.461	24	0.33	31	0.17675	45	0.042159
16	1.2451	25	0.30551	38	0.15771	52	0.027568
17	1.1191	26	0.28175	39	0.13912	56	0.013514
18	0.95972						

Table 3.7: Estimates of the cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of MZ male twins

Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)
11	2.9487	21	0.59449	30	0.21366	40	0.050075
12	2.6183	22	0.5312	31	0.18946	42	0.040377
13	2.191	23	0.47259	32	0.15857	45	0.03398
14	1.9795	24	0.439	33	0.13623	48	0.030795
15	1.5863	25	0.36715	34	0.11458	49	0.021401
16	1.3362	26	0.33266	35	0.096989	50	0.018288
17	1.1108	27	0.29105	36	0.079903	53	0.015193
18	0.93356	28	0.27745	37	0.069808	55	0.009068
19	0.77708	29	0.2387	38	0.066456	63	0.006024
20	0.65498						

Table 3.8: Estimates of the cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of DZ female twins

Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)
11	2.7992	21	0.75692	32	0.23619	47	0.077004
12	2.4844	22	0.66663	33	0.22766	48	0.06365
13	2.0408	23	0.54017	34	0.20289	49	0.057001
14	1.8905	24	0.47071	35	0.17905	50	0.043912
15	1.5622	25	0.42689	36	0.16353	52	0.037442
16	1.4305	26	0.37563	37	0.14827	54	0.031051
17	1.2145	27	0.33599	38	0.14072	58	0.024735
18	1.032	28	0.31671	42	0.12597	59	0.018475
19	0.93006	29	0.29795	43	0.11153	67	0.012253
20	0.87245	30	0.27082	45	0.090627	70	0.006098

Table 3.9: Estimates of the cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of DZ male twins

Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)	Age t	$\hat{M}_0$(t)
11	2.2406	15	1.1054	21	0.38313	26	0.13876
12	1.9715	16	0.92491	22	0.34179	35	0.05379
13	1.58	17	0.62013	24	0.3024	52	0.026316
14	1.2125	18	0.47116	25	0.20057		

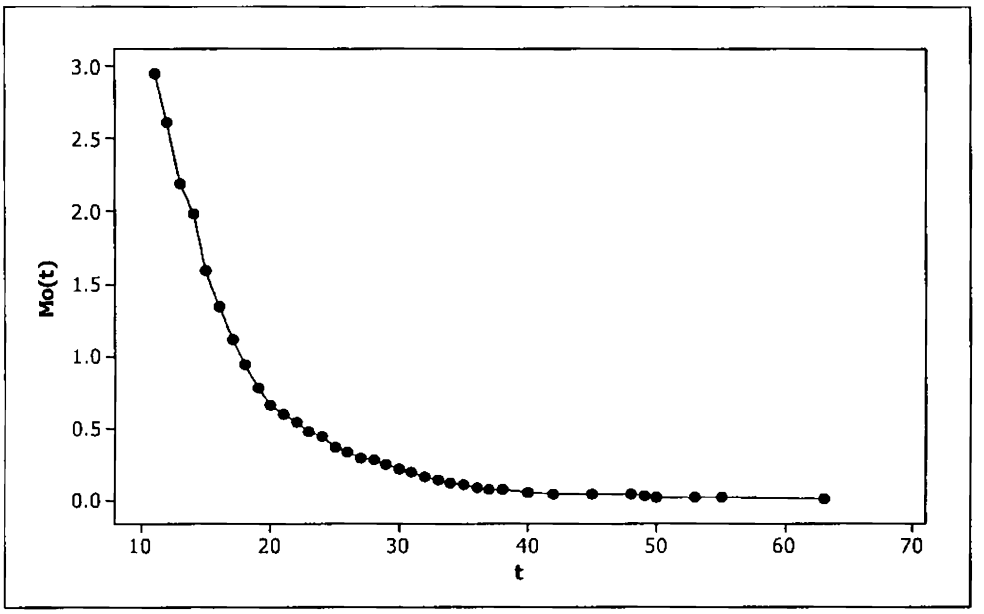

Figure 3.5 Estimate of cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of MZ male pairs

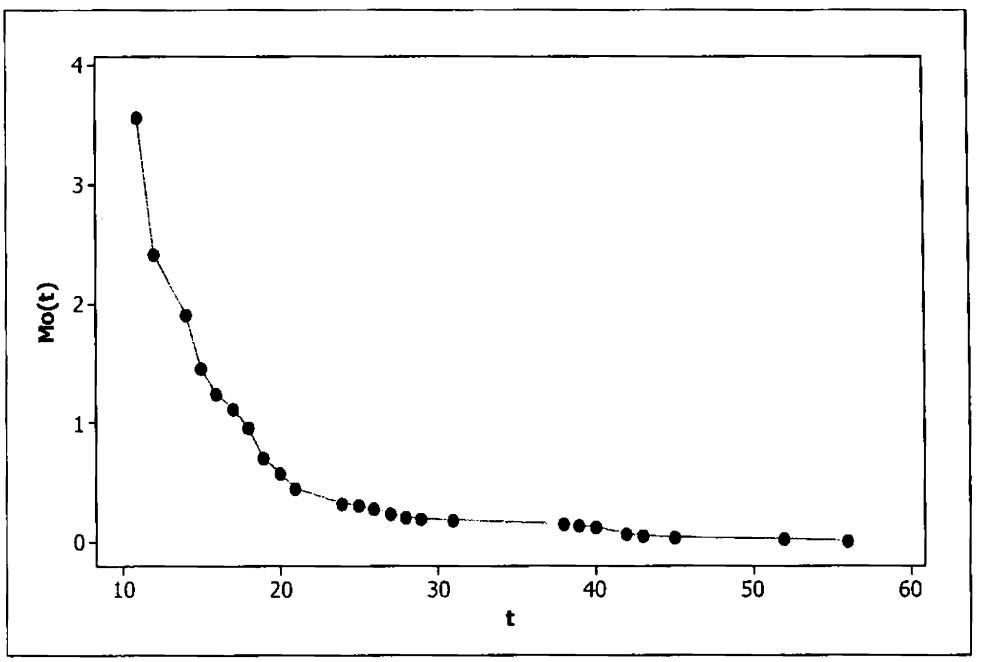

Figure 3.6 Estimate of cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of MZ female pairs

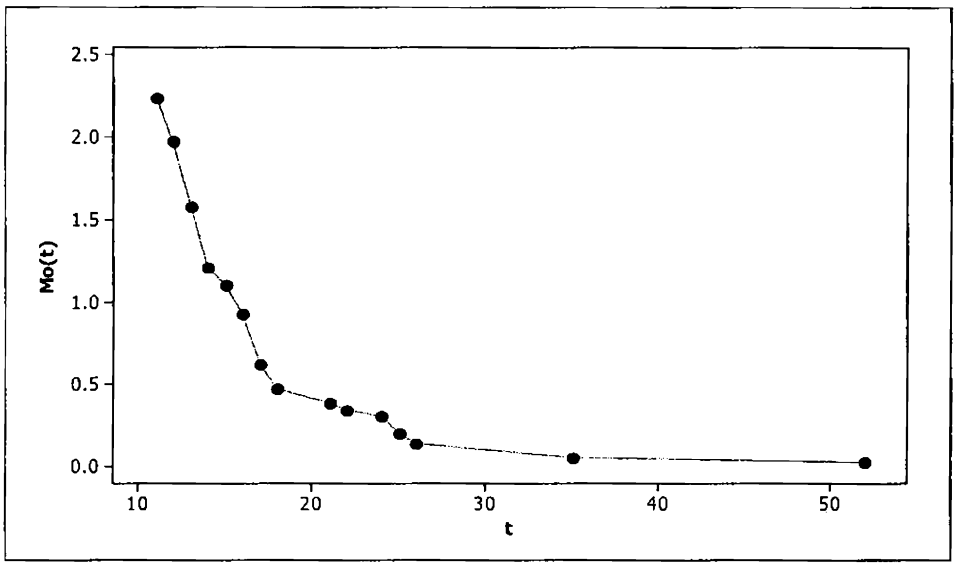

Figure 3.7 Estimate of cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of DZ male pairs

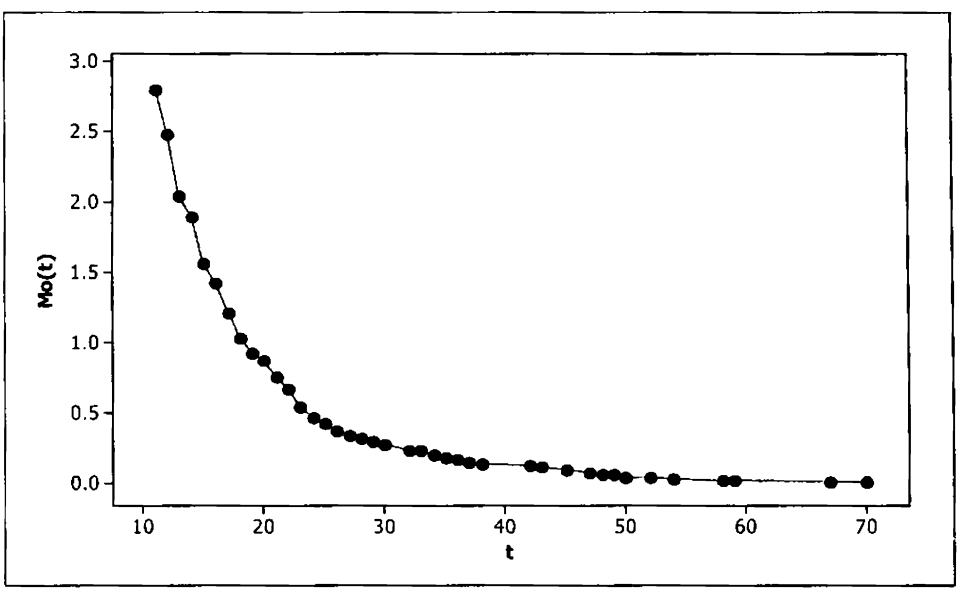

Figure 3.8 Estimate of cumulative baseline reversed hazard rate, $\hat{M}_0(t)$, for different age at onset of DZ female pairs

Table 3.10 Value of likelihood ratio test statistic for twins

	MZ	DZ
Female	1206.8	3602
Male	8004	152.44

3.6 Conclusion

The proportional reversed hazards frailty model, which is useful to model the dependence through common random effect in the context of left censoring was introduced. The shared frailty reversed hazards model was developed by considering the distribution of frailty variable as gamma distribution. The properties of the shared gamma frailty reversed hazards model were studied. The estimation of the parameter of the shared gamma frailty reversed hazards model via EM algorithm was discussed and properties of the estimators were studied. The shared gamma frailty reversed hazards model with covariates was also discussed. Finally, with monozygotic and dizygotic data in Duffy et al. (1990), the applicability of the model was well illustrated.

Chapter Four

Bivariate Correlated Gamma Frailty Reversed Hazards Model[‡]

4.1 Introduction

The shared frailty model describes the association between subjects within subgroups. It assumes the unobserved factors to be same within the subgroups, which may not always reflect reality. For example, it may be inappropriate to assume that all partners in a subgroup share all their unobserved risk factors. In the shared frailty model, the dependence between lifetimes within the subgroup is based on a marginal distribution of frailty. But in many applications, it is important to specify different marginal frailty distributions for related individuals, for example, in the analysis of unlike-sex twins, brothers and sisters, relatives from different generations etc. This is because the distribution of frailty may be different for males and females or individuals taken from different generations. Assuming gamma distribution as the distribution of frailty random variables, Yashin and Iachine (1995) introduced correlated gamma frailty model. The model is useful for the analysis of either complete or right censored bivariate lifetime data, but it is not appropriate for the analysis of left censored bivariate lifetime data. So in order to incorporate the situations where the bivariate lifetime data is left censored and individuals in a group have different but correlated frailties, we introduce bivariate correlated gamma frailty reversed hazards model.

[‡] Some results are published in the journal Metrika (see Sankaran and Gleeja (2007b)) and some other results are summarized in Sankaran and Gleeja (2008b) and communicated.

The rest of the chapter is organized in the following way. Section 4.2 introduces bivariate correlated gamma frailty reversed hazards model. The extension of the proposed model to the multivariate set up is given in Section 4.3. In Section 4.4, we present estimation of the parameters of the bivariate correlated gamma frailty reversed hazards model. The model is applied to dizygotic data in Duffy et al. (1990), in Section 4.5. Finally, a brief summary is given in Section 4.6.

4.2 Bivariate Correlated Gamma Frailty Reversed Hazards Model

Let $T=(T_1,T_2)$ be a nonnegative random vector representing lifetimes of two related individuals with an absolutely continuous distribution function $F(t_1,t_2)$ in the support of $D=[0,b_1]\times[0,b_2]$ where (b_1,b_2) is such that $b_j<\infty$ and $b_j=\inf\{t \mid F_j(t)=1\}$, $j=1,2$. Let $m_j(Z_j,t_j)=Z_j m_{0j}(t_j)$, $j=1,2$ be their individual RHRs given frailties Z_1 and Z_2 where $m_{0j}(t_j)$ are the baseline reversed hazard rate of T_j, $j=1,2$. We assume that Z_1 and Z_2 are correlated and lifetimes (T_1,T_2) are conditionally independent given frailties Z_1 and Z_2.

A bivariate correlated gamma frailty reversed hazards model can be constructed when the frailties Z_1 and Z_2 follow gamma distributions. In this case, the frailties of two individuals have different variances σ_1^2 and σ_2^2 and a correlation coefficient ρ. The following theorem specifies the bivariate correlated gamma frailty reversed hazards model, useful for the analysis of left censored data, which is analogous to the model given by Yashin and Iachine (1997) using hazard rates.

Theorem 4.1

For the proportional reversed hazards model of bivariate lifetime distribution, there exist a bivariate gamma frailty distribution with different

marginals that allows for representation of the bivariate distribution function $F(t_1,t_2)$ as

$$F(t_1,t_2)=\left[F_1(t_1)\right]^{1-\rho\frac{\sigma_1}{\sigma_2}}\left[F_2(t_2)\right]^{1-\rho\frac{\sigma_2}{\sigma_1}}\left[\left[F_1(t_1)\right]^{-\sigma_1^2}+\left[F_2(t_2)\right]^{-\sigma_2^2}-1\right]^{-\frac{\rho}{\sigma_1\sigma_2}} \quad (4.1)$$

where $0\le\rho\le\min\left(\frac{\sigma_2}{\sigma_1},\frac{\sigma_1}{\sigma_2}\right)$.

Proof

Let Y_i, $i=0,1,2$ be three independent gamma distributed random variables $\left(\Gamma(k_i,\lambda),i=0,1,2\right)$. Let α be a real positive number. Let $Z_1=Y_0+Y_1$ and $Z_2=\alpha(Y_0+Y_2)$. Then the random variables Z_1 and Z_2 are gamma distributed and correlated. Let us assume that Z_1 and Z_2 have means equal to one, variances σ_1^2 and σ_2^2 and a correlation coefficient ρ. Then parameters of gamma distributions are obtained from these assumptions as

$$\lambda=\frac{1}{\sigma_1^2},\ \alpha=\frac{\sigma_2^2}{\sigma_1^2},\ k_0=\frac{\rho}{\sigma_1\sigma_2},\ k_1=\frac{1}{\sigma_1^2}-\frac{\rho}{\sigma_1\sigma_2},\ \text{and}\ k_2=\frac{1}{\sigma_2^2}-\frac{\rho}{\sigma_1\sigma_2}. \quad (4.2)$$

Since $k_i\ge 0$, $i=0,1,2$, we get $0\le\rho\le\min\left(\frac{\sigma_2}{\sigma_1},\frac{\sigma_1}{\sigma_2}\right)$.

We can write (3.1) as

$$F\left(t_1,t_2\,|\,Z_1,Z_2\right)=\exp\left\{-Y_0\left[M_{01}(t_1)+\alpha M_{02}(t_2)\right]-Y_1\left[M_{01}(t_1)\right]-Y_2\left[\alpha M_{02}(t_2)\right]\right\}.$$

Integrating out Y_0, Y_1 and Y_2 we obtain $F(t_1,t_2)$ as

$$F(t_1,t_2)=\left[1+\frac{M_{01}(t_1)+\alpha M_{02}(t_2)}{\lambda}\right]^{-k_0}\left[1+\frac{M_{01}(t_1)}{\lambda}\right]^{-k_1}\left[1+\frac{\alpha M_{02}(t_2)}{\lambda}\right]^{-k_2}. \quad (4.3)$$

Substituting (4.2) in (4.3), we get $F(t_1,t_2)$ in terms of cumulative baseline RHRs as

$$F(t_1,t_2)=\left[1+\sigma_1^2 M_{01}(t_1)+\sigma_2^2 M_{02}(t_2)\right]^{-\frac{\rho}{\sigma_1\sigma_2}}$$
$$\left[1+\sigma_1^2 M_{01}(t_1)\right]^{-\left(\frac{1}{\sigma_1^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}\left[1+\sigma_2^2 M_{02}(t_2)\right]^{-\left(\frac{1}{\sigma_2^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}.$$

Further, we have the marginal distributions as

$$F_1(t_1)=\left[1+\frac{M_{01}(t_1)}{\lambda}\right]^{-(k_0+k_1)}=\left[1+\sigma_1^2 M_{01}(t_1)\right]^{-\left(\frac{1}{\sigma_1^2}\right)} \qquad (4.4)$$

and

$$F_2(t_2)=\left[1+\frac{\alpha M_{02}(t_2)}{\lambda}\right]^{-(k_0+k_2)}=\left[1+\sigma_2^2 M_{02}(t_2)\right]^{-\left(\frac{1}{\sigma_2^2}\right)}. \qquad (4.5)$$

Substituting (4.2), (4.4) and (4.5) in (4.3) we get (4.1), which is the correlated gamma frailty reversed hazards model in terms of marginal distribution functions.

Remark 4.1 The model (4.1) can be obtained from (2.34) if we substitute $\alpha_1=1-2\rho(\sigma_1/\sigma_2)$, $\alpha_2=1-2\rho(\sigma_2/\sigma_1)$, $\beta_1=\beta_2=1$ and

$$m(u,v)-k_1(u,v)k_2(u,v)=\frac{\rho\sigma_1\sigma_2 m_1(u)m_2(v)\left[F_1(u)\right]^{-\sigma_1^2}\left[F_2(v)\right]^{-\sigma_2^2}}{\left[\left[F_1(u)\right]^{-\sigma_1^2}+\left[F_2(v)\right]^{-\sigma_2^2}-1\right]^2}.$$

Remark 4.2 Substituting $\sigma_1^2=\sigma_2^2=\theta$ in (4.1), we obtain

$$F(t_1,t_2)=\left[F_1(t_1)\right]^{1-\rho}\left[F_2(t_2)\right]^{1-\rho}\left[\left[F_1(t_1)\right]^{-\theta}+\left[F_2(t_2)\right]^{-\theta}-1\right]^{-\rho/\theta}.$$

Remark 4.3 For $\sigma_1^2=\sigma_2^2=\theta$ and $\rho=1$, (4.1) reduces to (3.10), the shared gamma frailty reversed hazards model.

The identifiability of correlated gamma frailty reversed hazards model is established in the following theorem.

Theorem 4.2

Let $F(t_1,t_2)$ be a known bivariate distribution function given by (4.1) with parameters $\rho>0$, $\sigma_1^2>0$ and $\sigma_2^2>0$ and let the marginal reversed hazard rate $m_i(b_i)\neq 0$, $i=1,2$. Then the model is identifiable.

Proof

Let $F(t_1,t_2)$ be a known bivariate distribution function given by (4.1). Then $m_1(t_1)$, $m_2(t_2)$ and a function $\phi(t_1,t_2)$ are known, where $m_1(t_1)$ and $m_2(t_2)$ are marginal reversed hazard rate of T_1 and T_2 respectively and

$$\phi(t_1,t_2)=\frac{\partial^2 \log F(t_1,t_2)}{\partial t_1 \partial t_2}$$

$$=\frac{\rho\sigma_1\sigma_2 m_1(t_1)m_2(t_2)F_1(t_1)^{-\sigma_1^2}F_2(t_2)^{-\sigma_2^2}}{\left[F_1(t_1)^{-\sigma_1^2}+F_2(t_2)^{-\sigma_2^2}-1\right]^2}. \quad (4.6)$$

Since $F_i(b_i)=1$ and $m_i(b_i)\neq 0$, $i=1,2$, we have (4.6) as

$$\phi(b_1,b_2)=\rho\sigma_1\sigma_2 m_1(b_1)m_2(b_2),$$

which gives

$$\rho\sigma_1\sigma_2=\frac{\phi(b_1,b_2)}{m_1(b_1)m_2(b_2)}. \quad (4.7)$$

Let us define another function $\psi(t_1,t_2)$ as

$$\psi(t_1,t_2)=\frac{\phi(t_1,t_2)}{m_1(t_1)m_2(t_2)\rho\sigma_1\sigma_2}$$

$$=\frac{F_1(t_1)^{-\sigma_1^2}F_2(t_2)^{-\sigma_2^2}}{\left[F_1(t_1)^{-\sigma_1^2}+F_2(t_2)^{-\sigma_2^2}-1\right]^2}. \quad (4.8)$$

$\psi(t_1,t_2)$ is a known function and we obtain logarithmic derivatives of (4.8) as

$$\eta_i(t_1,t_2)=\frac{\partial \log \psi(t_1,t_2)}{\partial t_i}$$

$$=-\sigma_i^2 m_i(t_i)\left[1-\frac{2F_i(t_i)^{-\sigma_i^2}}{\left[F_1(t_1)^{-\sigma_1^2}+F_2(t_2)^{-\sigma_2^2}-1\right]}\right],\ i=1,2. \qquad (4.9)$$

Now, from (4.9), we obtain

$$\sigma_i^2=\frac{\eta_i(b_1,b_2)}{m_i(b_i)},\ i=1,2. \qquad (4.10)$$

Thus we obtain, from (4.7) and (4.10),

$$\rho=\frac{\phi(b_1,b_2)}{m_1(b_1)m_2(b_2)\sigma_1\sigma_2},$$

which completes the proof.

4.3 Multivariate Correlated Gamma Frailty Reversed Hazards Model

Multivariate lifetime models are important in the analysis of lifetime data with several events and in the analysis of lifetimes of related individuals. Accordingly, we can extend the correlated gamma frailty reversed hazards model for more than two lifetimes. Let T_i, $i=1,2,...,n$ be the life spans of n related individuals, $m_i(Z_i,t_i)=Z_i m_{0i}(t_i)$ be their individual RHRs given frailties Z_i where $m_{0i}(t_i)$ are the underlying baseline reversed hazards for $i=1,2,...,n$. Let $T_1,T_2,...,T_n$ be conditionally independent given frailties $Z_1,Z_2,...,Z_n$. In the following, we give the multivariate correlated gamma frailty reversed hazards model.

Theorem 4.3

For the proportional reversed hazards of the n-variate distribution function, there exist an n-variate gamma frailty distribution with different marginals that allows for representation of distribution function $F(t_1,t_2,...,t_n)$ as

$$F(t_1,t_2,...,n)=\left[\sum_{i=1}^{n}\left[F_i(t_i)\right]^{-\sigma_i^2}-n+1\right]^{-\frac{\rho_{ij}}{\sigma_i\sigma_j}}\prod_{i=1}^{n}\left[F_i(t_i)\right]^{1-\rho_{ij}\frac{\sigma_i}{\sigma_j}},$$

where ρ_{ij} are correlation coefficient between Z_i and Z_j, and σ_i^2 are the variances of Z_i, $i,j=1,2,...,n$, $i\neq j$ and $\rho_{ij}/\sigma_i\sigma_j$ is a constant not depending on i,j.

Proof

Let $Y_i, i=0,1,2,...,n$, be $n+1$ independent gamma distributed random variables $\left(\Gamma(k_i,\lambda), i=0,1,2,...,n\right)$. Let us define a set of random variables Z_i as $Z_i=\alpha_i(Y_0+Y_i)$, $i=1,2,...,n$ where α_i is a real positive number with $\alpha_1=1$. Then Z_i's are gamma distributed and dependent random variables. Let us assume that Z_i's have means equal to one and variances σ_i^2. Let the correlation coefficient between Z_i and Z_j be ρ_{ij}, $i,j=1,2,..,n$, $i\neq j$ and $\frac{\rho_{ij}}{\sigma_i\sigma_j}$ is a constant not depending on i,j. Then the parameters of the distribution of Y_i are $\lambda=\frac{1}{\sigma_1^2}$,

$\alpha_i=\frac{\sigma_i^2}{\sigma_1^2}$, $k_0=\frac{\rho_{ij}}{\sigma_i\sigma_j}$, $i,j=1,2,..,n$, $i\neq j$, $k_i=\frac{1}{\sigma_i^2}-\frac{\rho_{ij}}{\sigma_i\sigma_j}$, $i,j=1,2,..,n$, $i\neq j$

and $\frac{\rho_{ij}}{\sigma_i\sigma_j}=\frac{\rho_{kl}}{\sigma_k\sigma_l}$, $i\neq j$, $k\neq l$, $i,j,k,l=1,2,...,n$.

We have

$$F_i(t_i\mid Z_i)=\exp\left\{-\int_{t_i}^{b_i}m_i(Z_i,u)\,du\right\},\ i=1,2,...,n$$

and

$$F(t_1,t_2,...t_n\mid Z_1,Z_2,...,Z_n)=\exp\left\{-\sum_{i=1}^{n}\int_{t_i}^{b_i}m_i(Z_i,u)\,du\right\}.$$

Now, following the steps similar to the bivariate case, we get

$$F(t_1,t_2,...,t_n)=\left[\sum_{i=1}^{n}\left[F_i(t_i)\right]^{-\sigma_i^2}-n+1\right]^{-\frac{\rho_{ij}}{\sigma_i\sigma_j}}\prod_{i=1}^{n}\left[F_i(t_i)\right]^{1-\rho_{ij}\frac{\sigma_i}{\sigma_j}}. \qquad (4.11)$$

Remark 4.4 If the marginal frailty distributions of the multivariate model are identical ($\sigma_i^2 = \sigma^2$, $i = 1,2,...,n$) and all associations among frailty variables in the multivariate correlated gamma frailty reversed hazards model (4.11) are described by one parameter ρ, then

$$F(t_1,t_2,...,t_n) = \left[\sum_{i=1}^{n}\left[F_i(t_i)\right]^{-\sigma^2} - n + 1\right]^{-\rho/\sigma^2} \prod_{i=1}^{n}\left[F_i(t_i)\right]^{1-\rho}.$$

Remark 4.5 If $\sigma_i^2 = \theta$, $i = 1,2,...,n$ and $\rho_{ij} = 1, i, j = 1,2,...,n$, $i \neq j$ in the model (4.11), then we get

$$F(t_1,t_2,...,t_n) = \left[\sum_{i=1}^{n}\left[F_i(t_i)\right]^{-\theta} - n + 1\right]^{-1/\theta},$$

which is shared gamma frailty reversed hazards model extended to the multivariate setup.

4.4 Estimation of Correlated Gamma Frailty Reversed Hazards Model

We consider the situation where the data is left censored. The lifetime vector $T = (T_1, T_2)$ and censoring time vector $U = (U_1, U_2)$ are defined on a common probability space $(\Omega, \mathcal{F}, P)$. Then the observable random vectors are given by $\tilde{T} = (\tilde{T}_1, \tilde{T}_2)$ and $\delta = (\delta_1, \delta_2)$ where $\tilde{T}_j = max(T_j, U_j)$ and $\delta_j = I(T_j = \tilde{T}_j)$, $j = 1,2$. Suppose now that $\tilde{T}_i = (\tilde{T}_{1i}, \tilde{T}_{2i})$ and $\delta_i = (\delta_{1i}, \delta_{2i})$, $i = 1,2,...,n$ is an i.i.d. sample, each $(\tilde{T}_i, \delta_i)$ having the same distribution as $(\tilde{T}, \delta)$.

Let us define the counting process $\mathbf{Y}(t)$ as $\mathbf{Y}(t) = (Y_{ij}(t),\ i = 1,2,...,n,\ j = 1,2)$ where $Y_{ij}(t) = I\{\tilde{T}_{ij} \leq t\}$ and the counting process $\mathbf{N}(t)$ as $\mathbf{N}(t) = (N_{ij}(t),\ i = 1,2,...,n,\ j = 1,2)$ where $N_{ij}(t) = I\{\tilde{T}_{ij} \leq t, \delta_{ij} = 1\}$, $i = 1,2,...,n$ and $j = 1,2$.

We obtain the likelihood function $L(\rho,\sigma_1,\sigma_2)$ as

$$L(\rho,\sigma_1,\sigma_2)=$$

$$\prod_{i=1}^{n} p_0(w_{i0})\,p_1(w_{i1})\,p_2(w_{i2})\prod_{j=1}^{2}\prod_{t}\left(z_{ij}Y_{ij}(t)\,dM_0(t)\right)^{\Delta N_{ij}(t)}\exp\left\{-z_{ij}\int_0^b Y_{ij}(s)\,dM_0(s)\right\},$$

where $p_j(.)$ denotes the p.d.f of W_j, $j=0,1,2$ and $b=\max(b_1,b_2)$. Using binomial expansion (since $N_{ij}(u)\in\{0,1\}$, $j=1,2$, $i=1,2,...,n$, all binomial coefficients are equal to one), we can write

$$z_{i1}^{\,N_{i1}(b)}=\left(w_{i0}+w_{i1}\right)^{N_{i1}(b)}=\sum_{k=0}^{N_{i1}(b)} {w_{i0}}^{N_{i1}(b)-k}{w_{i1}}^{k}$$

and

$$z_{i2}^{\,N_{i2}(b)}=\left[\frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}\left(w_{i0}+w_{i2}\right)\right]^{N_{i2}(b)}=\left(\frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}\right)^{N_{i2}(b)}\sum_{h=0}^{N_{i2}(b)} {w_{i0}}^{N_{i2}(b)-h}{w_{i2}}^{h},\ i=1,2,...,n.$$

Thus

$$z_{i1}^{\,N_{i1}(b)}z_{i2}^{\,N_{i2}(b)}=\left(\frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}\right)^{N_{i2}(b)}\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} {w_{i0}}^{N_{i1}(b)+N_{i2}(b)-k-h}{w_{i1}}^{k}{w_{i2}}^{h}. \qquad (4.12)$$

Now using (4.12) and integrating out W_j, $j=0,1,2$, we get the likelihood function as

$$L(\rho,\sigma_1,\sigma_2)=\prod_{i=1}^{n}\prod_{j=1}^{2}\prod_{t}\left(Y_{ij}(t)\,dM_0(t)\right)^{\Delta N_{ij}(t)}\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G, \qquad (4.13)$$

where G is the product of gamma functions given as

$$G=\frac{\Gamma\left(N_{i1}(b)+N_{i2}(b)-k-h+\frac{\rho}{\sigma_1\sigma_2}\right)\Gamma\left(k+\frac{1}{\sigma_1^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}\right)\Gamma\left(h+\frac{1}{\sigma_2^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}\right)}{\Gamma\left(\frac{\rho}{\sigma_1\sigma_2}\right)\Gamma\left(\frac{1}{\sigma_1^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}\right)\Gamma\left(\frac{1}{\sigma_2^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}\right)\left(1+\sigma_1^{\,2}M_{i1}(0)\right)^{k+\frac{1}{\sigma_1^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}}}$$

$$\frac{\left(\sigma_1^{\,2}\right)^{N_{i1}(b)}\left(\sigma_2^{\,2}\right)^{N_{i2}(b)}}{\left(1+\sigma_2^{\,2}M_{i2}(0)\right)^{h+\frac{1}{\sigma_2^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}}\left(1+\sigma_1^{\,2}M_{i1}(0)+\sigma_2^{\,2}M_{i2}(0)\right)^{N_{i1}(b)+N_{i2}(b)-k-h+\frac{\rho}{\sigma_1\sigma_2}}}$$

and $M_{ij}(0)=\int_0^b Y_{ij}(s)dM_0(s)$, $i=1,2,...,n$, $j=1,2$.

We can use the EM algorithm to maximize (4.13). The E-step is to obtain the estimates of frailty variables as

$$\hat{w}_{0i}=\frac{\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G.\sigma_1^{\,2}.\left(\frac{N_{i1}(b)+N_{i2}(b)-k-h+\frac{\rho}{\sigma_1\sigma_2}}{1+\sigma_1^{\,2}M_{i1}(0)+\sigma_2^{\,2}M_{i2}(0)}\right)}{\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G},$$

$$\hat{w}_{1i}=\frac{\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G.\left(\frac{k+\frac{1}{\sigma_1^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}}{\frac{1}{\sigma_1^{\,2}}+M_{i1}(0)}\right)}{\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G}$$

and

$$\hat{w}_{2i}=\frac{\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G.\left(\frac{h+\frac{1}{\sigma_2^{\,2}}-\frac{\rho}{\sigma_1\sigma_2}}{\frac{1}{\sigma_2^{\,2}}+\frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}M_{i2}(0)}\right)}{\sum_{k=0}^{N_{i1}(b)}\sum_{h=0}^{N_{i2}(b)} G}, \text{ for } i=1,2,..,n.$$

Then estimates of z_{1i} and z_{2i} are

$$\hat{z}_{1i}=\hat{w}_{0i}+\hat{w}_{1i} \quad \text{and} \quad \hat{z}_{2i}=\frac{\sigma_2^{\,2}}{\sigma_1^{\,2}}(\hat{w}_{0i}+\hat{w}_{2i}), \text{ for } i=1,2,..,n.$$

The M step is to obtain the estimates of the parameters ρ, σ_1 and σ_2 from the likelihood function (4.13) and to obtain estimate of $M_0(t)$ as

$$\hat{M}_0(t)=\int_t^b \frac{dN_{..}(s)}{\sum_{i=1}^{n}\sum_{j=1}^{2}\hat{z}_{ij}Y_{ij}(s)},$$

where $N..(t)=\sum_{i=1}^{n}\sum_{j=1}^{2}N_{ij}(t)$.

Then we obtain

$$\hat{M}_{ij}(0)=\int_0^b Y_{ij}(s)\,d\hat{M}_0(s),\ \text{for } i=1,2,..,n,\ j=1,2.$$

The EM steps are repeated until the estimates converge.

The consistency and asymptotic normality of the estimators follows from Parner (1998) by incorporating appropriate changes in the conditions.

Remark 4.6 For the correlated gamma frailty model with covariates, individual RHRs given frailties Z_1 and Z_2 is $m_j\left(Z_j,t_j \mid \underline{x}_j\right)=Z_j e^{\underline{\beta}'\underline{x}_j} m_{0j}(t_j)$ where $m_{0j}(t_j)$ are the baseline reversed hazard rate of T_j, $j=1,2$, $\underline{\beta}=(\beta_1,\beta_2,...,\beta_p)'$ is $px1$ vector of regression parameters and $\underline{x}_j=(x_{j1},x_{j2},...,x_{jp})'$, $j=1,2$ is $px1$ vector of covariates. Then retaining the assumptions and proceeding similar to that of correlated gamma frailty reversed hazards model in Section 4.2, we get the distribution function $F(t_1,t_2)$ in terms of cumulative baseline RHRs as

$$\begin{aligned}F(t_1,t_2)=&\left[1+\sigma_1^2 e^{\underline{\beta}'\underline{x}_1}M_{01}(t_1)+\sigma_2^2 e^{\underline{\beta}'\underline{x}_2}M_{02}(t_2)\right]^{-\frac{\rho}{\sigma_1\sigma_2}}\\&\left[1+\sigma_1^2 e^{\underline{\beta}'\underline{x}_1}M_{01}(t_1)\right]^{-\left(\frac{1}{\sigma_1^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}\left[1+\sigma_2^2 e^{\underline{\beta}'\underline{x}_2}M_{02}(t_2)\right]^{-\left(\frac{1}{\sigma_2^2}-\frac{\rho}{\sigma_1\sigma_2}\right)}.\end{aligned}\qquad(4.14)$$

The distribution function (4.14), when expressed in terms of marginal distribution functions, reduces to (4.1). For the estimation of the parameters of the correlated gamma frailty reversed hazards model with covariates note that the cumulative reversed hazard rate is

$$M_{ij}(0)=\int_0^b Y_{ij}(s)e^{\underline{\beta}'\underline{x}_{ij}}dM_0(s),\ i=1,2,...,n,\ j=1,2.$$

Then estimation is done by maximum likelihood method via EM algorithm as given above. The estimation of parameter vector $\underline{\beta}$ is done as in Section 3.4 of Chapter 3.

4.5 Data Analysis

To depict the importance of correlated gamma frailty reversed hazards model, we apply the model (4.1) to the DZ data in Duffy et al. (1990) which consists of information on the age at appendectomy of monozygotic (MZ) and dizygotic (DZ) twins. The DZ data consist of the same-sex male pairs, same-sex female pairs and opposite-sex pairs. The individuals with missing age at onset are the left censored observations. We consider pair of twins with uncensored age at appendectomy and individuals having age at onset less than 11 are considered as left censored observation. We analyse the female-male pairs, male-male pairs and female-female pairs separately. Table 4.1 shows the estimates of the parameters of the model and we can see that correlation between the frailty variables for female-male pairs is lower than that of male-male and female-female pairs. It might be because of gender difference. Figure 4.1 shows the graph of two frailty variables for female-male pair. Figures 4.2 and 4.3 show the graphs of the two frailty variables for male-male pair and female-female pair respectively. From Figure 4.1, we see that there is some difference between the values of two frailty variables for female-male pair. But in Figures 4.2 and 4.3, the values of frailty variables are almost same. The difference between frailty variables in Figure 4.1 may be because of the effect of difference in gender.

Table 4.1 Estimates of the parameters

DZ	$\hat{\rho}$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$
Female Male	0.734332	0.260224	0.140324
Male Male	0.946333	0.978372	1.092486
Female Female	0.983083	0.676444	0.699925

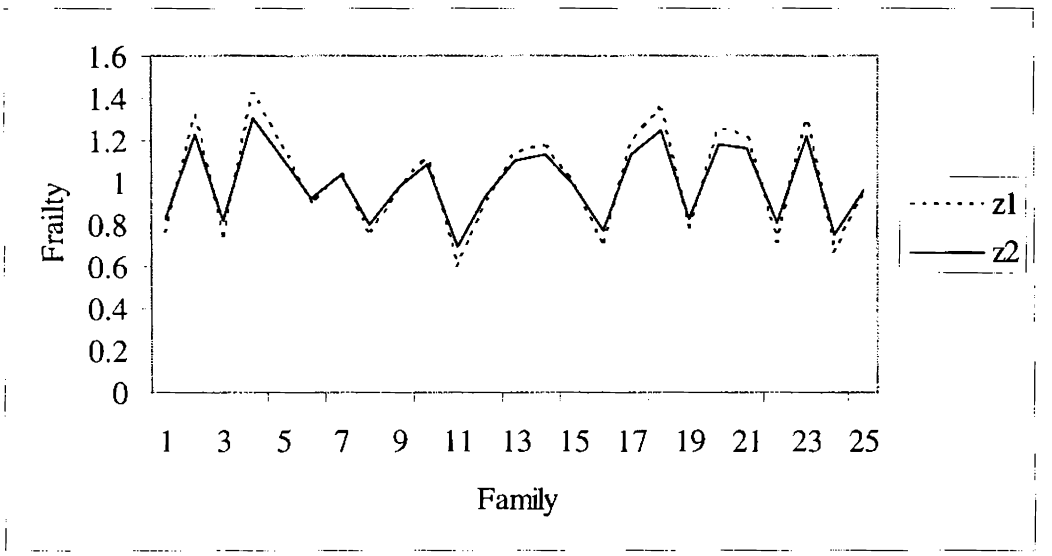

Figure 4.1 The frailty variables for the female-male pair

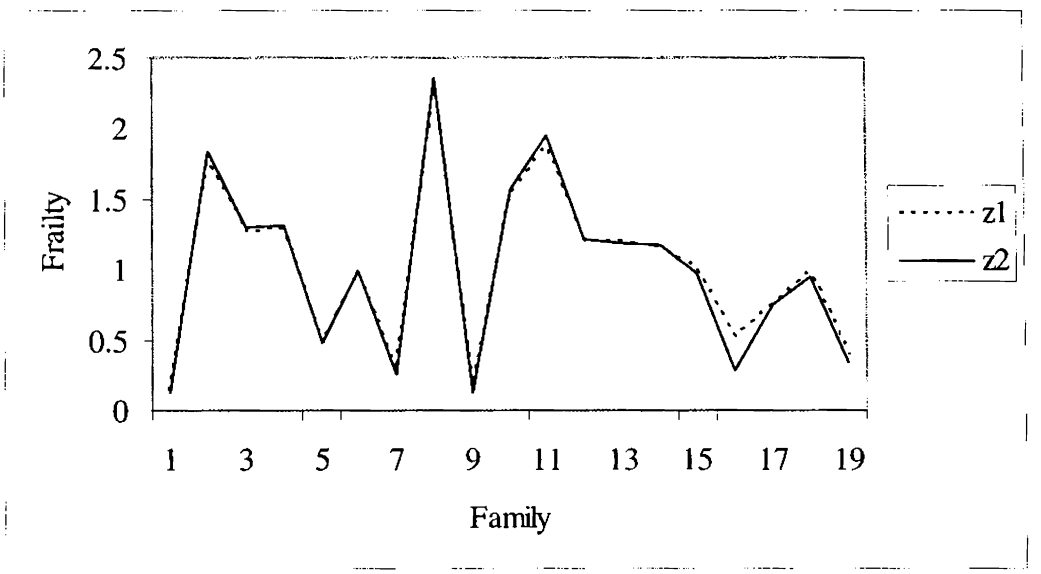

Figure 4.2 The frailty variables for the male-male pair

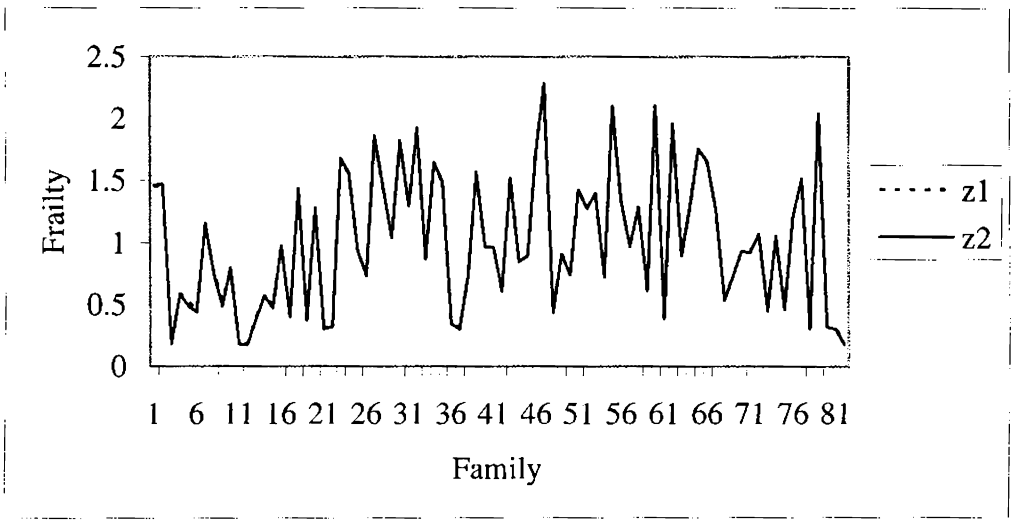

Figure 4.3 The frailty variables for the female-female pair

4.6 Conclusion

Correlated gamma frailty reversed hazards model, which is a proportional reversed hazards model, was introduced for the analysis of left censored bivariate lifetime data of related individuals. The estimation of the parameters of the correlated gamma frailty reversed hazards model via EM algorithm was discussed. Finally, using dizygotic data in Duffy et al. (1990), utility of the correlated gamma frailty reversed hazards model was illustrated.

Chapter Five

Association Measures[§]

5.1 Introduction

Dependence relations between random variables are one of the most widely studied subjects in Probability and Statistics. There are several global dependence measures such as Karl Pearson's coefficient of correlation, Kendall's τ and Spearman's rank correlation coefficient, that are commonly used to study the dependence among random variables. Although it is customary to compute a correlation coefficient, the dependence between a pair of continuous random variable is often more complex than single scalar dependence measure can reflect. Therefore, a global dependence measure such as the correlation coefficient will not convey the complete dependence structure. Accordingly, various local dependence measures are developed in literature. We discussed those measures, in Chapter one, which depends on hazard rate or survivor function and thus appropriate for the analysis of complete or right censored data.

In lifetime studies, there are many situations that require measure of association among variables, when lifetime data is left censored. In the bivariate set up, Chu et al. (2005) considered data on plasma and saliva viral loads in the Women's Interagency Human Immunodeficiency Virus (HIV) oral study. 195 seropositive subjects who had been on highly active antiretroviral therapy were randomly selected from the oral health component of the Women's Interagency HIV study to examine the correlations in HIV RNA copies per millilitre found in

[§] Some results are published in Journal of the Japan Statistical Society (see Sankaran and Gleeja (2006)) and some other results are published in Communications in Statistics-Theory and Methods (see Sankaran and Gleeja (2008a)).

serum versus saliva samples obtained within 14 days of each other. HIV RNA was measured by using the NucliSens® quantification assay, which at the time of the analyses had a lower detection limit of 80 copies per millilitre. In the data given, the HIV RNA copies per millilitre, which are below the lower detection limit of 80 copies per millilitre, is denoted as 56.57 copies per millilitre. That is, the value 56.57 indicates left censored observations and the data is left censored. We revisit the data in Section 5.5. In such situations, as pointed out by Andersen et al. (1993), hazard rates, survivor functions or mean residual life functions are not suitable for the analysis of lifetime data, but the reversed hazard rate (RHR) and distribution function are very useful. Accordingly, the reversed hazard rate and mean waiting time are more appropriate to measure the association among variables on such situations. Motivated by these facts, we introduce four association measures using bivariate distribution function, bivariate reversed hazard rates and bivariate mean waiting time.

The rest of the chapter is organised as follows. In Section 5.2, we introduce four association measures using bivariate RHR and bivariate mean waiting time and study their properties. Section 5.3 discusses association measures in terms of frailty. The estimators of the association measures are developed in Section 5.4. In Section 5.5, with two data sets, of which one is complete and other is left censored, estimation of association measures is carried out, in order to specify the usefulness of these association measures in the analysis of both types of data. We conclude the chapter in Section 5.6.

5.2. Association Measures

Let $T=(T_1,T_2)$ be a nonnegative random vector representing lifetimes of two components of a system with an absolutely continuous distribution function $F(t_1,t_2)$ in the support of $D=[0,b_1]\times[0,b_2]$ where (b_1,b_2) is such that $b_j<\infty$ and $b_j=\inf\{t\mid F_j(t)=1\}$, $j=1,2$. Suppose that the probability density function (p.d.f.) of T, $f(t_1,t_2)$ exists. Analogous to the cross

ratio function in Clayton (1978), we define a new local dependence measure in terms of reversed hazard rates, given by

$$\lambda(t_1,t_2)=\frac{F(t_1,t_2)f(t_1,t_2)}{\left[\partial F(t_1,t_2)/\partial t_1\right]\left[\partial F(t_1,t_2)/\partial t_2\right]},$$

which can be expressed as

$$\lambda(t_1,t_2)=\frac{m(t_1,t_2)}{k_1(t_1,t_2)k_2(t_1,t_2)}, \tag{5.1}$$

where $m(t_1,t_2)$ and $k_i(t_1,t_2)$, $i=1,2$ are bivariate reversed hazard rates defined in Chapter 2. Obviously $\lambda(t_1,t_2)>0$ for all $t_1,t_2>0$.

Now we study various properties of $\lambda(t_1,t_2)$.

Theorem 5.1

$\lambda(t_1,t_2)=1$ if and only if T_1 and T_2 are independent.

Proof

If T_1 and T_2 are independent, then

$$F(t_1,t_2)=F_1(t_1)F_2(t_2),$$

which leads to

$$\lambda(t_1,t_2)=\frac{F_1(t_1)F_2(t_2)f_1(t_1)f_2(t_2)}{F_2(t_2)f_1(t_1)F_1(t_1)f_2(t_2)}=1.$$

Conversely, if $\lambda(t_1,t_2)=1$, then $m(t_1,t_2)=k_1(t_1,t_2)k_2(t_1,t_2)$.

Substituting the above in (2.34), we get

$$F(t_1,t_2)=F_1(t_1)F_2(t_2),$$

which completes the proof.

Let $k_i^*(t_1,t_2)$, $i=1,2$ denote bivariate reversed rates defined in Chapter 2.

Theorem 5.2

$$\lambda(t_1,t_2)=\frac{k_1^*(t_1,t_2)}{k_1(t_1,t_2)}=\frac{k_2^*(t_1,t_2)}{k_2(t_1,t_2)}.$$

Proof

$$\text{Consider } \frac{\partial F(t_1,t_2)}{\partial t_2}=\int_0^{t_1} f(u,t_2)\,du = f_2(t_2)\,F\left(t_1|T_2=t_2\right),$$

which gives

$$F\left(t_1 \mid T_2=t_2\right)=\frac{\dfrac{\partial F(t_1,t_2)}{\partial t_2}}{f_2(t_2)}. \tag{5.2}$$

Now substituting (5.2) in (2.6), we get

$$k_1^*(t_1,t_2)=\frac{f\left(t_1 \mid T_2=t_2\right) f_2(t_2)}{\dfrac{\partial F(t_1,t_2)}{\partial t_2}}=\frac{f(t_1,t_2)}{\dfrac{\partial F(t_1,t_2)}{\partial t_2}}$$

$$=\frac{m(t_1,t_2)}{k_2(t_1,t_2)}. \tag{5.3}$$

Thus, from (5.3) and (5.1), we get

$$\frac{k_1^*(t_1,t_2)}{k_1(t_1,t_2)}=\lambda(t_1,t_2). \tag{5.4}$$

Similarly, we can obtain

$$\frac{k_2^*(t_1,t_2)}{k_2(t_1,t_2)}=\lambda(t_1,t_2), \tag{5.5}$$

which completes the proof.

Remark 5.1 $\lambda(t_1,t_2)$ can be interpreted as the ratio of the reversed hazard rate of the conditional distribution of T_1 given $T_2=t_2$ to that of T_1 given $T_2<t_2$. By symmetry, a similar interpretation holds with (T_1,T_2) interchanged.

Remark 5.2 $\lambda(t_1,t_2)=\theta+1$, where $\theta>0$, uniquely determines the shared frailty reversed hazard model (3.10).

Theorem 5.3

The bivariate RHRs and $\lambda(t_1,t_2)$ are related by

$$k_1(t_1,t_2)k_2(t_1,t_2)\left[\lambda(t_1,t_2)-1\right] = k_2(t_1,t_2)\left[k_1^*(t_1,t_2)-k_1(t_1,t_2)\right]$$
$$= k_1(t_1,t_2)\left[k_2^*(t_1,t_2)-k_2(t_1,t_2)\right].$$

Proof

From (5.4), we obtain,

$$k_2(t_1,t_2)\left[k_1^*(t_1,t_2)-k_1(t_1,t_2)\right] = k_2(t_1,t_2)\left[k_1(t_1,t_2)\lambda(t_1,t_2)-k_1(t_1,t_2)\right],$$

which gives,

$$k_2(t_1,t_2)\left[k_1^*(t_1,t_2)-k_1(t_1,t_2)\right] = k_1(t_1,t_2)k_2(t_1,t_2)\left[\lambda(t_1,t_2)-1\right].$$

Similarly, from (5.5), we get,

$$k_1(t_1,t_2)\left[k_2^*(t_1,t_2)-k_2(t_1,t_2)\right] = k_1(t_1,t_2)k_2(t_1,t_2)\left[\lambda(t_1,t_2)-1\right]$$

which completes the proof.

Definition 5.1 (Shaked, 1977) Let T_1 and T_2 have joint p.d.f. $f(t_1,t_2)$. Then $f(t_1,t_2)$ is said to be totally positive of order 2 (TP2) if for all $t_1 < t_1'$, $t_2 < t_2'$

$$f(t_1,t_2)f(t_1',t_2') \geq f(t_1,t_2')f(t_1',t_2).$$

Definition 5.2 (Brindley and Thompson, 1972) Random variables T_1 and T_2 are said to be Left Corner Set Decreasing (Increasing) (LCSD (LCSI)) if $P(T_1 \leq t_1', T_2 \leq t_2' \mid T_1 \leq t_1, T_2 \leq t_2)$ is decreasing (increasing) in (t_1,t_2) for every choice of (t_1',t_2').

Definition 5.3 (Brindley and Thompson, 1972) T_2 is Left Tail Decreasing in T_1 (LTD($T_2|T_1$)) if $P(T_2 \leq t_2 \mid T_1 \leq t_1)$ is decreasing in t_1 for all t_2.

Definition 5.4 (Lehmann, 1966) The pair (T_1,T_2) or its distribution function $F(t_1,t_2)$ is positively (negatively) quadrant dependent (PQD (NQD)) if

$$P(T_1 \le t_1, T_2 \le t_2) \ge (\le) P(T_1 \le t_1) P(T_2 \le t_2) \text{ for all } (t_1,t_2) \in D.$$

The dependence is strict if inequality holds for at least some pair.

Remark 5.3 TP2 $\Leftrightarrow$ LCSD $\Rightarrow$ LTD $(T_2 | T_1) \Rightarrow$ PQD.

From the definitions, we can easily prove that (T_1,T_2) is LCSD if $\lambda(t_1,t_2) \ge 1$ for all $(t_1,t_2) \in D$ and consequently TP2, LTD and PQD. If $\lambda(t_1,t_2) \le 1$ for all $(t_1,t_2) \in D$, then (T_1,T_2) is Left Corner Set Increasing.

Conditional expectations describe the dependence between random variables T_1 and T_2. For example, in analysis of positively associated twin data, the expected waiting time for twin 1 given the death time of twin 2 will be less than that without any information about twin 2. Motivated by this, we define an association measure as

$$\varphi_1(t_1,t_2) = \frac{E(t_1 - T_1 | T_1 \le t_1, T_2 \le t_2)}{E(t_1 - T_1 | T_1 \le t_1)}. \tag{5.6}$$

The numerator of (5.6) is the expected waiting time of T_1 given $T_1 \le t_1$ and $T_2 \le t_2$ and the denominator is the expected waiting time for T_1 given $T_1 \le t_1$. Values of $\varphi_1(t_1,t_2)$ much different from one indicate strong influence of T_2 and therefore strong association between T_1 and T_2. If T_1 and T_2 are positively associated, as t_2 increases $\varphi_1(t_1,t_2)$ should increase. We can define $\varphi_2(t_1,t_2)$ by interchanging T_1 and T_2. The measure (5.6) is time dependent, but not symmetric with respect to each coordinate. When T_1 and T_2 respectively denote first and second recurrence times of a disease for an individual, the knowledge that $T_1 \le t_1$ influences the expectation of T_2 and hence $\varphi_2(t_1,t_2)$ measures the association between T_1 and T_2.

The following theorem shows that the $\varphi_i(t_1,t_2)$ and marginal reversed hazard rates $m_i(t_i)$, $i=1,2$ determine the bivariate distribution function.

Theorem 5.4

The marginal reversed hazard rates and $\varphi_i(t_1,t_2)$, $i=1,2$ determine the bivariate distribution function of $T=(T_1,T_2)$.

Proof

Let $m_i(t_i)$ be the marginal reversed hazard rate function of T_i, $i=1,2$. It is well known that $m_i(t_i)$ determines $F_i(t_i)$ uniquely. Define the marginal waiting time of T_1 as

$$\mu_1(t_1)=E[t_1-T_1 \mid T_1\le t_1]=\frac{\int_0^{t_1}F_1(u)\,du}{F_1(t_1)}.$$

If $m_1(t_1)$ is known, then we can obtain $\mu_1(t_1)$. Thus, from (5.6), we have

$$E(t_1-T_1 \mid T_1\le t_1,T_2\le t_2)=\varphi_1(t_1,t_2)\mu_1(t_1). \tag{5.7}$$

It is to be noted that $E(t_1-T_1 \mid T_1\le t_1,T_2\le t_2)$ is nothing but the expected waiting time of T_1 given $T_1\le t_1$ and $T_2\le t_2$. From (5.7), we can obtain the conditional distribution function of T_1 given $T_2\le t_2$. Thus, the marginal distribution of T_2 and conditional distribution function of T_1 given $T_2\le t_2$ jointly provide the bivariate distribution of $T=(T_1,T_2)$.

Analogous to the dependence measure in terms of conditional probability in Anderson et al. (1992), we can define a dependence measure as

$$\beta(t_1,t_2)=\frac{P(T_1\le t_1 \mid T_2\le t_2)}{P(T_1\le t_1)}=\frac{F(t_1,t_2)}{F_1(t_1)F_2(t_2)}=\exp\{A(t_1,t_2)\}. \tag{5.8}$$

where

$$A(t_1,t_2)=\log\left[\frac{F(t_1,t_2)}{F_1(t_1)F_2(t_2)}\right]. \tag{5.9}$$

Since $A(t_1,t_2)$ is symmetric, $\beta(t_1,t_2)$ is also symmetric. When T_1 and T_2 are independent, $\beta(t_1,t_2)=1$ for all $(t_1,t_2)\in D$. $\beta(t_1,t_2)>1$ indicates positive dependence between T_1 and T_2, since $\beta(t_1,t_2)>1$ only when $T_2\le t_2$ greatly increases $P(T_1\le t_1)$.

From the representation of bivariate distribution function (2.7), we have

$$F(t_1,t_2)=F_1(t_1)F_2(t_2)\exp\{A(t_1,t_2)\}. \tag{5.10}$$

It is obvious that $A(t_1,t_2)$ or $\exp\{A(t_1,t_2)\}$ defines association between T_1 and T_2, with $A(t_1,b_2)=A(b_1,t_2)=0$ for all $(t_1,t_2)\in D$.

Differentiating (5.10) with respect to t_1 and t_2, we get

$$\lambda(t_1,t_2)=1+\frac{\dfrac{\partial^2 A(t_1,t_2)}{\partial t_1\partial t_2}}{\left[m_1(t_1)+\dfrac{\partial A(t_1,t_2)}{\partial t_1}\right]\left[m_2(t_2)+\dfrac{\partial A(t_1,t_2)}{\partial t_2}\right]}, \tag{5.11}$$

where $m_i(t_i)=\dfrac{f_i(t_i)}{F_i(t_i)}$ is the reversed hazard rate of T_i, $i=1,2$. The identity (5.11) connects $\lambda(t_1,t_2)$ with $A(t_1,t_2)$.

Remark 5.4 From the definitions (2.1), (5.8), (5.9) and (5.11), it is easy to see that

$$\frac{\partial^2\log\beta(t_1,t_2)}{\partial t_1\partial t_2}=\left(\lambda(t_1,t_2)-1\right)k_1(t_1,t_2)k_2(t_1,t_2). \tag{5.12}$$

If $\beta(t_1,t_2)\le 1$, for all $(t_1,t_2)\in D$, then (T_1,T_2) is negative quadrant dependent. In a similar manner, if $\beta(t_1,t_2)\ge 1$, for all $(t_1,t_2)\in D$, then (T_1,T_2) is positive quadrant dependent.

As in Theorem 5.4, we can prove that the marginal reversed hazard rates and $\beta(t_1,t_2)$ together determine the bivariate distribution function of $T=(T_1,T_2)$.

Now we propose another measure of association using product moment of waiting times. Product moment of waiting times is defined as

$$M(t_1,t_2)=E\left[(t_1-T_1)(t_2-T_2)\,|\,T_1\le t_1,T_2\le t_2\right]=\frac{\int_0^{t_1}\int_0^{t_2}F(u_1,u_2)du_2du_1}{F(t_1,t_2)}.$$

Then the measure of association $\alpha(t_1,t_2)$ is defined as

$$\alpha(t_1,t_2)=\frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}, \tag{5.13}$$

where $\mu_i(t_1,t_2)=E(t_i-T_i\,|\,T_1\le t_1,T_2\le t_2)$ is the expected waiting time of T_i given $T_1\le t_1$ and $T_2\le t_2$, $i=1,2$. We say that (T_1,T_2) is positively (negatively) associated if $\alpha(t_1,t_2)\ge(\le)1$ for all $(t_1,t_2)\in D$ and there is no association between T_1 and T_2 if $\alpha(t_1,t_2)=1$, for all $(t_1,t_2)\in D$.

Now we consider examples of positively and negatively associated random variables. For the bivariate uniform distribution with density

$$f(t_1,t_2)=1+\theta(2t_1-1)(2t_2-1),\ 0<t_1,t_2<1,\ \theta>0, \tag{5.14}$$

we get

$$\lambda(t_1,t_2)=\frac{A_1+5t_1t_2\theta}{A_1+4t_1t_2\theta}>1, \qquad \beta(t_1,t_2)=1+\theta(t_1-1)(t_2-1)>1,$$

$$\varphi_1(t_1,t_2)=1+\frac{\theta t_1(1-t_2)}{3+3\theta(1-t_1)(1-t_2)}>1,\ \varphi_2(t_1,t_2)=1+\frac{\theta t_2(1-t_1)}{3+3\theta(1-t_1)(1-t_2)}>1,$$

and $$\alpha(t_1,t_2)=\frac{A_2+13\theta t_1t_2}{A_2+12\theta t_1t_2}>1,$$

where $A_1 = 1+\theta(t_1-1)(t_2-1)(2t_1-1)(2t_2-1)+\theta(2-3t_1-3t_2)$ and

$$A_2 = 9+18\theta-15\theta t_1-15\theta t_2+25\theta^2 t_1t_2-10\theta^2 t_1^2t_2-10\theta^2 t_1t_2^2+4\theta^2 t_1^2t_2^2 \\ -15\theta^2 t_2+6\theta^2 t_1^2+6\theta^2 t_2^2+9\theta^2.$$

Since all the four association measures described above are greater than one for (5.14), (T_1,T_2) is positively associated or (T_1,T_2) is positively quadrant dependent. Since $\lambda(t_1,t_2)>1$, (T_1,T_2) is TP2 and hence it is LCSD.

However, for the bivaraite Dirichlet distribution with density

$$f(t_1,t_2)=6(1-t_1-t_2),\ 0<t_1,t_2<1,\ 0<t_1+t_2<1,$$

we get

$$\lambda(t_1,t_2)=\frac{B_1+4t_1t_2}{B_1+5t_1t_2}<1, \qquad \beta(t_1,t_2)=\frac{3(2-t_1-t_2)}{(3-3t_1+t_1^2)(3-3t_2+t_2^2)}<1,$$

$$\varphi_1(t_1,t_2)=\frac{2(6-2t_1-3t_2)(3-3t_1+t_1^2)}{3(2-t_1-t_2)(6-4t_1+t_1^2)}<1,$$

$$\varphi_2(t_1,t_2)=\frac{2(6-2t_2-3t_1)(3-3t_2+t_2^2)}{3(2-t_1-t_2)(6-4t_2+t_2^2)}<1 \quad \text{and} \quad \alpha(t_1,t_2)=\frac{B_2+6t_1t_2}{B_2+\frac{13}{2}t_1t_2}<1,$$

where $B_1=4-6t_1-6t_2+2t_1^2+2t_2^2$ and $B_2=18-15t_1-15t_2+3t_1^2+3t_2^2$.

In this case (T_1,T_2) is negatively associated or (T_1,T_2) is negative quadrant dependent since all association measures are less than one and since $\lambda(t_1,t_2)<1$, (T_1,T_2) is LCSI.

Theorem 5.5

The following statements are equivalent

(a) T_1 and T_2 are independent.

(b) $\lambda(t_1,t_2)=1$, for all $(t_1,t_2)\in D$

(c) $\varphi_i(t_1,t_2)=1$, $i=1,2$, for all $(t_1,t_2)\in D$

(d) $\beta(t_1,t_2)=1$, for all $(t_1,t_2)\in D$

(e) $\alpha(t_1,t_2)=1$, for all $(t_1,t_2)\in D$.

Proof

We first prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d). If T_1 and T_2 are independent, then $F(t_1,t_2)=F_1(t_1)F_2(t_2)$, which provides

$$\lambda(t_1,t_2)=\varphi_i(t_1,t_2)=\beta(t_1,t_2)=\alpha(t_1,t_2)=1,\ i=1,2.$$

The proof for the result $\lambda(t_1,t_2)=1$ implies T_1 and T_2 are independent, is given in Theorem 5.1.

To prove (c) $\Rightarrow$ (a), for $i=1$, we have

$$\mu_1(t_1,t_2)=\mu_1(t_1,b_2). \tag{5.15}$$

Differentiating (5.9) with respect to t_1, we get

$$1-k_1(t_1,t_2)\mu_1(t_1,t_2)=1-k_1(t_1,b_2)\mu_1(t_1,b_2), \tag{5.16}$$

where $k_1(t_1,t_2)$ is the reversed hazard rate of T_1 given $T_2\le t_2$.

The identities (5.15) and (5.16) together provide that $k_1(t_1,t_2)=k_1(t_1,b_2)$ or $k_1(t_1,t_2)$ is independent of T_2. Similarly, we can show that reversed hazard rate of T_2 given $T_1\le t_1$, $k_2(t_1,t_2)$, is independent of T_1. Thus T_1 and T_2 are independent. The proof for $i=2$ is similar.

When $\beta(t_1,t_2)=1$, it follows that $F(t_1,t_2)=F_1(t_1)F_2(t_2)$ and thus T_1 and T_2 are independent.

Finally, we prove (e) $\Rightarrow$ (a). From $\alpha(t_1,t_2)=1$, we get

$$\int_0^{t_1}\int_0^{t_2} F(u_1,u_2)du_2du_1=F(t_1,t_2)\mu_1(t_1,t_2)\mu_2(t_1,t_2). \tag{5.17}$$

Differentiating (5.17) with respect to t_1 and dividing by $F(t_1,t_2)$, we obtain

$$\mu_2(t_1,t_2)=k_1(t_1,t_2)\mu_1(t_1,t_2)\mu_2(t_1,t_2)+\left[\frac{d\mu_1(t_1,t_2)}{dt_1}\right]\mu_2(t_1,t_2)$$
$$+\mu_1(t_1,t_2)\left[\frac{d\mu_2(t_1,t_2)}{dt_1}\right].$$

Since $k_1(t_1,t_2)=\dfrac{1-\dfrac{d\mu_1(t_1,t_2)}{dt_1}}{\mu_1(t_1,t_2)}$, we get $\dfrac{d\mu_2(t_1,t_2)}{dt_1}\mu_1(t_1,t_2)=0$ or $\mu_2(t_1,t_2)$ is a function of t_2 only.

Similarly, we can show that $\mu_1(t_1,t_2)$ is a function of t_1 only. Therefore $k_i(t_1,t_2)$ is independent of t_j, $i,j=1,2,\ i\neq j$. Thus T_1 and T_2 are independent.

Theorem 5.6

If $\lambda(t_1,t_2)>1$ for all $(t_1,t_2)\in D$, then $\varphi_i(t_1,t_2)>1$, $i=1,2$ for all $(t_1,t_2)\in D$.

Proof

We prove the result for $i=1$. The proof for $i=2$ is similar.

From (2.10) and (5.11), we get

$$\frac{dk_1(t_1,t_2)}{dt_2}=k_1(t_1,t_2)k_2(t_1,t_2)\left[\lambda(t_1,t_2)-1\right].$$

Thus $\lambda(t_1,t_2)>1$ implies $\dfrac{dk_1(t_1,t_2)}{dt_2}>0$ or $k_1(t_1,t_2)$ is increasing in t_2. We also have

$$k_1(t_1,t_2)=\lim_{\Delta t_1\to 0}\frac{1-\left[\dfrac{F(t_1-\Delta t_1,t_2)}{F(t_1,t_2)}\right]}{\Delta t_1}.$$

When $k_1(t_1,t_2)$ is increasing in t_2, we have $\dfrac{F(t_1-\Delta t_1,t_2)}{F(t_1,t_2)}$ is decreasing in t_2 or $\dfrac{\int_0^{t_1} F(v,t_2)\,dv}{F(t_1,t_2)}$ is decreasing in t_2. Thus, we obtain $\mu_1(t_1,t_2)$ is decreasing in t_2 and hence

$$\frac{\mu_1(t_1,t_2)}{\mu_1(t_1,b_2)}=\varphi_1(t_1,t_2)>1.$$

Remark 5.5 Measures described above do not determine the distribution uniquely. This follows from the fact that $\lambda(t_1,t_2)=\beta(t_1,t_2)=\alpha(t_1,t_2)=\varphi_i(t_1,t_2)=1,\ i=1,2$ for any bivariate distribution with independent marginals.

Theorem 5.7

$\alpha(t_1,t_2)=k\ \ (k\neq 1)$, a constant independent of t_1 and t_2 if and only if the distribution function of $T=(T_1,T_2)$ has the representation

$$F(t_1,t_2)=\frac{\mu_2(t_1,b_2)}{\mu_2(t_1,t_2)}\frac{\mu_1(b_1,b_2)}{\mu_1(t_1,b_2)}\left[\frac{\mu_1(t_1,t_2)}{\mu_1(t_1,b_2)}\frac{\mu_2(t_1,b_2)}{\mu_2(b_1,b_2)}\right]^{\frac{k}{1-k}} \tag{5.18}$$

$$=\frac{\mu_1(b_1,t_2)}{\mu_1(t_1,t_2)}\frac{\mu_2(b_1,b_2)}{\mu_2(b_1,t_2)}\left[\frac{\mu_2(t_1,t_2)}{\mu_2(b_1,t_2)}\frac{\mu_1(b_1,t_2)}{\mu_1(b_1,b_2)}\right]^{\frac{k}{1-k}}. \tag{5.19}$$

Proof

Suppose that $\alpha(t_1,t_2)=k$. Then from the definition, we have

$$F(t_1,t_2)M(t_1,t_2)=\int_0^{t_1}\int_0^{t_2}F(u_1,u_2)du_2du_1. \tag{5.20}$$

Differentiating (5.20) with respect to t_1 and dividing by $F(t_1,t_2)$, we get

$$k_1(t_1,t_2)M(t_1,t_2)+\frac{dM(t_1,t_2)}{dt_1}=\mu_2(t_1,t_2). \tag{5.21}$$

Using $k_1(t_1,t_2)=\dfrac{1-\dfrac{d\mu_1(t_1,t_2)}{dt_1}}{\mu_1(t_1,t_2)}$, (5.21) can be written as

$$\left(1-\frac{d\mu_1(t_1,t_2)}{dt_1}\right)M(t_1,t_2)+\mu_1(t_1,t_2)\frac{dM(t_1,t_2)}{dt_1}=\mu_1(t_1,t_2)\mu_2(t_1,t_2). \tag{5.22}$$

Since $\alpha(t_1,t_2)=k$, we have $M(t_1,t_2)=k\mu_1(t_1,t_2)\mu_2(t_1,t_2)$, so that

$$\frac{dM(t_1,t_2)}{dt_1}=k\left(\mu_1(t_1,t_2)\frac{d\mu_2(t_1,t_2)}{dt_1}+\mu_2(t_1,t_2)\frac{d\mu_1(t_1,t_2)}{dt_1}\right),$$

and hence (5.22) reduces to

$$k\mu_2(t_1,t_2)+\mu_1(t_1,t_2)\frac{d\mu_2(t_1,t_2)}{dt_1}=\mu_2(t_1,t_2),$$

or

$$\frac{d\log\mu_2(t_1,t_2)}{dt_1}=\frac{1-k}{k\mu_1(t_1,t_2)}. \tag{5.23}$$

Using the representation for $\mu_1(t_1,t_2)$ in terms of distribution function, (5.23) leads to

$$\frac{d\log\mu_2(t_1,t_2)}{dt_1}=\left(\frac{1-k}{k}\right)\left(\frac{d\log\int_0^{t_1}F(u,t_2)\,du}{dt_1}\right),$$

or

$$\mu_2(t_1,t_2)=\left[\mu_1(t_1,t_2)F(t_1,t_2)\right]^{\frac{1-k}{k}}A_1(t_2), \tag{5.24}$$

where $A_1(t_2)$ is a function of t_2 only.

As $t_1\to b_1$, we obtain

$$A_1(t_2)=\left(\mu_1(b_1,t_2)F(b_1,t_2)\right)^{-\frac{1-k}{k}}\mu_2(b_1,t_2). \tag{5.25}$$

From (5.24) and (5.25), we get

$$\left[F(t_1,t_2)\right]^{\frac{1-k}{k}} = \frac{\mu_2(t_1,t_2)}{\mu_2(b_1,t_2)}\left[\frac{\mu_1(b_1,t_2)F(b_1,t_2)}{\mu_1(t_1,t_2)}\right]^{\frac{1-k}{k}} \tag{5.26}$$

Similarly, differentiating $M(t_1,t_2)$ with respect to t_2, we arrive at

$$\left[F(t_1,t_2)\right]^{\frac{1-k}{k}} = \frac{\mu_1(t_1,t_2)}{\mu_1(t_1,b_2)}\left[\frac{\mu_2(t_1,b_2)F(t_1,b_2)}{\mu_2(t_1,t_2)}\right]^{\frac{1-k}{k}} \tag{5.27}$$

As $t_2 \to b_2$, (5.26) becomes

$$\left[F(t_1,b_2)\right]^{\frac{1-k}{k}} = \frac{\mu_2(t_1,b_2)}{\mu_2(b_1,b_2)}\left[\frac{\mu_1(b_1,b_2)}{\mu_1(t_1,b_2)}\right]^{\frac{1-k}{k}} \tag{5.28}$$

Substituting (5.28) in (5.27), we obtain $F(t_1,t_2)$ as required in (5.18). The identity (5.27) provides the expression for $F(b_1,t_2)$ and then substituting $F(b_1,t_2)$ in (5.26), we obtain (5.19). The proof of converse part is direct.

5.3 Association Measures in terms of Frailty

Bivariate reversed hazard rates were expressed in terms of frailty variables in Chapter 3. Now, from (3.4), (3.5) and (5.1), we can represent the association measure $\lambda(t_1,t_2)$ in terms of frailty parameters as

$$\begin{aligned}\lambda(t_1,t_2) &= \frac{E\left(m_1(Z_1,t_1)m_2(Z_2,t_2)\,|\,T_1 \le t_1, T_2 \le t_2\right)}{E\left(m_1(Z_1,t_1)\,|\,T_1 \le t_1, T_2 \le t_2\right)E\left(m_2(Z_2,t_2)\,|\,T_1 \le t_1, T_2 \le t_2\right)} \\ &= \frac{E\left(Z_1Z_2\,|\,T_1 \le t_1, T_2 \le t_2\right)}{E\left(Z_1\,|\,T_1 \le t_1, T_2 \le t_2\right)E\left(Z_2\,|\,T_1 \le t_1, T_2 \le t_2\right)}.\end{aligned} \tag{5.29}$$

Remark 5.6 For the shared gamma frailty reversed hazards model (3.10), we get

$$\lambda(t_1,t_2) = \frac{E\left(Z^2\,|\,T_1 \le t_1, T_2 \le t_2\right)}{\left[E\left(Z\,|\,T_1 \le t_1, T_2 \le t_2\right)\right]^2} = 1 + \frac{V\left(Z\,|\,T_1 \le t_1, T_2 \le t_2\right)}{\left[E\left(Z\,|\,T_1 \le t_1, T_2 \le t_2\right)\right]^2}.$$

$\beta(t_1,t_2)$ can be obtained in terms of frailty parameters by substituting (3.2) and (3.3) in (5.8). To represent other association measures in terms of frailty parameters, note that

$$E(t_i - T_i \mid T_i \le t_i) = \frac{\int_0^{t_i}\int_0^{\infty} F_i(u \mid z_i)\, g(z_i)\, dz_i du}{\int_0^{\infty} F_i(t_i \mid z_i)\, g(z_i)\, dz_i}, \tag{5.30}$$

$$E(t_i - T_i \mid T_i \le t_i, T_j \le t_j) = \frac{\int_0^{t_i}\int_0^{\infty}\int_0^{\infty} F(u, t_j \mid z_1, z_2)\, g(z_1, z_2)\, dz_1 dz_2 du}{\int_0^{\infty}\int_0^{\infty} F(t_i, t_j \mid z_1, z_2)\, g(z_1, z_2)\, dz_1 dz_2}, i \ne j,\ i, j = 1,2 \tag{5.31}$$

and

$$M(t_1,t_2) = \frac{\int_0^{t_1}\int_0^{t_2}\int_0^{\infty}\int_0^{\infty} F(u_1,u_2 \mid z_1,z_2)\, g(z_1,z_2)\, dz_1 dz_2 du_2 du_1}{\int_0^{\infty}\int_0^{\infty} F(t_1,t_2 \mid z_1,z_2)\, g(z_1,z_2)\, dz_1 dz_2}. \tag{5.32}$$

Now, by substituting (5.30) and (5.31) in (5.6), we get $\varphi_1(t_1,t_2)$ in terms of frailty parameters. Similarly, we can obtain $\varphi_2(t_1,t_2)$. Substituting (5.31) and (5.32) in (5.13), we get $\alpha(t_1,t_2)$ in terms of frailty parameters.

5.4 Estimation

To apply the association measures in a practical situation we need to find the estimators of the measures. For this purpose, nonparametric estimators of the quantities involved in the measures can be used. First, we consider the estimation of joint distribution function and marginal distribution functions.

For complete sample, let (T_{1i}, T_{2i}), $i = 1,2,...,n$ be a random sample having same distribution as (T_1, T_2). Then, the estimates of the bivariate distribution

function $F(t_1,t_2)$ and marginal distribution function $F_j(t_j)$, $j=1,2$ could be obtained using the empirical distribution function as,

$$\hat{F}(t_1,t_2)=\frac{1}{n}\sum_{i=1}^{n} I(T_{1i}\le t_1,T_{2i}\le t_2), \tag{5.33}$$

and

$$\hat{F}_j(t_j)=\frac{1}{n}\sum_{i=1}^{n} I(T_{ji}\le t_j),\ j=1,2. \tag{5.34}$$

For the estimation of the distribution functions for left censored data, let $C=(C_1,C_2)$ be a pair of censoring variables with distribution function $G(t_1,t_2)$. The observable random vectors, under the bivariate left censoring, are $Y=(Y_1,Y_2)$ and $\delta=(\delta_1,\delta_2)$ where $Y_i=max(T_i,C_i)$, $i=1,2$ and $\delta_i=I(Y_i=T_i)$, $i=1,2$. Let $H(t_1,t_2)$ be the distribution function of Y. Assume that censoring mechanism is independent of failure time. This implies that $H(t_1,t_2)=F(t_1,t_2)G(t_1,t_2)$. Let $\Lambda(t_1,t_2)=(\Lambda_1(dt_1,t_2),\Lambda_2(t_1,dt_2),\Lambda_{12}(dt_1,dt_2))$ be the reversed hazard rate vector, where

$$\Lambda_1(dt_1,t_2)=\frac{P(T_1\in dt_1,T_2\le t_2)}{P(T_1\le t_1,T_2\le t_2)}=\frac{F(dt_1,t_2)}{F(t_1,t_2)}, \tag{5.35}$$

$$\Lambda_2(t_1,dt_2)=\frac{P(T_1\le t_1,T_2\in dt_2)}{P(T_1\le t_1,T_2\le t_2)}=\frac{F(t_1,dt_2)}{F(t_1,t_2)} \tag{5.36}$$

and

$$\Lambda_{12}(dt_1,dt_2)=\frac{P(T_1\in dt_1,T_2\in dt_2)}{P(T_1\le t_1,T_2\le t_2)}=\frac{F(dt_1,dt_2)}{F(t_1,t_2)}. \tag{5.37}$$

When the joint density function of (T_1,T_2), $f(t_1,t_2)$ exists, we can write $\Lambda_1(dt_1,t_2)=k_1(t_1,t_2)dt_1$, $\Lambda_2(t_1,dt_2)=k_2(t_1,t_2)dt_2$ and $\Lambda_{12}(dt_1,dt_2)=m(t_1,t_2)dt_1dt_2$, where $k_i(t_1,t_2)$, $i=1,2$ is the bivariate reversed hazard vector component given by (2.1) and $m(t_1,t_2)$ is the scalar bivariate reversed hazard rate given by (2.4).

Denote

$$H(t_1,t_2)=P(Y_1\le t_1,Y_2\le t_2), \tag{5.38}$$

$$K_1(t_1,t_2)=P(Y_1\le t_1,Y_2\le t_2,\delta_1=1), \tag{5.39}$$

$$K_2(t_1,t_2)=P(Y_1\le t_1,Y_2\le t_2,\delta_2=1) \tag{5.40}$$

and

$$K_{12}(t_1,t_2)=P(Y_1\le t_1,Y_2\le t_2,\delta_1=1,\delta_2=1). \tag{5.41}$$

From the equations (5.35) - (5.41), we obtain

$$\Lambda_1(t_1,t_2)=\int_{t_1}^{b_1}\frac{K_1(du,t_2)}{H(u,t_2)}, \qquad \Lambda_2(t_1,t_2)=\int_{t_2}^{b_2}\frac{K_2(t_1,dv)}{H(t_1,v)}$$

and $$\Lambda_{12}(t_1,t_2)=\int_{t_1}^{b_1}\int_{t_2}^{b_2}\frac{K_{12}(du,dv)}{H(u,v)}.$$

Then the distribution functions $F_i(t_i)$, $i=1,2$ and $F(t_1,t_2)$ can be written as

$$F_1(t_1)=\exp\{-\Lambda_1(t_1,b_2)\}, \tag{5.42}$$

$$F_2(t_2)=\exp\{-\Lambda_2(b_1,t_2)\} \tag{5.43}$$

and

$$F(t_1,t_2)=F_1(t_1)F_2(t_2)\exp\{\Lambda_{12}(t_1,t_2)-\Lambda_1(t_1,t_2)\Lambda_2(t_1,t_2)\}. \tag{5.44}$$

To estimate joint distribution function and marginal distribution functions in the left censored set up, suppose that $(Y_{1i},Y_{2i},\delta_{1i},\delta_{2i})$, $i=1,2,...,n$ is a random sample having the same distribution as $(Y_1,Y_2,\delta_1,\delta_2)$. Then the estimator of cumulative reversed hazard vector components is obtained as

$$\hat{\Lambda}_1(t_1,t_2)=\int_{t_1}^{b_1}\frac{\hat{K}_1(du,t_2)}{\hat{H}(u,t_2)}, \qquad \hat{\Lambda}_2(t_1,t_2)=\int_{t_2}^{b_2}\frac{\hat{K}_2(t_1,dv)}{\hat{H}(t_1,v)}$$

and $$\hat{\Lambda}_{12}(t_1,t_2)=\int_{t_1}^{b_1}\int_{t_2}^{b_2}\frac{\hat{K}_{12}(du,dv)}{\hat{H}(u,v)},$$

where

$$\hat{H}(t_1,t_2)=\frac{1}{n}\sum_{i=1}^{n} I(Y_{1i}\le t_1, Y_{2i}\le t_2),\quad \hat{K}_1(t_1,t_2)=\frac{1}{n}\sum_{i=1}^{n} I(Y_{1i}\le t_1, Y_{2i}\le t_2, \delta_{1i}=1),$$

$$\hat{K}_2(t_1,t_2)=\frac{1}{n}\sum_{i=1}^{n} I(Y_{1i}\le t_1, Y_{2i}\le t_2, \delta_{2i}=1) \text{ and}$$

$$\hat{K}_{12}(t_1,t_2)=\frac{1}{n}\sum_{i=1}^{n} I(Y_{1i}\le t_1, Y_{2i}\le t_2, \delta_{1i}=1, \delta_{2i}=1).$$

Thus, from (5.42), (5.43) and (5.44), we get the estimators of distribution functions as

$$\hat{F}_1(t_1)=\exp[-\hat{\Lambda}_1(t_1,b_2)], \tag{5.45}$$

$$\hat{F}_2(t_2)=\exp[-\hat{\Lambda}_2(b_1,t_2)] \tag{5.46}$$

and

$$\hat{F}(t_1,t_2)=\hat{F}_1(t_1)\hat{F}_2(t_2)\exp\left\{\hat{\Lambda}_{12}(t_1,t_2)-\hat{\Lambda}_1(t_1,t_2)\hat{\Lambda}_2(t_1,t_2)\right\}. \tag{5.47}$$

Thus, we derived estimates of marginal and joint distribution functions for complete sample set up and for left censored data.

Now, to estimate the association measures, we use estimates (5.33) and (5.34) for complete sample set up and (5.45), (5.46) and (5.47) for left censored data. Then we obtain the estimator of association measure, $\beta(t_1,t_2)$ as

$$\hat{\beta}(t_1,t_2)=\frac{\hat{F}(t_1,t_2)}{\hat{F}_1(t_1)\hat{F}_2(t_2)}.$$

To obtain the estimator of $\lambda(t_1,t_2)$, note that $\hat{\Lambda}_1(dt_1,t_2)$ is the estimator of the reversed hazard rate of T_1 given $T_2\le t_2$. On similar lines, we can develop the estimator of reversed hazard rate of T_1 given $T_2=t_2$. Thus we get a nonparametric estimator of $\lambda(t_1,t_2)$, by substituting these estimates in (5.4).

To derive the estimator of $\varphi_1(t_1,t_2)$, note that

$$\hat{\mu}_1(t_1,t_2)=\frac{\int_0^{t_1}\hat{F}(u,t_2)\,du}{\hat{F}(t_1,t_2)}.$$

Then we get, $\hat{\varphi}_1(t_1,t_2)=\frac{\hat{\mu}_1(t_1,t_2)}{\hat{\mu}_1(t_1,b_2)}$. Similarly we can obtain $\hat{\varphi}_2(t_1,t_2)$.

The non-parametric estimator of $\hat{\alpha}(t_1,t_2)$ is obtained as

$$\hat{\alpha}(t_1,t_2)=\frac{\hat{M}(t_1,t_2)}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)},$$

where $\hat{M}(t_1,t_2)=\frac{\int_0^{t_1}\int_0^{t_2}\hat{F}(u_1,u_2)\,du_2du_1}{\hat{F}(t_1,t_2)}$ is the nonparametric estimator of $M(t_1,t_2)$.

The strong consistency and asymptotic normality of the estimators of cumulative reversed hazard rates can easily be established by extending the proofs of those for hazard rates given in Andersen et al. (1993). Then, the asymptotic normality of estimators $\hat{F}_1(t_1),\hat{F}_2(t_2)$ and $\hat{F}(t_1,t_2)$ can be proved by functional delta method (see van der Vaart and Wellner, 1996 and Gürler, 2004). Finally, we can establish the strong consistency and asymptotic normality of the estimators $\hat{\lambda}(t_1,t_2),\hat{\beta}(t_1,t_2),\hat{\alpha}(t_1,t_2)$ and $\hat{\varphi}_i(t_1,t_2)$, $i=1,2$. The following theorems establish the strong consistency and asymptotic normality of $\hat{\alpha}(t_1,t_2)$ and proofs for other estimators are similar.

Theorem 5.8

$\left\|\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right\|_D\to 0$ almost surely. That is, $\hat{\alpha}(t_1,t_2)$ is uniformly strong consistent.

Proof

To prove the strong consistency of $\hat{\alpha}(t_1,t_2)$, we have $\left\|\hat{\Lambda}_i(t_1,t_2)-\Lambda_i(t_1,t_2)\right\|_D \to 0$ almost surely for $i=1,2$ and $\left\|\hat{\Lambda}_{12}(t_1,t_2)-\Lambda_{12}(t_1,t_2)\right\|_D \to 0$ almost surely, proof of which follows easily by extending the proofs for hazard rates given in Darbowska (1998). The strong consistency of $\hat{F}_1(t_1)$, $\hat{F}_2(t_2)$ and $\hat{F}(t_1,t_2)$ also can be proved by following the steps similar to one given in Darbowska (1998) for survivor functions. Now, we can write

$$\left\|\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right\|_D = \left\|\frac{\hat{M}(t_1,t_2)}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)}-\frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right\|_D$$

$$= \left\|\frac{\hat{M}(t_1,t_2)}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)}-\frac{\hat{M}(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}+\frac{\hat{M}(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}-\frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right\|_D$$

$$\le \left\|\frac{\hat{M}(t_1,t_2)}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)}-\frac{\hat{M}(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right\|_D + \left\|\frac{\hat{M}(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}-\frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right\|_D . \qquad (5.48)$$

Since $\left\|\hat{F}(t_1,t_2)-F(t_1,t_2)\right\|_D \to 0$ almost surely, we can establish that $\left\|\hat{M}(t_1,t_2)-M(t_1,t_2)\right\|_D \to 0$ and $\left\|\hat{\mu}_i(t_1,t_2)-\mu_i(t_1,t_2)\right\|_D \to 0$, $i=1,2$ almost surely. Now (5.48) can be written as

$$\left\|\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right\|_D \le \frac{M(t_1,t_2)}{{\mu_1}^2(t_1,t_2){\mu_2}^2(t_1,t_2)}\left\{\left\|\hat{\mu}_1(t_1,t_2)\left[\hat{\mu}_2(t_1,t_2)-\mu_2(t_1,t_2)\right]\right\|_D + \mu_2(t_1,t_2)\left\|\hat{\mu}_1(t_1,t_2)-\mu_1(t_1,t_2)\right\|_D\right\} + \frac{1}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\left\|\hat{M}(t_1,t_2)-M(t_1,t_2)\right\|_D ,$$

which provides

$$\left\|\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right\|_D \le \frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2{}^2(t_1,t_2)}\left\|\hat{\mu}_2(t_1,t_2)-\mu_2(t_1,t_2)\right\|_D$$
$$+\frac{M(t_1,t_2)}{\mu_1{}^2(t_1,t_2)\mu_2(t_1,t_2)}\left\|\hat{\mu}_1(t_1,t_2)-\mu_1(t_1,t_2)\right\|_D$$
$$+\frac{1}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\left\|\hat{M}(t_1,t_2)-M(t_1,t_2)\right\|_D.$$

Since $\left\|\hat{M}(t_1,t_2)-M(t_1,t_2)\right\|_D \to 0$ and $\left\|\hat{\mu}_i(t_1,t_2)-\mu_i(t_1,t_2)\right\|_D \to 0$, $i=1,2$ almost surely, $\left\|\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right\|_D \to 0$.

Theorem 5.9

For fixed $(t_1,t_2)\in D$, $\sqrt{n}\left(\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right)$ is asymptotically normal with zero mean.

Proof

First we prove the asymptotic normality of $\sqrt{n}\left(\hat{\Lambda}_1(t_1,t_2)-\Lambda_1(t_1,t_2)\right)$.

$$\sqrt{n}\left(\hat{\Lambda}_1(t_1,t_2)-\Lambda_1(t_1,t_2)\right)=\sqrt{n}\left(\int_{t_1}^{h_1}\frac{\hat{K}_1(du,t_2)}{\hat{H}(u,t_2)}-\int_{t_1}^{h_1}\frac{K_1(du,t_2)}{H(u,t_2)}\right)$$
$$=\sqrt{n}\left(\int_{t_1}^{h_1}\frac{\hat{K}_1(du,t_2)}{\hat{H}(u,t_2)}-\frac{\hat{K}_1(du,t_2)}{H(u,t_2)}+\frac{\hat{K}_1(du,t_2)}{H(u,t_2)}-\frac{K_1(du,t_2)}{H(u,t_2)}\right). \quad (5.49)$$

Since $\left\|\hat{H}(t_1,t_2)-H(t_1,t_2)\right\|_D \to 0$ and $\left\|\hat{K}_1(t_1,t_2)-K_1(t_1,t_2)\right\|_D \to 0$, (5.49) is asymptotically equal to

$$\sqrt{n}\left(\int_{t_1}^{h_1}\frac{K_1(du,t_2)\left(H(u,t_2)-\hat{H}(u,t_2)\right)}{H^2(u,t_2)}\right)+\sqrt{n}\left(\int_{t_1}^{h_1}\frac{\left(\hat{K}_1(du,t_2)-K_1(du,t_2)\right)}{H(u,t_2)}\right). \quad (5.50)$$

By multivariate central limit theorem, each item in the simple brackets of (5.50) converges to a normal variate with mean zero. Thus for fixed $(t_1,t_2)\in D$, the

asymptotic normality of $\sqrt{n}\left(\hat{\Lambda}_1(t_1,t_2)-\Lambda_1(t_1,t_2)\right)$ follows from the delta method using the maps $(a_1,a_2)\to a_1+a_2$ (see van der Vaart and Wellner, 1996). The asymptotic variance is given by

$$\sigma_1^2(t_1,t_2)=E\left[\left(\int_{t_1}^{b_1}\frac{I(Y_{1i}\le t_1,Y_{2i}\le t_2,\delta_{1i}=1)}{H(u,t_2)}\right)-\left(\int_{t_1}^{b_1}\frac{\left(K_1(du,t_2)\,I(Y_{1i}\le t_1,Y_{2i}\le t_2)\right)}{H^2(u,t_2)}\right)\right]^2.$$

Similarly we can prove the asymptotic normality of $\sqrt{n}\left(\hat{\Lambda}_2(t_1,t_2)-\Lambda_2(t_1,t_2)\right)$ and $\sqrt{n}\left(\hat{\Lambda}_{12}(t_1,t_2)-\Lambda_{12}(t_1,t_2)\right)$. The asymptotic variances thus obtained will be denoted by $\sigma_2^2(t_1,t_2)$ and $\sigma_{12}^2(t_1,t_2)$.

Next, we consider

$$\hat{F}(t_1,t_2)=\hat{F}_1(t_1)\,\hat{F}_2(t_2)\,\hat{F}_3(t_1,t_2),$$

where $\hat{F}_3(t_1,t_2)=\exp\left\{\hat{\Lambda}_{12}(t_1,t_2)-\hat{\Lambda}_1(t_1,t_2)\hat{\Lambda}_2(t_1,t_2)\right\}$. The asymptotic normality of $\sqrt{n}\left(\hat{\Lambda}_i(t_1,t_2)-\Lambda_i(t_1,t_2)\right)$, $i=1,2$ and $\sqrt{n}\left(\hat{\Lambda}_{12}(t_1,t_2)-\Lambda_{12}(t_1,t_2)\right)$ carries over the asymptotic normality of $\sqrt{n}\left(\hat{F}_i(t_i)-F_i(t_i)\right)$, $i=1,2$ and $\sqrt{n}\left(\hat{F}_3(t_1,t_2)-F_3(t_1,t_2)\right)$ and hence to $\sqrt{n}\left(\hat{F}(t_1,t_2)-F(t_1,t_2)\right)$.

To establish the asymptotic normality of $\sqrt{n}\left(\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right)$, write

$$\sqrt{n}\left(\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right)=\sqrt{n}\left(\frac{\hat{M}(t_1,t_2)}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)}-\frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right)$$

$$=\sqrt{n}\left(\frac{\hat{M}(t_1,t_2)}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)}-\frac{\hat{M}(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}+\frac{\hat{M}(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}-\frac{M(t_1,t_2)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right)$$

$$=\sqrt{n}\left(\frac{\hat{M}(t_1,t_2)\left[\mu_1(t_1,t_2)\mu_2(t_1,t_2)-\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)\right]}{\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)\mu_1(t_1,t_2)\mu_2(t_1,t_2)}+\frac{\left[\hat{M}(t_1,t_2)-M(t_1,t_2)\right]}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right).$$

(5.51)

Since $\underset{D}{Sup}\left\|\hat{M}(t_1,t_2)-M(t_1,t_2)\right\| \xrightarrow{a.s.} 0$ and $\underset{D}{Sup}\left\|\hat{\mu}_i(t_1,t_2)-\mu_i(t_1,t_2)\right\| \xrightarrow{a.s.} 0$, $i=1,2$, (5.51) is asymptotically equal to

$$\sqrt{n}\left(\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right)$$

$$=\sqrt{n}\left(\frac{M(t_1,t_2)\left(\mu_1(t_1,t_2)\mu_2(t_1,t_2)-\hat{\mu}_1(t_1,t_2)\hat{\mu}_2(t_1,t_2)\right)}{\mu_1^2(t_1,t_2)\mu_2^2(t_1,t_2)}-\frac{\left(\hat{M}(t_1,t_2)-M(t_1,t_2)\right)}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}\right)$$

$$=\frac{\sqrt{n}\left[\hat{M}(t_1,t_2)-M(t_1,t_2)\right]}{\mu_1(t_1,t_2)\mu_2(t_1,t_2)}-\frac{\sqrt{n}M(t_1,t_2)\left[\hat{\mu}_1(t_1,t_2)-\mu_1(t_1,t_2)\right]}{\mu_1^2(t_1,t_2)\mu_2(t_1,t_2)}$$

$$-\frac{\sqrt{n}M(t_1,t_2)\left[\hat{\mu}_2(t_1,t_2)-\mu_2(t_1,t_2)\right]}{\mu_1(t_1,t_2)\mu_2^2(t_1,t_2)}. \qquad (5.52)$$

Now,

$$\sqrt{n}\left(\hat{M}(t_1,t_2)-M(t_1,t_2)\right)=$$

$$\sqrt{n}\left[\frac{\int_0^{t_1}\int_0^{t_2}\hat{F}(u_1,u_2)\,du_2du_1}{\hat{F}(t_1,t_2)}-\frac{\int_0^{t_1}\int_0^{t_2}\hat{F}(v_1,v_2)\,dv_2dv_1}{F(t_1,t_2)}+\frac{\int_0^{t_1}\int_0^{t_2}\hat{F}(v_1,v_2)\,dv_2dv_1}{F(t_1,t_2)}-\frac{\int_0^{t_1}\int_0^{t_2}F(v_1,v_2)\,dv_2dv_1}{F(t_1,t_2)}\right]. \qquad (5.53)$$

Since $\left\|\hat{F}(t_1,t_2)-F(t_1,t_2)\right\|_D \to 0$ almost surely, (5.53) is asymptotically equal to

$$\sqrt{n}\left(\hat{M}(t_1,t_2)-M(t_1,t_2)\right)=\frac{\sqrt{n}}{F(t_1,t_2)}\left[\int_0^{t_1}\int_0^{t_2}\left(\hat{F}(u_1,u_2)-F(u_1,u_2)\right)du_2du_1\right]$$

$$-\frac{\sqrt{n}}{F^2(t_1,t_2)}\left[\int_0^{t_1}\int_0^{t_2}F(t_1,t_2)\left(\hat{F}(u_1,u_2)-F(u_1,u_2)\right)du_2du_1\right].$$

On similar lines, we can write $\sqrt{n}\left(\hat{\mu}_1(t_1,t_2)-\mu_1(t_1,t_2)\right)$ and $\sqrt{n}\left(\hat{\mu}_2(t_1,t_2)-\mu_2(t_1,t_2)\right)$ in terms of $\sqrt{n}\left(\hat{F}(t_1,t_2)-F(t_1,t_2)\right)$. Thus the asymptotic normality of $\sqrt{n}\left(\hat{\alpha}(t_1,t_2)-\alpha(t_1,t_2)\right)$ follows easily from (5.52). The asymptotic variance can be calculated, but it will be in a complex form. In practice, the bootstrap method of resampling the observed data with

$(Y_{1i}, Y_{2i}, \delta_{1i}, \delta_{2i})$, $i = 1, 2, .., n$ with replacement can be used for the estimation of variance of the estimators (see Efron and Tibshirani, 1993).

Remark 5.7 If the bivariate distribution is of the form (3.10), we can obtain a non-parametric estimator of $\lambda(t_1, t_2)$ using Kendall's (1962) coefficient of concordance. Let (T_{1i}, T_{2i}), $i = 1, 2..., n$ be a random sample from bivariate distribution (3.10). Then, for $1 \le i < j \le n$, define $X_{ij} = 1$ or $X_{ij} = -1$ according as $T_{1i} < T_{1j}$ or $T_{1i} > T_{1j}$. Define Y_{ij} similarly for T_{2i} and T_{2j} and let $Z_{ij} = X_{ij} Y_{ij}$. Set $U = \dfrac{\sum_{i=1}^{n} \sum_{j>i}^{n} Z_{ij}}{\binom{n}{2}}$. Since the probability of concordance is $\dfrac{\lambda(t_1, t_2)}{\lambda(t_1, t_2) + 1}$, U is an unbiased estimator of $\dfrac{\lambda(t_1, t_2) - 1}{\lambda(t_1, t_2) + 1}$. Thus, we propose a nonparametric estimator of $\lambda(t_1, t_2)$ by $\hat{\lambda}(t_1, t_2) = \dfrac{1 + U}{1 - U}$. The asymptotic normality of U follows from the results of Hoeffding (1948).

Remark 5.8 We can obtain the estimates of $\lambda(t_1, t_2)$, from the identity (5.29), by substituting the estimates of $E(Z_1 Z_2 \mid T_1 \le t_1, T_2 \le t_2)$, $E(Z_1 \mid T_1 \le t_1, T_2 \le t_2)$ and $E(Z_2 \mid T_1 \le t_1, T_2 \le t_2)$. The estimates of these quantities can be obtained using the conditional density function $g(z_1, z_2 \mid T_1 \le t_1, T_2 \le t_2)$. Similar technique can be used to obtain the estimates of $\beta(t_1, t_2)$, $\alpha(t_1, t_2)$ and $\varphi_i(t_1, t_2)$, $i = 1, 2$.

5.5 Data Analysis

We illustrate the use of association measures with two real life data sets, the first one is a complete sample and the second one is a left censored data. First we consider the cancer recurrence data given in Kulkarni and Rattihalli (2002). The data consist of observations on 19 patients having bladder tumours when they

entered the trial. These tumours were removed and patients were given a treatment called 'placebo pills'. At subsequent follow-up visits, any tumours found were removed, and the treatment was continued. The variables observed are time (in months) to first recurrence of tumour (T_1) and second recurrence of tumour (T_2). The estimates of the association measures at all the time points are computed using the method given in Section 5.4 and are given in Table 5.1. In Table 5.1, estimate of the variances of the estimators, computed using bootstrap method, are given in brackets. The values of estimators of association measures at different time points are greater than one, which is an indication of positive association. But at certain time points, for which t_1 is very small compare to t_2, estimates are less than one. (For example, association measures at the time points (2,17), (2,26) and (3,16).) As mentioned in Section 5.2, $\varphi_2(t_1,t_2)$ is more appropriate association measure for this data. Surface plot of the data for various association measures are given in Figures 5.1-5.5.

Table 5.1: Estimates of association measures for complete data

t_1	t_2	$\hat{\lambda}(t_1,t_2)$	$\hat{\beta}(t_1,t_2)$	$\hat{\varphi}_1(t_1,t_2)$	$\hat{\varphi}_2(t_1,t_2)$	$\hat{\alpha}(t_1,t_2)$
2	8	1.0000 (0.000)	1.58330 (0.000)	1.00000 (0.000)	0.75000 (0.000)	1.00000 (0.000)
2	15	1.0000 (0.000)	0.79167 (1.428×10^{-29})	1.00000 (0.000)	1.02440 (5.447×10^{-30})	0.28571 (1.691×10^{-29})
2	17	1.0000 (0.000)	0.89063 (0.000)	1.00000 (0.000)	0.87562 (8.349×10^{-30})	0.27273 (1.93×10^{-30})
2	26	1.0000 (0.000)	1.05560 (0.000)	1.00000 (0.000)	0.78440 (4.940×10^{-32})	0.10526 (8.923×10^{-31})
3	6	1.0000 (0.000)	1.90000 (1.428×10^{-29})	2.50000 (0.000)	1.00000 (0.000)	1.00000 (0.000)
3	9	1.3333 (0.000)	1.90000 (3.006×10^{-28})	0.62500 (1.47×10^{-29})	1.00000 (0.000)	0.57143 (0.000)
3	15	0.9333 (2.501×10^{-29})	1.10830 (3.602×10^{-29})	0.71429 (4.902×10^{-29})	1.29620 (0.000)	0.79032 (3.569×10^{-30})
3	16	1.3333 (1.961×10^{-28})	1.08570 (2.613×10^{-29})	0.62500 (0.000)	1.25470 (1.000×10^{-28})	0.94737 (4.459×10^{-30})
5	14	5.0000 (0.000)	1.23380 (3.873×10^{-29})	0.82500 (2.391×10^{-29})	1.11180 (7.904×10^{-31})	1.23460 (4.940×10^{-30})
7	10	5.0000 (3.34×10^{-29})	1.58330 (4.002×10^{-30})	0.88696 (0.000)	1.00000 (0.000)	1.22990 (1.235×10^{-28})
9	11	6.0000 (0.000)	1.35710 (0.000)	0.90000 (1.039×10^{-29})	1.00000 (0.000)	1.31940 (3.162×10^{-30})
9	17	3.2500 (3.34×10^{-29})	1.10270 (0.000)	0.96923 (0.000)	1.13890 (5.711×10^{-29})	1.04170 (2.076×10^{-29})
10	15	2.7500 (0.089279)	1.16110 (0.0058551)	0.97403 (2.204×10^{-05})	1.09090 (0.00012247)	1.10890 (0.0001738)
12	15	2.4000 (0.0125)	1.11760 (0.0017174)	1.01900 (7.124×10^{-06})	1.00000 (0.000)	1.17790 (9.395×10^{-05})
12	16	3.5000 (0.000)	1.11760 (0.000)	0.96930 (4.155×10^{-29})	1.00000 (0.000)	1.19590 (3.34×10^{-29})
16	19	17.0000 (0.000)	1.05560 (1.075×10^{-29})	0.97738 (1.045×10^{-28})	1.00000 (0.000)	1.13150 (4.155×10^{-29})
28	30	19.0000 (0.000)	1.00000 (4.013×10^{-29})	1.00000 (0.000)	1.00000 (0.000)	1.06850 (0.000)

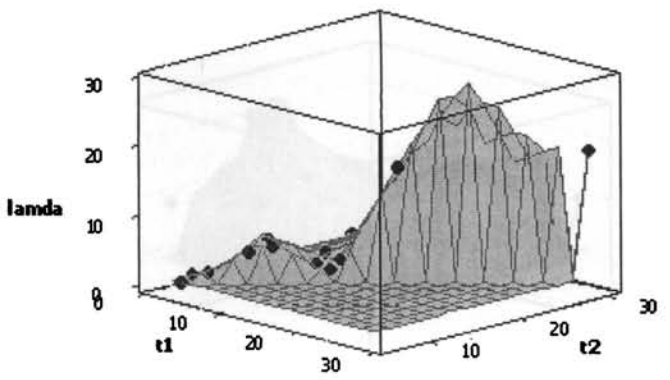

Figure 5.1: Plot of $\hat{\lambda}(t_1, t_2)$ for complete data.

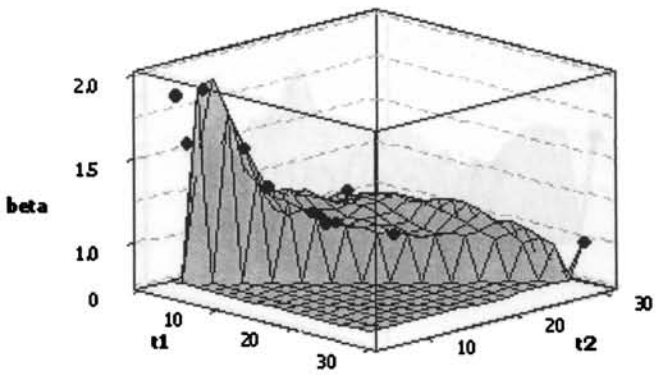

Figure 5.2: Plot of $\hat{\beta}(t_1, t_2)$ for complete data.

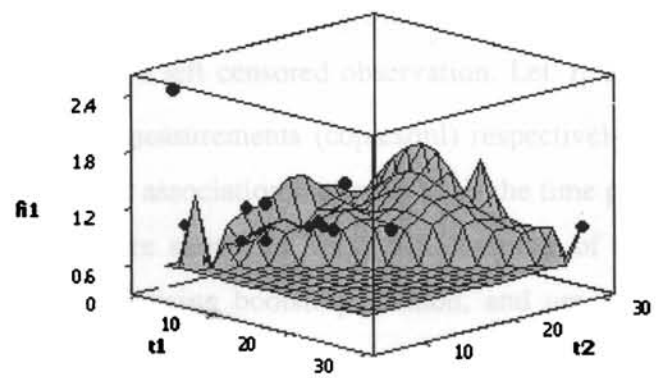

Figure 5.3: Plot of $\hat{\varphi}_1(t_1, t_2)$ for complete data.

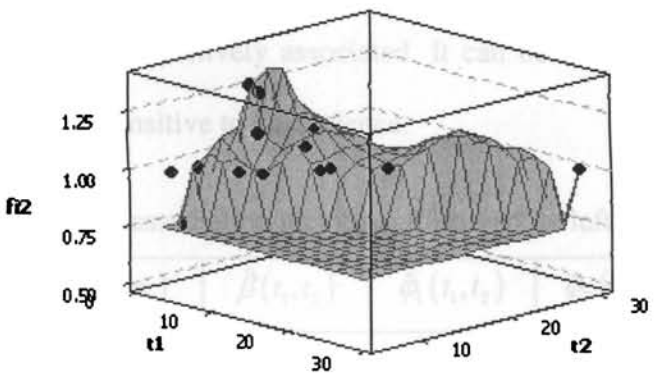

Figure 5.4: Plot of $\hat{\varphi}_2(t_1,t_2)$ for complete data.

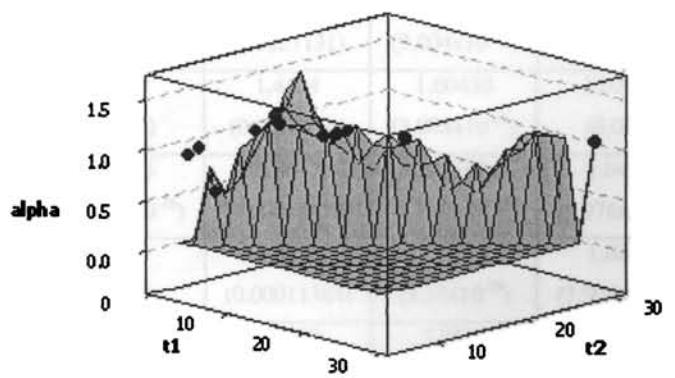

Figure 5.5: Plot of $\hat{\alpha}(t_1,t_2)$ for complete data.

Now we consider the bivariate left censored data, given in Chu et al. (2005). The data is explained in Section 5.1 and we take the value 56.57 as 80 itself and note that it is a left censored observation. Let T_1 and T_2 denote saliva and plasma viral load measurements (copies/ml) respectively. We then estimate the values of the various association measures at all the time points and estimators at certain time points are shown in Table 5.2. Estimate of the variances of the estimators, is computed using bootstrap method, and are given in brackets, in Table 5.2. Surface plot of the various association measures are given in Figures 5.6-5.10. From the figures, we can see that behaviours of $\lambda(t_1,t_2)$, $\beta(t_1,t_2)$, $\alpha(t_1,t_2)$, $\varphi_1(t_1,t_2)$ and $\varphi_2(t_1,t_2)$ exhibits same pattern. The association measures $\lambda(t_1,t_2)$, $\beta(t_1,t_2)$, $\varphi_1(t_1,t_2)$, $\varphi_2(t_1,t_2)$ and $\alpha(t_1,t_2)$ have values greater than

one at almost all the time points. This indicates that (T_1,T_2) is positive quadrant dependent or (T_1,T_2) is positively associated. It can be noted that, in both data sets, $\lambda(t_1,t_2)$ is more sensitive to dependence.

Table 5.2: Estimates of association measures for bivariate left censored data.

t_1	t_2	$\hat{\lambda}(t_1,t_2)$	$\hat{\beta}(t_1,t_2)$	$\hat{\varphi}_1(t_1,t_2)$	$\hat{\varphi}_2(t_1,t_2)$	$\hat{\alpha}(t_1,t_2)$
80	80	0.000 (0.000)	1.4247 (0.00587)	1.00000 (0.000)	1.00000 (0.000)	1.00000 (0.000)
80	590000	0.000 (0.000)	1.0351 ($1.336\text{x}10^{-28}$)	1.00000 (0.000)	1.02020 ($9.564\text{x}10^{-29}$)	1.0237 ($7.904\text{x}10^{-29}$)
130	3200	55 (0.000)	1.4177 (0.0027131)	0.9943 ($9.004\text{x}10^{-30}$)	1.0311 ($1.388\text{x}10^{-28}$)	1.0069 ($4.446\text{x}10^{-29}$)
380	80	1.000 (0.000)	1.4424 (0.0030754)	1.00430 ($3.088\text{x}10^{-29}$)	1.00000 (0.000)	1.0411 ($1.045\text{x}10^{-28}$)
510	160000	48.333 ($1.462\text{x}10^{-26}$)	1.0667 (0.00014474)	1.0042 ($2.613\text{x}10^{-29}$)	1.0421 ($1.976\text{x}10^{-29}$)	1.0103 ($1.601\text{x}10^{-29}$)
1700	56000	75 (0.000)	1.0849 (0.00011384)	1.0138 ($3.34\text{x}10^{-29}$)	1.0294 ($7.904\text{x}10^{-29}$)	1.02 ($9.564\text{x}10^{-29}$)
2900	2000	38 ($5.059\text{x}10^{-29}$)	1.1912 (0.00029446)	1.0513 ($5.38\text{x}10^{-29}$)	0.99955 ($5.711\text{x}10^{-29}$)	1.0108 ($6.403\text{x}10^{-29}$)
3800	29000	146 (0.000)	1.1057 (0.00019831)	1.0169 ($9.135\text{x}10^{-29}$)	1.015 ($4.748\text{x}10^{-29}$)	1.0144 ($7.904\text{x}10^{-31}$)
4000	840000	171 (0.000)	1.0225 ($7.114\text{x}10^{-30}$)	1.0002 ($7.114\text{x}10^{-30}$)	1.0099 ($1.976\text{x}10^{-29}$)	1.0087 ($1.138\text{x}10^{-28}$)
4300	5100	130.00 (0.000)	1.1480 (0.00022893)	1.03850 ($8.349\text{x}10^{-30}$)	0.99921 ($1.783\text{x}10^{-29}$)	1.0199 ($5.059\text{x}10^{-29}$)
6700	11000	47 ($4.255\text{x}10^{-26}$)	1.0873 (0.00018845)	1.0337 ($4.748\text{x}10^{-29}$)	1.0049 ($1.549\text{x}10^{-28}$)	1.0243 ($5.711\text{x}10^{-29}$)
10000	870000	91 (0.000)	1.0107 ($1.441\text{x}10^{-05}$)	1.0021 ($6.763\text{x}10^{-29}$)	1.0089 ($7.904\text{x}10^{-31}$)	1.0088 ($1.235\text{x}10^{-30}$)
16000	410000	90.5 (0.000)	1.0161 ($7.179\text{x}10^{-05}$)	1.0057 ($5.978\text{x}10^{-30}$)	1.0025 ($9.683\text{x}10^{-30}$)	1.0101 ($9.683\text{x}10^{-30}$)
17000	17000	75 (0.000)	1.0489 ($7.437\text{x}10^{-05}$)	1.0291 ($1.112\text{x}10^{-29}$)	1 (0.000)	1.0145 ($4.446\text{x}10^{-31}$)
30000	730000	187 ($1.882\text{x}10^{-25}$)	1.0106 ($2.360\text{x}10^{-05}$)	1.0025 ($1.494\text{x}10^{-28}$)	1.0009 ($5.711\text{x}10^{-29}$)	1.0084 ($1.976\text{x}10^{-31}$)
58000	130000	89.5 (0.000)	1.0055 (3.6015e-029)	1.0092 (4.7476e-029)	1 (0.000)	1.0104 (1.899e-028)

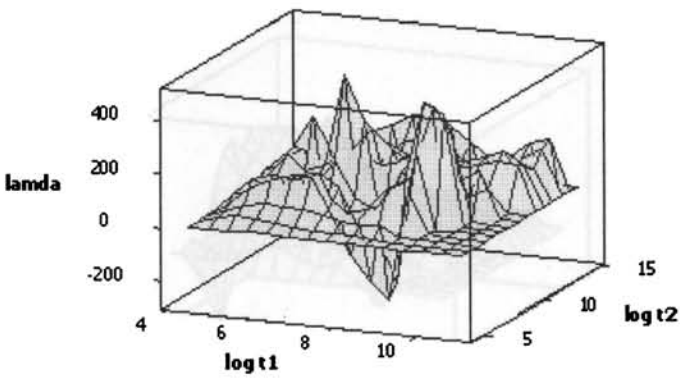

Figure 5.6: Plot of $\hat{\lambda}(t_1, t_2)$ for left censored data.

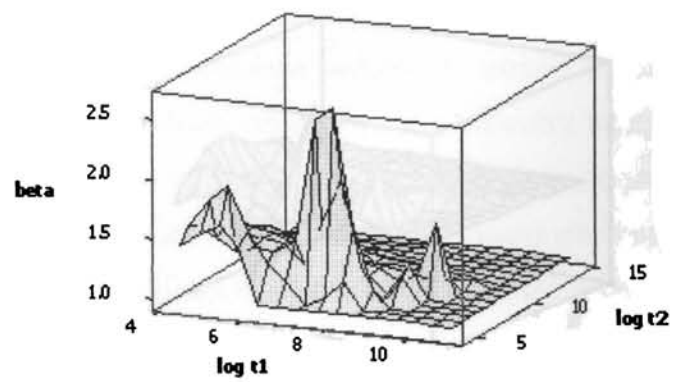

Figure 5.7: Plot of $\hat{\beta}(t_1, t_2)$ for left censored data.

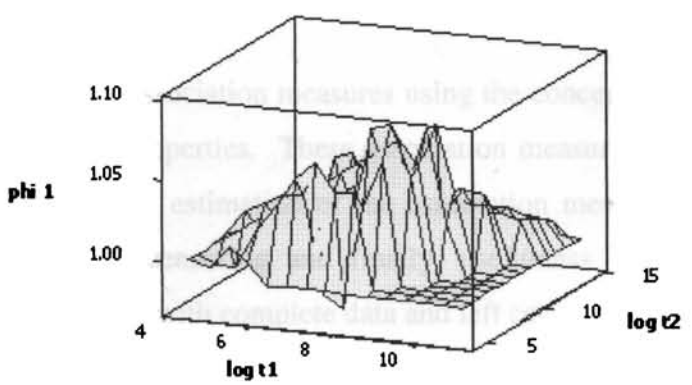

Figure 5.8: Plot of $\hat{\varphi}_1(t_1, t_2)$ for left censored data.

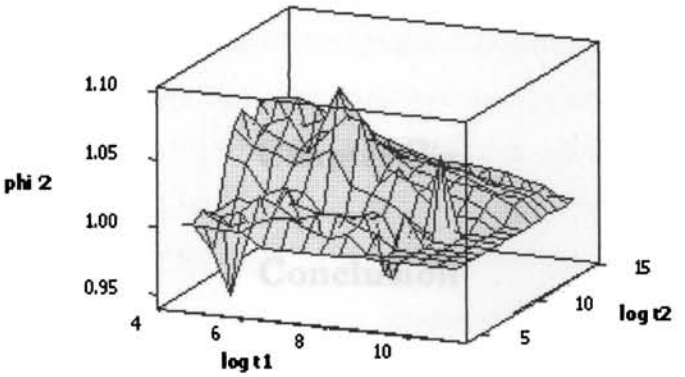

Figure 5.9: Plot of $\hat{\varphi}_2(t_1,t_2)$ for left censored data.

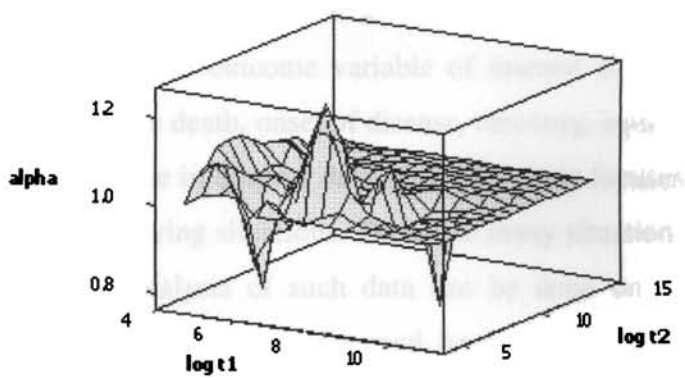

Figure 5.10: Plot of $\hat{\alpha}(t_1,t_2)$ for left censored data.

5.6 Conclusion

We developed four association measures using the concept of reversed hazard rates and studied their properties. These association measures were discussed in terms of frailty also. The estimation of the association measures was discussed under independent left censoring and finally, usefulness of these association measures were illustrated with complete data and left censored data.

Chapter Six

Conclusion

6.1 Introduction

Lifetime data analysis refers to a collection of statistical procedures for data analysis, in which the outcome variable of interest is time until an event occurs. The event may be death, onset of disease, recovery, equipment breakdown etc. The existing literature in lifetime data analysis mainly focuses on the analysis of data under right censoring situations. There are many situations where the data is left censored. The analysis of such data can be done on the basis of new stochastic models developed using reversed hazard rates. Accordingly, in the present work, we developed new stochastic models that enable us to analyze such type of data. In chapter 2, various definitions of bivariate reversed hazard rates and their importance in the analysis and modelling of lifetime data was explored. Based on bivariate reversed hazard rates discussed, a unique representation for bivariate distribution was given and a new class of distributions was proposed. The proposed model was used to develop models for the analysis of data on parallel systems. As applications of the proposed model, the frailty model and a bivariate proportional reversed hazards model were also derived.

The concept of frailty, to account for unobserved heterogeneity, is attracting increasing attention in the literature because individuals of interest usually differ in susceptibility to causes of death or disease, response to treatment and influence of risk factors. In Chapter 3, based on reversed hazard rate, we introduced proportional reversed hazards frailty models. The univariate gamma

frailty reversed hazards model and shared gamma frailty reversed hazards model were discussed. The estimation of the parameters of the shared gamma frailty reversed hazards model via EM algorithm was presented. The properties of the model were also discussed. The model was applied to monozygotic and dizygotic data given in Duffy et al. (1990). A correlated gamma frailty reversed hazards model was introduced in Chapter 4, to incorporate the situations where frailties of individuals differ. The estimation of the parameters of correlated gamma frailty reversed hazards model via EM algorithm was presented and applicability of the model was illustrated with real data sets.

Studies on dependence among random variables are increasingly important in survival analysis, actuarial science, reliability analysis and other areas related to probability and statistics. Various concepts of dependence, developed over recent decades, had allowed researchers in modelling multivariate random variables. In Chapter 5, we developed four new local dependence measures and studied their properties. These association measures were also discussed in terms frailty variables. Estimation of these association measures was presented and properties of those estimates were discussed. Finally, the usefulness of association measures was discussed using two real data sets.

6.2 Future Works

In the present work, the estimation of the parameters of the proportional reversed hazards models and association measures was done based on the assumption that the lifetime vector and censoring time vector are independent, to ensure the identifiability of marginal distribution functions. There are many situations, in which censoring time depends on lifetime. For example, Cui (1999) discussed Australian AIDS data diagnosed from 1984 and reported by the end of 1993. The time to AIDS diagnosis and the time to entry on to AIDS registry were the lifetimes under study. The reporting dates were recorded only from November 23, 1990. Before this date, only the fact that a case had been registered, but not the exact date of registration was known. Therefore time to entry on to AIDS registry

was left censored. Moreover, left censoring of the time to entry on to AIDS registry depends on the time to AIDS diagnosis. The analysis of left censored data, under dependent censoring is an area to be explored.

The estimation of parameters of the proportional reversed hazards frailty models, in Chapters 3 and 4, was done via EM algorithm. Other methods of estimation like penalized likelihood, pseudo likelihood, Bayesian technique etc can be used and comparison between the estimates can be carried out to determine the most suitable method.

The choice of frailty distribution is an important problem in modelling frailty random variables. In the present work, because of mathematical convenience, the gamma distribution with mean one is chosen as distribution of the frailty random variable. Other distributions such as positive stable, Weibull, lognormal etc can be considered as distribution of the frailty random variables. A study in this direction is an area to be explored.

In Chapters 3 and 4, we discussed proportional reversed hazards frailty models, which incorporate positive association among the individuals in the group. Negative association may exist between individuals in a group. For example, in the analysis of lifetime data on adopted children (Nielsen et al. (1992)) negative estimated dependence was found between adoptee and adoptive mother. Proportional reversed hazards frailty models to incorporate negative dependence is an area of research yet to be studied.

Based on the association measures described in Chapter 5, we can develop tests for independence among variables, which would be a topic of further research.

In the present work, we introduced new stochastic models for the analysis of left censored lifetime data. There are situations where the lifetime data is right truncated and left censored. For example, consider the bipolar affective disorder data given in Mclnnis et al. (1993). 125 families were ascertained via eighteen

hundred Probands screened for bipolar I or bipolar II. Age at onset and current ages of 34 parent –child pairs from 34 families among those 125 families were available. The remaining 91 families were excluded because they either showed clinical evidence of bilineality or did not have at least one interviewed, affected individual in each of two successive generations. Among the 34 parent-child pairs, there were 9 parent-child pairs with either missing age at onset or current age. The modelling and analysis of such data is also an area of interest.

References

1. Andersen, P. K., Borgan, Ø., Gill, R. D. and Keilding, N. (1993) Statistical Models Based on Counting Processes. Springer- Verlag, New York.

2. Anderson, J. E., Louis, T. A., Holm, N. V. and Harvald, B. (1992) *Time dependent association measures for bivariate survival distributions*, Journal of the American Statistical Association, **87**, 641-650.

3. Arnold, B. C. and Zahedi, H. (1988) *On multivariate mean residual life functions*, Journal of Multivariate Analysis, **25**, 1-9.

4. Asadi, M. (2006). *On the mean past lifetime of the components of a parallel system*, Journal of Statistical Planning and Inference, **136** (4), 1197-1206.

5. Bairamov, I., Kotz, S. and Kozubowski, J. J. (2003) *A new measure of linear local dependence*, Statistics, **37** (3), 243-258.

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

7. Barlow, R. E., Marshall, A. W., and Proshan, F. (1963) *Properties of probability distributions with monotone hazard rate*, The Annals of Mathematical Statistics, **34** (2), 375-389.

8. Bartoszewicz, J. and Skolimowska, M. (2006) *Preservation of classes of life distributions and stochastic orders under weighting*, Statistics and Probability Letters, **76** (6), 587-596.

9. Basu, A. P. (1971) *Bivariate failure rate*. Journal of the American Statistical Association, **66**, 103-104.

10. Basu, A. P. and Sun, K. (1997) *Multivariate exponential distributions with constant total failure rates*, Journal of Multivariate Analysis, **61**, 159-169.

11. Bismi, G. (2005) Bivariate Burr Distributions, PhD Thesis, Cochin University of Science and Technology, India.

12. Bjerve, S. and Doksum, K. (1992) *Correlation curves: Measures of association as functions of covariate values*, The Annals of Statistics, **21** (2), 890-902.

13. Block, H. W., Savits, T. H. and Singh, H. (1998) *The reversed hazards rate function*, Probability in the Engineering and Informational Sciences, **12**, 69-90.

14. Brindley, E. C. and Thompson, W. A. (1972) *Dependence and aging aspects of multivariate survival*, Journal of the American Statistical Association, **67**, 822-830.

15. Chandra, N. K. and Roy, D. (2005) *Classification of distributions based on reversed hazard rate*, Calcutta Statistical Association Bulletin, **56**, 221-224.

16. Chandra, N. K. and Roy, D. (2001) *Some results on reversed hazard rate*, Probability in the Engineering and Informational Science, **15** (1), 95-102.

17. Chen, Y. Q., Wang, M. C. and Huang, Y. (2004) *Semiparametric regression analysis on longitudinal pattern of recurrent gap times*, Biostatistics, **5** (2), 277-290.

18. Chu, H., Moulton, L. H., Mack, W. J., Passaro, D. J., Barroso, P. F. and Monoz, A. (2005) *Correlating two continuous variables subject to detection limits in the context of mixture distributions*, Journal of the Royal Statistical Society, Series C, **54** (5), 831-845.

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

20. Clayton, D. G. (1978) *A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence*, Biometrika, **65** (1), 141-151.

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

22. Crouchley, R. and Pickles A. R. (1995) *Multivariate survival models for repeated and correlated events*, Journal of Statistical Planning and Inference, **47** (1), 95-110.

23. Cui, J.(1999) *Nonparametric estimation of a delay distribution based on left censored and right truncated data*, Biometrics, **55** (2), 345-349.

24. Darbowska, D. M. (1988) *Kaplan-Meier estimate on the plane*, The Annals of Statistics, **16** (4), 1475-1489.

25. Di Crescenzo, A. (2000) *Some results on the proportional reversed hazards model*, Statistics and Probability Letters, **50** (4), 313-321.

26. Duffy, D. L., Martin, N. G. and Mathews, J. D. (1990) *Appendectomy in Australian twins*, American Journal of Human Genetics. **47** (3), 590-592.

27. Efron, B. and Tibshirani, R. J. (1993) An Introduction to the Bootstrap. Chapman and Hall, New York.

28. Fan, J. J., Hsu, I. and Prentice, R. L. (1998) *Dependence estimation over a finite bivariate failure time region*, Technical Report **336**. Division of Statistics, University of California at Davis, U.S.A.

29. Fan, J.J., Prentice, R. L. and Hsu, I. (2000). *A class of weighted dependence measures for bivariate failure time data*, Journal of the Royal Statistical Society, Series B, **62** (1), 181-190.

30. Finkelstein, M. S. (2002) *On the reversed hazard rate*, Reliability Engineering and System Safety, **78** (1), 71-75.

31. Finkelstein, M. S. (2003) *On one class of bivariate distributions*, Statistics and Probability Letters, **65** (1), 1-6.

32. Galambos, J. and Kotz, S. (1978) Characterization of Probability Distributions, Springer-Verlag, New York.

33. Gumbel, E. J. (1960) *Bivariate exponential distributions*, Journal of the American Statistical Association, **55**, 698-707.

34. Gumbel, E. J. (1961) *Bivariate logistic distributions*, Journal of the American Statistical Association, **56**, 335-349.

35. Gupta, R. C. (2003) *On some association measures in bivariate distributions and their relationships*, Journal of Statistical Planning and Inference, **117** (1), 83-98.

36. Gupta, R. C. and Gupta, R. D. (2007) *Proportional reversed hazard rate model and its applications.* Journal of Statistical Planning and Inference, **137** (11), 3525-3536.

37. Gupta, R. C. and Wu, H. (2001) *Analyzing survival data by proportional reversed hazard model*, International Journal of Reliability and Applications, **2** (1), 1-26.

38. Gupta, R. C., Gupta, R.D., and Gupta, P.L. (1998) *Modelling failure time data by Lehman alternatives*, Communication in Statistics- Theory and Methods, **27** (4), 887-904.

39. Gupta, R. D and Nanda, A. K. (2001) *Some results on reversed hazard rate ordering*, Communications in Statistics-Theory and Methods, **30** (11), 2447-2457.

40. Gürler, Ü. (1996) *Bivariate estimation with right-truncated data*, Journal of the American Statistical Association, **91**, 1152-1165.

41. Gürler, Ü. (1997) *On testing independence with right truncated data*, Nonlinear Analysis, Theory, Methods and Applications, **30** (5), 3201-3206.

42. Gürler, Ü. (2004) *Nonparametric bivariate estimation with randomly truncated observations*, Handbook of Statistics, **23**, 195-207.

43. Hoeffding, W. (1948) *A class of statistics with asymptotically normal distribution*, Annals of Mathematical Statistics, **19** (3), 273-325.

44. Holland, P. W. and Wang, Y. J. (1987) *Dependence function for continuous bivariate densities*, Communication in Statistics-Theory and Methods, **16**, 863-876.

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

46. Hougaard, P. (2000) Analysis of Multivariate Survival Data, Springer-Verlag, New York.

47. Iachine, I. A. (2001) The use of Twin and Family Survival Data in the Population Studies of Aging: Statistical Methods Based on Multivariate Survival Models, Ph.D Thesis, University of Southern Denmark.

48. Johnson, N. L. and Kotz, S. (1975) *A vector multivariate hazard rate*, Journal of Multivariate Analysis, **5**, 53-66.

49. Jones, G. and Rocke, D. M. (2002) *Multivariate survival analysis with doubly censored data: application to the assessment of Accutane treatment for fibrodysplasia ossificans progressiva*, Statistics in Medicine, **21**, 2547-2562.

50. Kalbfleisch, J. D. and Lawless, J. F. (1989) *Inference based on retrospective ascertainment: an analysis of the data based on transfusion-related AIDS*, Journal of the American Statistical Association, **84**, 360-372.

51. Kalbfleisch, J. D. and Prentice, R. L. (2002) The Statistical Analysis of Failure Time Data, John Wiley and Sons, New York.

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

53. Kayid, M. and Ahmad, I. A. (2004) *On the mean inactivity time ordering with reliability applications*, Probability in the Engineering and Informational Sciences, **18**, 395-409.

54. Keilson, J. and Sumita, U. (1982) *Uniform stochastic ordering and related inequalities*, The Canadian Journal of Statistics, **10** (3), 181-198.

55. Kendall, M. G. (1938) *A new measure of rank correlation*, Biometrika, **30** (1), 81-93.

56. Kendall, M. G. (1962) Rank Correlation Methods, 3rd Edition, Griffin, London.

57. Klein, J. P. (1992) *Semi parametric estimation of random effects using Cox model based on EM algorithm*, Biometrics, **48** (3), 795-806.

58. Klein, J. P. and Moeschberger, M. L. (2003) Survival Analysis: Techniques for Censored and Truncated Data, Springer-Verlag, New York.

59. Kulkarni, H. V. and Rattihalli, R. N. (2002) *Nonparametric estimation of a bivariate mean residual life function*, Journal of the American Statistical Association, **97**, 907-917.

60. Lawless, J. F. (2003) Statistical Models and Methods for Lifetime Data, Wiley, New York.

61. Lehmann, E. L. (1953) *The power of rank tests*, Annals of Mathematical Statistics, **24**, 23-43.

62. Lehmann, E. L. (1966) *Some concepts of dependence*, Annals of Mathematical Statistics, **37** (5), 1137-1153.

63. Li, X. and Lu, J. (2003) *Stochastic comparisons on residual life and inactivity time of series and parallel systems*, Probability in the Engineering and Informational Sciences, **17** (2), 267-275.

64. Marshall, A. W. (1975) *Some comments on the hazard gradient*, Stochastic Processes and their Applications, **3**, 293-300.

65. Martz, H. F. and Waller, R. A. (1982) Bayesian Reliability Analysis, Wiley, New York.

66. McKay, A. T. (1934) *Sampling from batches*, Journal of the Royal Statistical Society, Supplement, **1**, 207-216.

67. Mclnnis, M. G., McMahon, F. J., Chase, G. A., Simpson, S. G., Ross, C. A., and DePaulo, J. R. (1993) *Anticipation in bipolar affective disorder*, American Journal of Human Genetics, **53** (2), 385-390.

68. Morgenstern, D. (1956) *Einfache Beispiele Zweidimensionaler Verteilungen*, Mitteilingsblatt fur Mathematische Statistik, **8**, 234-235.

69. Murphy, S. A. (1994) *Consistency in a proportional hazards model incorporating a random effect*, The Annals of Statistics, **22** (2), 712-731.

70. Murphy, S. A. (1995) *Asymptotic theory for the frailty model*, The Annals of Statistics, **23** (1), 182-198.

71. Nair, N. U. and Asha, G. (2004) *Characterizations using failure and reversed failure rates*, Journal of the Indian Society for Probability and Statistics, **8**, 45-56.

72. Nair, N. U. and Sankaran, P. G. (2008) *A new measure of association for bivariate survival data*, (personal communication).

73. Nair, N. U., Asha, G. and Sankaran, P. G. (2004) *On covariate residual lives*, Statistica, **64** (3), 573-585.

74. Nair, N. U., Sankaran, P. G. and Asha, G. (2005) *Characterizations of distributions using reliability concepts*, Journal of Applied Statistical Science, **14** (3), 237-241.

75. Nanda, A. K., Bhattacharjee, S. and Alam, S. S. (2006) *On upshifted reversed mean residual life order*, Communications in Statistics- Theory and Methods, **35** (8), 1513-1523.

76. Nanda, A. K., Singh, H., Misra, N. and Paul, P. (2003) *Reliability properties of reversed residual lifetime*, Communications in Statistics-Theory and Methods, **32** (10), 2031-2042.

77. Nelson, W. B. (1969). *Hazard plotting methods for incomplete failure data*, Journal of Quality Technology, **1**, 27-52.

78. Nielsen, G. G., Gill, R. D., Andersen, P. K. and SØrensen, T. I. A. (1992) *A counting process approach to maximum likelihood estimation in frailty models*, Scandinavian Journal of Statistics, **19**, 25-44.

79. Oakes, D. (1982) *A model for association in bivariate survival data*, Journal of the Royal Statistical Society, Series B, **44** (3), 414-422.

80. Oakes, D. (1989) *Bivariate survival models induced by frailties*, Journal of the American Statistical Association, **84**, 487-493.

81. Parner, E. (1998) *Asymptotic theory for the correlated gamma-frailty model*, The Annals of Statistics, **26** (1), 183-214.

82. Roy, D. (2002) *A characterization of model approach for generating bivariate life distributions using reversed hazard rates*, Journal of the Japan Statistical Society, **32** (2), 239-245.

83. Sankaran, P. G and Gleeja, V. L. (2007b) *Proportional reversed hazard and frailty models*, Metrika, DOI 10.1007/s00184-007-0165-0.

84. Sankaran, P. G. and Gleeja, V. L (2006) *On bivariate reversed hazard rates*, Journal of the Japan Statistical Society, **36** (2), 213-224.

85. Sankaran, P. G. and Gleeja, V. L. (2007a) *Nonparametric estimation of reversed hazard rate*, Calcutta Statistical Association Bulletin, **59**, 55-68.

86. Sankaran, P. G. and Gleeja, V. L. (2008a) *Association measures for bivariate lifetime data*, Communications in Statistics- Theory and Methods (to appear).

87. Sankaran, P. G. and Gleeja, V. L. (2008b) *Proportional reversed hazards frailty models.* (communicated)

88. Sengupta, D., Singh, H. and Nanda, A. K. (1998) *The proportional reversed hazard model*, Technical report, Applied Statistics Division, Calcutta: Indian Statistical Institute.

89. Shaked, M. (1977) *A family of concepts of dependence for bivariate distributions*, Journal of the American Statistical Association, **72**, 642-650.

90. Shaked, M. and Shanthikumar, J. G. (1994) Stochastic Orders and their Applications, New York, Academic Press.

91. Shanbag, D. N. and Kotz, S. (1987) *Some new approaches to multivariate probability distributions*, Journal of Multivariate Analysis, **22** (2), 189-211.

92. Sinha S. K. (1986) Reliability and Life Testing, Wiley Eastern Limited, New Delhi.

93. van der Vaart, A. W. and Wellner, J. A. (1996) Weak Convergence and Empirical Processes with Application to Statistics, Springer-Verlag, New York.

94. Vaupel, J. W., Manton, K. G. and Stallard, E. (1979) *The impact of heterogeneity in individual frailty on the dynamics of mortality*, Demography **16** (3), 439-454.

95. Ware, J. H. and DeMets, D. L. (1976) *Reanalysis of some baboon descent data*, Biometrics, **32** (2), 459-463.

96. Wienke A. (2003) *Frailty Models*, Max-Plank Institute for Demographic Research. Working Paper 2003-032.

97. 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.

98. Yashin, A. I. and Iachine, I. A. (1995) *Genetic analysis of durations: Correlated frailty model applied to survival of Danish twins*, Genetic Epidemiology, **12** (5), 529-538.

99. Yashin, A. I. and Iachine, I. A. (1997) *How frailty models can be used for evaluating longevity limits: taking advantage of an interdisciplinary approach*, Demography, **34** (1), 31-48.

100. Yashin, A. I. and Iachine, I. A. (1999 a) *Dependent hazards in multivariate survival problems*, Journal of Multivariate Analysis, **71** (2), 241-261.

101. Yashin, A. I. and Iachine, I. A. (1999 b) *What difference does the dependence between durations make? Insights for population studies of aging*. Lifetime Data Analysis, **5** (1), 5-22.

102. Zdravkovic, S., Wienke, A., Pedersen, N. L., Marenberg, M. E., Yashin A. I. and de Faire, U. (2004) *Genetic influences on CHD-death and the impact of known risk factors: Comparison of two frailty models*, Behavior Genetics, **34** (6), 585-592.