2018

A dependent competing risks model

Yiqing Wang

ABSTRACT

A DEPENDENT COMPETING RISKS MODEL

Yiqing Wang, Ph.D.
Department of Mathematical Sciences
Northern Illinois University, 2018
Sanjib Basu and Nader Ebrahimi, Co-Directors

The competing risks model considers the setting where subjects or units are exposed to multiple risks, one of which may eventually cause the occurrence of the event, such as failure or recurrence or death. There is a substantial literature on identifiability and inference in both parametric and nonparametric models for competing risks. In this dissertation, we propose a parametric model for dependent competing risks that can be motivated by a frailty approach as well as by a copula approach. We establish identifiability conditions for this proposed model. We also consider competing risks regression framework and establish identifiability and methods for statistical inference in this framework. This proposed model has been further extended to analysis of semi-competing data while we again establish identifiability and statistical inference. The proposed models have been illustrated in extensive simulation studies and we apply these models to analyze competing risks data from a Tamoxifen trial on breast cancer patients and to analyze semi-competing risks data from a trial on tuberculous pericarditis collected in eight countries in Africa.

NORTHERN ILLINOIS UNIVERSITY
DE KALB, ILLINOIS

MAY 2018

A DEPENDENT COMPETING RISKS MODEL

BY

YIQING WANG

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICAL SCIENCES

Dissertation Co-Directors:
Sanjib Basu and Nader Ebrahimi

ACKNOWLEDGEMENTS

First and foremost, I would like to show my deepest gratitude to my advisor, Dr. Sanjib Basu, for his encouragement, inspiration and guidance on my dissertation study. I also appreciate my co-advisor, Dr. Nader Ebrahimi, for his suggestions and assistance on my study. At the meanwhile, I am indebted to Dr. Ananda Sen for his advice to my dissertation topic. I would also like to extend my sincerest thanks toward the rest of the committee members, including Dr. Alan M. Polansky, Dr. Duchwan Ryu and Dr. Jeffrey Thunder, for their agreement to serve as the committee members as well as their comments and suggestions. I would like to thank Dr. Desel Pena sincerely for his detailed review of my dissertation.

I would also like to thank the Division of Statistics and the Department of Mathematical Sciences for giving me the chance to study at NIU and accomplishing my dissertation study. And I owe my great thanks to my friends colleagues for their assistance throughout the whole dissertation process as well as the five years' life at DeKalb. I would say I cannot carry on finishing this task without their help and encouragement.

At the last and the most importantly, I would like to thank my parents for their supporting and encouraging.

DEDICATION

I dedicate this dissertation to my parents for their encouraging and supporting

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Competing risks data are usually encountered in biomedical studies. Such data arise when an event occurs because of several simultaneous possible causes (say K causes). The occurrence of an event due to one specific cause precludes one from observing the occurrence of events due to the other causes. Associated with each cause j, for $j = 1, ..., K$, there is a latent failure time T_j, which is the time to the occurrence of the event from cause j. In the competing risks setting, one observes only the minimum of the latent failure times T and a variable C indicating the cause of the event. That is, $T = \min(T_1, ..., T_K)$ and $C = \arg\min(T_1, \ldots, T_K)$. In survival analysis, T is the time to death and C is the cause of death. Such model has been discussed by David & Moeschberger (1978) and Cox & Oakes (1984). Competing risks models are popular not only in biomedicine but also in economics, engineering applications (reliability analysis) and actuarial sciences. In economics, T may be the duration of unemployment and C is the index of reason for leaving unemployment, such as getting a job or dropping out of the workforce. Flinn & Heckman (1982) applied a competing risks model to analyze such an unemployment problem. In engineering, for instance, T can be the lifetime of a computer and C is the cause of failure, such as failure of power source, CPU or display. Competing risks models in reliability have been studied by Schick & Wolverton (1978). A topic of substantial interest in competing risks literature is the identifiability issues in latent failure time approach related to identifying the sub-distribution of latent failure times T_j's from the distribution of the identified minimum.

Tsiatis (1975) and Cox (1959) showed that for any joint distribution of the failure times there exists a joint distribution with independent failure times which gives the same distribution of the identified minimum. Crowder (1991) showed that even when the marginal distribution of risks is

known, the joint distribution is not identifiable. This is not a desirable property of the competing risks model and Crowder (1991) called it the "identifiability crisis in competing risks analysis."

Several authors explored identifiability under model restrictions. A. P. Basu & Ghosh (1978) established identifiability of competing risks in the bivariate normal and the bivariate exponential distributions introduced in Marshall & Olkin (1967) and Gumbel (1960). Under weak assumptions of a known censoring time and a specified hazard ratio, Slud & Rubinstein (1983) constructed bounds on the marginal survival function. Heckman & Honoré (1989) established identifiability within a Cox proportional hazard model structure. Abbring & Van den Berg (2003) provided weaker conditions for a more restrictive model. Ebrahimi et al. (2003) provided identifiability conditions involving partial derivative of the conditional survival function of the event time given a restricted censoring time.

There is a substantial literature on modeling the dependent structure in competing risks using a copula. Sklar (1959) provided general results on flexibility of copula in modeling dependence. Zheng & Klein (1995) and Rivest & Wells (2001) proposed models using the Archimedean copula to model such dependent structure and the copula graphic estimator of marginal survival function when the dependence parameter in the copula model is known. Klein & Moeschberger (1988) also developed a model under the hypothesis of a known copula; more specifically, they applied a Clayton copula with known parameters.

In this dissertation, we propose a parametric modeling of the dependence between latent failure times $T_1, \ldots, T_K$. There are several advantages of the proposed model. First, it is flexible since the model can handle the dependent competing risks data and the independent competing risks data as well. Second, it can be represented via a frailty approach so that the conditional independent property can be applied. Third, it could be motivated via a copula approach as well so that the property of copula can be put into use. Finally, we establish conditions such that the model is identifiable.

The rest of this dissertation is organized as follows. In Chapter 2, we provide a literature review of competing risks models. We introduce our new dependent competing risks model including its representations under copula approach and frailty approach in Chapter 3. In Chapter 4, we discuss identifiability of the proposed dependent competing risks models and establish the conditions under which these models are identifiable. We show the performance of our model via simulation and tamoxifen data in Chapter 5 and extend the work to semi-competing risks field with associated simulation study and application on Mycobacterium data. In Chapter 7 we conclude with future research plan.

CHAPTER 2

COMPETING RISKS MODEL

2.1 Introduction

Competing risks is a generalization of standard survival analysis with multiple risks, any of which may cause the occurrence of the event. These risks "compete" with each other to become the cause of event. That is, the event of interest occurs from the risk which occurs first. One example of competing risks is when a patient suffers from various types of diseases simultaneously and dies from one of these diseases. There are several methods to analyze data involving competing risks; see for example David & Moeschberger (1978), Prentice et al. (1978), Crowder (2001), Pintilie (2006) and many others. The two primary approaches are:

1. Cause-Specific Hazard Approach

2. Latent Failure Time Approach

In general, cause-specific hazard approach models the joint distribution of failure time and the specific cause. There is an extension of cause-specific hazard approach which is induced by regressors and introduced by Cox (1972). The general latent failure time approach is to model the joint distribution of all potential failure times, including dependence modeling for competing risks. We will first introduce the formulations of cause-specific hazard and the latent failure time approaches, followed by competing risks via copula approaches. The dependent competing risks models via frailty approaches will be introduced in this chapter as well. Also, competing risks model with regressors will also be discussed.

2.2 Cause-Specific Hazard Approach

The cause-specific hazard approach has been explored by Altshuler (1970), Chiang (1970) and Prentice et al. (1978) and developed further by Prentice & Breslow (1978), Pepe & Mori (1993) and Kalbfleisch & Prentice (2011). It is a popular approach in biomedical applications.

Suppose failure time T is a non-negative continuous random variable and cause C takes values in the finite set $\{1,...,K\}$. It is assumed that a subject can fail from one and only one cause. The joint distribution of (T,C) is fully specified by either the cause-specific hazard function, $h(j,t)$, or by the cumulative incidence function, $F(j,t)$.

The *cause-specific hazard function* or *sub-hazard function* due to cause j is defined as

$$h(j,t) = \lim_{\triangle t \to 0} \frac{P(T \leq t + \triangle t, C = j | T > t)}{\triangle t}, \ for \ j = 1,...,K \tag{2.2.1}$$

which is a conditional probability that the event occurs from cause j at time t, given that the event of interest has not occurred up to time t. Some people also refer to this as the marginal hazard function. The function $h(j,t)$ represents the instantaneous failure rate from cause j at time t in the presence of other competing risks.

The survival function $S(j,t)$ of cause j, also referred to as *sub-survival function*, is defined as

$$S(j,t) = P(T > t, C = j), \ for \ j = 1,...,K \tag{2.2.2}$$

In addition, the *sub-density function* corresponding to cause j is given by

$$f(j,t) = -\frac{d}{dt} S(j,t), \ for \ j = 1,...,K \tag{2.2.3}$$

and the *cumulative incidence function* $F(j,t)$ from cause j is defined as

$$F(j,t) = P(T \leq t, C = j) = \int_0^t f(j,u)du, \ for \ j = 1,...K \tag{2.2.4}$$

which is the probability that event will occur from cause j within time t.

The *overall hazard function* $h(t)$ and the *overall survival function* $S(t)$ can be presented, respectively, in terms of the cause-specific functions as

$$h(t) = \lim_{\triangle t \longrightarrow 0} \frac{P(t < T \leq t + \triangle t | T > t)}{\triangle t} = \sum_{j=1}^{K} h(j,t) \tag{2.2.5}$$

$$S(t) = P(T > t) = \sum_{j=1}^{K} P(T > t, C = j) = \sum_{j=1}^{K} S(j,t) \tag{2.2.6}$$

The fact that the event of interest has not occurred up to time t indicates that the subject has survival from all K risks up to time t. Hence, there is a relationship among the cause-specific hazard function, sub-density function and overall survival function, which is

$$h(j,t) = \frac{f(j,t)}{S(t)} \tag{2.2.7}$$

The *overall cumulative hazard function* $H(t)$ from all K risks combined is expressed in terms of the sum of all the *cause-specific cumulative hazard functions*:

$$H(t) = \sum_{j=1}^{K} H(j,t) \tag{2.2.8}$$

where $H(j,t) = \int_0^t h(j,u)du$ is the cause-specific cumulative hazard function.

Therefore, applying the commonly used survival analysis relations, the overall survival function for the competing risk model (2.2.6) can be written as

$$S(t) = \exp[-H(t)] = \exp[-\sum_{j=1}^{K} H(j,t)] \tag{2.2.9}$$

Suppose $\{(t_i, C_i, d_i),\ for\ i = 1, ..., n\}$ are observations from n subjects, where t_i is the observed event time of i^{th} subject, C_i is the cause of the event of subject i, and d_i is a censoring indicator which takes the value one if the event occurs to i^{th} subject and zero otherwise. We only consider right censoring, which implies the true failure time is greater than or equal to the observed failure time.

The likelihood function for the competing risk model is written in terms of the sub-density function and the overall survival function as

$$L = \prod_{i \in E_1} f(C_i, t_i) \prod_{i \in E_2} S(t_i) \tag{2.2.10}$$

where E_1 is the set of subjects who has experienced the event and E_2 is the set of subjects who has not experienced the event up to the follow-up time t_i.

Further, using (2.2.7), (2.2.8) and (2.2.9), we can also write the likelihood (2.2.10) as

$$\begin{aligned} L &= \prod_{i\in E_1} h(C_i,t_i)S(t_i)\cdot \prod_{i\in E_2} S(t_i) \\ &= \prod_{i\in E_1} h(C_i,t_i)\exp[-H(t_i)]\cdot \prod_{i\in E_2} \exp[-H(t_i)] \\ &= \prod_{i\in E_1} h(C_i,t_i)\cdot \prod_{E_1\bigcup E_2} \exp[-H(t_i)] \\ &= \prod_{i=1}^{n}[h(C_i,t_i)]^{d_i}\cdot \prod_{i=1}^{n}\exp[-H(t_i)] \\ &= \prod_{i=1}^{n}[h(C_i,t_i)]^{d_i}\cdot \exp[-\sum_{j=1}^{K}\int_0^{t_i} h(j,u)du] \\ &= \prod_{i=1}^{n}\left([h(C_i,t_i)]^{d_i}\cdot \prod_{j=1}^{K}\exp[-\int_0^{t_i} h(j,u)du]\right) \end{aligned} \tag{2.2.11}$$

It can be seen that the likelihood function (2.2.11) is completely expressed in terms of cause-specific function $h(j,t),\ j=1,...,K$. Furthermore, the likelihood function can be factored into K separate components as

$$L=\prod_{j=1}^{K}\prod_{i=1}^{n}[h(j,t_i)]^{d_{ij}}\cdot\exp[-\int_0^{t_i} h(j,u)du] \tag{2.2.12}$$

where $d_{ij}=I(C_i=j),\ j=1,\ldots,K$. Note that each likelihood component is specifically the same as the likelihood function that would be obtained by censoring events from all other causes. Hence, the likelihood factorization along classical survival data approaches make it clear that each $h(j,t)$ is identifiable. The cause-specific hazards approach was proposed by Prentice et al. (1978) and Kalbfleisch & Prentice (1980). Statistical inference under the cause-specific model is discussed by Crowder (2001).

The general cause-specific approach allows for dependence among the risks within a general dependence structure but may not provide specific modeling of dependence among the risks.

2.3 Latent Failure Time Approach

The latent failure time approach models the potential failure time of different risks or possible causes, it has been discussed by Cox (1959), Moeschberger & David (1971) and Gail (1975). This model is popular in engineering applications for the study of reliability of systems.

Suppose there exist K competing risks in a study and potential failure times denoted by T_j (where $j = 1, ..., K$) corresponding to each cause of event. The model is often formulated in terms of these latent failure times $T_1, T_2, ..., T_K$. Each $T_j, j = 1, ..., K$ is a non-negative random variable. It is also assumed that the latent failure time $T_j, j = 1, ..., K$ is continuous and no ties occur, that is, $P(T_j = T_{j^*}) = 0$ for $j = j^*$. The *marginal survival* is

$$S_j(t) = P(T_j > t), j = 1, ..., K \tag{2.3.1}$$

and *marginal hazard function* is

$$h_j(t) = -\frac{d}{dt} \log S_j(t), j = 1, ..., K \tag{2.3.2}$$

The marginal survival functions are known as the net survival function in demographics and actuarial terminology. The joint survival function of $T_1, T_2, ..., T_K$ is defined as

$$\begin{aligned} S(t_1, ...t_K) &= P(T_1 > t_1, ..., T_K > t_K) \\ &= \int_{t_1}^{\infty} \cdots \int_{t_K}^{\infty} g(s_1, ..., s_K) ds_1 \ldots ds_K \end{aligned} \tag{2.3.3}$$

where $g(\cdot)$ is the joint density function of $T_1, T_2, ..., T_K$. The actual observed failure time T and cause C in the competing risks setting are defined as $T = \min(T_1, T_2, ..., T_K)$ and $C = \arg\min(T_1, \ldots, T_K)$.

Historically, the concept of a latent failure time was developed in Cox (1959) and Moeschberger & David (1971) in which the failure time T_j of cause j would be observed only if all failure of causes other than j are removed from the experimental unit or subject. This is why these are termed as latent or potential failure times. In reliability studies, this assumption is often tenable as the competing risks are from individual components of a system. The absence or existence of one risk may lead to different effects of other risks which has been mentioned by Cornfield (1957). Gail (1975) introduced a model which assumes the presence of $T_1, T_2, ..., T_K$ on each experimental unit or subject under certain study conditions. The joint survival function $S(\cdot)$ will demonstrate the interrelationship among the risks. The relationship between risk j and risk k could be either dependent or independent and is expressed in terms of the stochastic dependence or independence of T_j and T_k. The survival function of T_j is defined to be the marginal survival function as $S_j(t)$ in 2.3.1.

If all risks are independent of one other, the relationship between joint survival function and marginal survival function can be exhibited as following:

$$S(t_1, ..., t_K) = \prod_{j=1}^{K} S_j(t_j) \tag{2.3.4}$$

and the survival function of T is given by

$$S(t) = S(t, ..., t) = \prod_{j=1}^{K} S_j(t) \tag{2.3.5}$$

David & Moeschberger (1978) presented forms of the likelihood functions when all failure times are completely observed and censoring failure times are grouped. The likelihood function and its derivatives for independent Weibull and exponential latent failure times have also been explored in their studies.

It is often difficult to obtain accurate information about the marginal distributions of individual event times or the dependence structure among them given only the observed data (T,C). Both Tsiatis (1975) and Crowder (1991) have established that one can always construct a joint survival function with independent failure time distributions such that it has the same cause-specific hazard function as a joint survival function with an arbitrary dependence structure among the causes of the event. Therefore, neither the marginal survival function $S_j,\ j = 1,...,K$ nor the joint survival function S is identifiable from observed competing risks data. Further, the independence assumption of the marginal failure times cannot be tested in this general setting.

Frailty models form an important branch of survival analysis; see the book by Duchateau & Janssen (2007). Zheng & Klein (1995) have shown that if the copula function is known, the marginal distributions of failure times can be identified based on competing risks observations of time and cause. Under the case of discrete failure times, competing risks with regressor models and parametric models can make the identifiability possible. Crowder (2001) has a detailed survey on this topic.

2.4 Dependent Competing Risks Approach

The cause-specific approach allows an implicit model of dependence among the competing risks. Other approaches were also developed to provide explicit dependence model among the latent failure times of the competing risks, including:

1. Dependent Competing Risks Model Using Copula

2. Frailty Models for Dependent Competing Risks

2.4.1 A Dependent Competing Risks Model Using Copula

Copulas have been introduced in Sklar (1959) and discussed at greater length in Sklar (1973) and Schweizer & Sklar (1974). Copulas have been defined by Schweizer & Sklar (2011) as well, which are used to characterize the dependence between random variables, so that a multivariate joint distribution function is linked to its univariate margins. Various copula functions and their properties have been discussed by Joe (1997) and Nelsen (1999). Specifically, let G be a joint (two-dimensional) distribution function of two random variables X_1, X_2 and let $F_1,\ F_2$ be their marginal cumulative distribution functions, respectively. Then there is a continuous function C from the unit square $[0,1]\times[0,1]$ onto the unit interval $[0,1]$ such that for all x_1, x_2,

$$G(x_1,x_2) = P(X_1 \leq x_1, X_2 \leq x_2) = C[F_1(x_1), F_2(x_2)] \tag{2.4.1.1}$$

Such a function C is called a *copula*. More generally, in probability theory and statistics, a function $C(x_1,...,x_m)$ is called a copula function if it is a multivariate probability distribution for which the marginal probability distribution of each variable is uniformly distributed on $[0,1]$. That is, $C(x) = P(X_1 \leq x_1,...,X_m \leq x_m)$, where $x = (x_1,...,x_m)$ and each random variable X_i (where $i = 1,...,m$) is uniformly distributed on $[0,1]$.

Sklar (1959) established that any joint distribution function G with marginal distribution function $F_1,...,F_m$ can always be described as

$$G(\mathbf{x}) = C[F_1(x_1),...,F_m(x_m)] \tag{2.4.1.2}$$

where $\mathbf{x} = (x_1,...,x_m)$. The copula C is unique when the marginal distribution functions $F_1,...F_m$ are strictly monotonically increasing. Genest & Nešlehová (2007) mentioned that when one or more marginal distribution functions are discrete, the uniqueness is no longer true. The copula model

2.4.1.2, nevertheless, provides a model for defining a multivariate distribution function G using an admissible copula C. If the copula C can be presented as a product, $G(\mathbf{x}) = \prod[F_1(x_1), ..., F_m(x_m)]$, then the random variables $X_i,\ for\ i = 1, ..., m$ are no longer dependent.

Theorem 2.4.1. *Suppose $X_1, ..., X_m$ are continuous random variables, then $X_i,\ for\ i = 1, ..., m$ are independent if and only if the copula function C in 2.4.1.2 is a multiplication, that is,*

$$G(\mathbf{x}) = C[F_1(x_1), ..., F_m(x_m)] = \prod_{i=1}^{m} F_i(x_i)$$

In many applications, the random variables of interest represent the lifetime of individuals or objects in some population. The probability of an individual living or surviving beyond time x is given by the survival function $S(x) = P(X \leq x) = 1 - F(x)$. Nelsen (1999) has mentioned the survival copulas in his research. It is a copula which links the joint survival function to its univariate margins in a manner completely analogous to the way in which a traditional copula connects the joint distribution function to its margins. That is, the joint survival function can be written in terms of its margins associated with a survival copula:

$$S(\mathbf{x}) = S(x_1, ..., x_m) = C[S_1(x_1), ..., S_m(x_m)] \tag{2.4.1.3}$$

There is a wide range of families of copulas. Archimedean copulas are popular because they allow modeling dependence in arbitrarily high dimensions via a univariate function (Table 2.1). A copula C is called an Archimedean copula if it admits the following expression:

$$C(x_1, ..., x_m) = \psi[\psi^{-1}(x_1) + ... + \psi^{-1}(x_m)] \tag{2.4.1.4}$$

where ψ is a generator function and ψ^{-1} is its inverse function.

Table 2.1: Examples of Archimedean Copulas

name	generator $\psi(t)$	generator inverse$\psi^{-1}(t)$	parameter
Clayton	$(1+\theta t)^{-1/\theta}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$\theta \in [0,\infty)$
Gumbel	$\exp(-t^{\delta})$	$(-\log t)^{\frac{1}{\delta}}$	$\delta \in (0,1]$

Moreover, ψ is a continuous, strictly decreasing and convex function such that $\psi(1) = 0$. More generally, the C in 2.4.1.4 yields a copula if and only if ψ is d-monotone on $[0,\infty)$. That is, it is $d-2$ times differentiable and the derivatives satisfy

$$(-1)^k \psi^{(k)}(x) \geq 0$$

for all $x \geq 0,\ k = 0,1,...,d-2$, and $(-1)^{d-2}\psi^{(d-2)}(x)$ is non-increasing and convex.

In this section, we will first consider a bivariate dependent competing risks model under copula and then extend the model to multivariate case. In the bivariate case, we assume there are only two risks that may lead to the occurrence of the event: the latent failure time T_1 is the event time to risk 1, and T_2 is the event time to risk 2. Note that T_1 and T_2 are not independent. Klein & Moeschberger (1988) and Zheng & Klein (1995) proposed a bivariate competing risks model based on copulas that applies to a large class of copulas, and its performance has been checked for some popular copulas, such as the Gumbel copula (Gumbel (1960)) and Clayton model (Clayton (1978)). The model under Gumbel copula will be mainly discussed next.

Define $S_j(t)$ (where $j = 1,2$) as the marginal survival function of latent failure time T_j (where $j = 1,2$). The corresponding joint survival function $S(t_1,t_2)$ via the Gumbel copula is written as

$$\begin{aligned} S(t_1,t_2) &= P(T_1 > t_1, T_2 > t_2) \\ &= \psi[\psi^{-1}(S_1(t_1)) + \psi^{-1}(S_2(t_2))] \\ &= \exp\left(-\left[(-\log[S_1(t_1)])^{\frac{1}{\delta}} + (-\log[S_2(t_2)])^{\frac{1}{\delta}}\right]^{\delta}\right) \end{aligned} \tag{2.4.1.5}$$

where $\delta \in (0,1]$ represents the positive association coefficient between T_1 and T_2. The parameter δ is the degree of association and provides a one-to-one correspondence to Kendalls τ. The independence copula, corresponding to independent competing risks, is given by $C(u,v) = uv$. Thus, the case $\delta = 1$ yields independent competing risks.

In competing risks framework, one only observes the minimum between T_1 and T_2, namely $T = \min(T_1,T_2)$ and cause $C = \arg\min(T_1,T_2)$. Thus, the joint survival function based on the observed data can be written as

$$S(t,t) = P(T \geq t) = P(T_1 \geq t, T_2 \geq t) \tag{2.4.1.6}$$

which, for the bivariate joint survival function via Gumbel copula, is given by

$$S(t,t) = \exp\left(-\left[(-\log[S_1(t)])^{\frac{1}{\delta}} + (-\log[S_2(t)])^{\frac{1}{\delta}}\right]^{\delta}\right) \tag{2.4.1.7}$$

In order to extend the model to multivariate case, the observed data would be $T = \min(T_1, ..., T_K)$ and $C = \arg\min(T_1, \ldots, T_K)$; the corresponding multivariate joint survival function via Gumbel copula is given by

$$\begin{aligned} S(t,...,t) &= \psi[\psi^{-1}(S_1(t)) + ... + \psi^{-1}(S_K(t))] \\ &= \exp\left(- \big[\sum_{j=1}^{K} (-\log[S_j(t)])^{\frac{1}{\delta}} \big]^{\delta} \right) \end{aligned} \tag{2.4.1.8}$$

2.4.2 Frailty Models for Dependent Competing Risks

Another class of models, known as *frailty models*, can be used to model dependent competing risks as well. Frailty models are based on hazard functions and have a multiplicative frailty factor. Assume the hazard functions of K risks share a common random effect or *frailty factor* Z. The frailty factor is random and therefore a frailty distribution needs to be specified in the frailty model. For a subject, the frailty factor Z changes over the population of subjects with a distribution function D on $(0, \infty)$. Assume for a subject that the latent failure time T_j (where $j = 1, .., K$) is conditionally independent, given a positive random variable Z. That is, $(T_1, \ldots, T_K) \,|\, Z$ are conditionally independent and the corresponding conditional marginal survival function of T_j, given $Z = z$, is

$$S_j(t_j|z) = P(T_j > t_j | Z = z) = [S_{0_j}(t_j)]^z \tag{2.4.2.1}$$

where $S_{0_j}(t_j)$ is a continuous baseline survival function. Then the marginal cumulative hazard function of $T_j|Z = z$, for $j = 1, ..., K$, equivalently, can be written as

$$H_j(t_j|z) = H_{0_j}(t_j) \cdot z \tag{2.4.2.2}$$

and the joint survival function conditional on $Z = z$ can be written as

$$\begin{aligned} S(t_1,...,t_K|z) &= \prod_{j=1}^{K} S_j(t_j|z) && (2.4.2.3) \\ &= \prod_{j=1}^{K} \exp[-H_{0_j}(t_j)\cdot z] \\ &= \exp\left[-z\cdot\sum_{j=1}^{K} H_{0_j}(t_j)\right] && (2.4.2.4) \end{aligned}$$

The unconditional joint survival function can be obtained by integrating over the distribution of frailty factor Z, which is

$$\begin{aligned} S(t_1,...,t_K) &= \int_0^{\infty} \prod_{j=1}^{K} [S_{0_j}(t_j)]^z dD(z) \\ &= \int_0^{\infty} \exp\left(-\left\{\sum_{j=1}^{K} -\log S_{j_0}(t_j)\right\} z\right) dD(z) \\ &= E\left[\exp\left(-\left\{\sum_{j=1}^{K} -\log S_{j_0}(t_j)\right\} z\right)\right] && (2.4.2.5) \end{aligned}$$

which can also be regarded as the moment-generating function $M_Z(s)$ of the variable Z evaluated at $s = \left\{\sum_{j=1}^{K} \log S_{0_j}(t_j)\right\}$. And the joint survival function 2.4.2.5 can also be expressed in terms of cumulative hazard function as

$$S(t_1,...,t_K) = \int_0^{\infty} \exp\left[-z\cdot\sum_{j=1}^{K} H_{0_j}(t_j)\right] dD(z) \qquad (2.4.2.6)$$

which can also be regarded as the moment-generating function $M_Z(s)$ of the variable Z evaluated at $s = -\sum_{j=1}^{K} H_{0_j}(t_j)$.

Under the construction of the frailty model, a multivariate distribution of dependent lifetimes of the different components, or in other words the dependent competing risks, can be obtained. Oakes

(1989) proposed a bivariate survival model induced by frailties. See also the book by Duchateau & Janssen (2007) that discusses gamma, positive stable and other frailty models.

2.4.2.1 Model with Gamma Frailty

If the frailty Z follows a Gamma distribution with scale parameter α and shape parameter β, where $\alpha > 0,\ \beta > 0$, then the unconditional joint distribution function is given as the moment-generating function $M_Z(s)$ of the variable Z evaluated at $s = \left\{ \sum_{j=1}^{K} \log S_{0_j}(t_j) \right\}$, which is

$$S(t_1,...,t_K) = \left[1 + \alpha \left\{ \sum_{j=1}^{K} -\log S_{0_j}(t_j) \right\} \right]^{-\beta} \tag{2.4.2.1.1}$$

for $\left\{ \sum_{j=1}^{K} \log S_{0_j}(t_j) \right\} < \frac{1}{\alpha}$.

2.4.2.2 Model with Positive Stable Frailty

If the frailty factor Z follows a positive stable distribution indexed by the parameter δ with the Laplace transform given by $L(s) = E[\exp(-sZ)] = \exp(-s^{\delta})$, where $\delta \in (0,1]$, then the unconditional joint survival function can be expressed as the moment-generating function $M_Z(s)$ of the variable Z evaluated at $s = \sum_{j=1}^{K} \log S_{0_j}(t_j)$, which is

$$S(t_1,...,t_K) = \exp\left[-\left\{ \sum_{j=1}^{K} \log S_{0_j}(t_j) \right\}^{\delta} \right] \tag{2.4.2.2.1}$$

More details regarding the properties of the stable distribution can be found in the book by Duchateau & Janssen (2007).

2.5 Competing Risks Regression Approach

The widely popular proportional hazards (PH) regression model for lifetime data was proposed by Cox (1972). It can be utilized to model effects of regression variables on cause-specific hazard functions, see in Prentice & Breslow (1978) and Holt (1978).

Suppose the observed data for the i^{th} subject is $(t_i, C_i, \delta_i, \mathbf{x_i})$, where $i = 1, ..., n$; t_i denotes the observed failure time to event for subject i and C_i is the corresponding cause of the event. The indicator variable δ_i with value 1 implies that the observed failure time is uncensored and value 0 otherwise. The possible covariate $\mathbf{x_i}$ is a set of explanatory variables for subject i, which represents a collection of predictor variables that is being modeled to predict a subject's hazard. The proportional hazards regression model for cause-specific hazard can be written as

$$h(j,t|\mathbf{x}) = h_0(j,t) \cdot \exp(\mathbf{x}'\boldsymbol{\eta}_\mathbf{j}),\ for\ j = 1, ..., K \tag{2.5.1}$$

where $h_0(j,t)$ is the baseline hazard function of risk j only involving t and η_j is the cause-specific regression coefficient to be estimated from the data. In contrast, the exponential expression involves only $\mathbf{x}$, but not t. That is, the covariates $x's$ are time-independent variables. The observed likelihood function 2.2.12,

$$L = \prod_{i=1}^{n} \prod_{j=1}^{K} \left([h(C_i, t_i|\mathbf{x_i})]^{d_{ij}} \cdot \exp\big[- \int_0^{t_i} h(j,u|\mathbf{x_i}) du \big] \right)$$

for Cox PH model can be written as

$$\begin{aligned} L &= \prod_{i=1}^{n} \prod_{j=1}^{K} \left([h_0(C_i,t_i) \cdot \exp(\mathbf{x_i'}\boldsymbol{\eta}_\mathbf{j})]^{d_{ij}} S(j,t_i|\mathbf{x_i}) \right) \\ &= \prod_{i=1}^{n} \prod_{j=1}^{K} \left([h_0(C_i,t_i) \cdot \exp(\mathbf{x_i'}\boldsymbol{\eta}_\mathbf{j})]^{d_{ij}} \cdot \exp\big[- \int_0^{t_i} h_0(C_i,u) \cdot \exp(\mathbf{x_i'}\boldsymbol{\eta}_\mathbf{j}) du \big] \right) \end{aligned} \tag{2.5.2}$$

where $d_{ij} = 1$ if $C_i = j$ and $\delta_{ij} = 0$ otherwise. Note that no assumption is required on the inter-relation among the causes of failure, that is, competing risks. The usual semiparametric approach to this model leaves the baseline hazard functions as unspecified and proceeds with the analysis based on the "partial" likelihood, which has been introduced by Cox (1975), for inference about η_j, which is derived using ordered failure times as

$$L^* = \prod_{i=1}^{n} \prod_{j=1}^{K} \frac{\exp(\mathbf{x}'_{\mathbf{i(j)}} \eta_{\mathbf{j}})}{\sum_{l \in R_j} \exp(\mathbf{x}'_{\mathbf{i(l)}} \eta_{\mathbf{j}})} \tag{2.5.3}$$

where R_j denotes the risk set at time t_j, that is, the set of subjects who have not failed or been censored by that time. $x_{i(j)}$ denotes the covariates' value for the subject i failing at time $t_{(j)}$.

Nonparametric estimation of cause-specific baseline hazard function $h_0(j,t)$ may be obtained by the approaches of Cox (1972), Kalbfleisch & Prentice (1973) and Breslow (1974). In particular, Aalen (1976) has mentioned that the Kaplan-Meier estimator may be utilized to estimate the cause-specific hazard function $h_(j,t)$ if there are no regressor variables present.

Cox PH model can be used to incorporate time-dependent covariates, but such a model no longer satisfies the PH assumption. Alternatively, the accelerated failure (AFT) regression model can address such a case. More general models are proposed in Wei (1992) and allow the predictors to change over time.

2.6 Identifiability and Non-Identifiability Issues in Competing Risks

It is known that if the latent failure times $T_1, ..., T_K$ are independent, then the cause-specific hazard approach and latent failure time approach lead to identical statistical inference, see in Crowder (2001) and Kalbfleisch & Prentice (2011). For any competing risks model with a joint survival function $S(t_1, ..., t_K)$ where $T_1, ..., T_K$ may be dependent, there exists a different joint survival func-

tion in which $T_1,...,T_K$ are independent such that the same set of sub-survival functions $S(j,t)$ or crude survival functions come into being from these two models. It is a well-known problem of identifiability which has been discussed by Cox (1959), Tsiatis (1975), Crowder (2001) and many others. This problem makes it impossible to differentiate the dependent model and independent model only based on the observed data (T,C).

Specifically, Tsiatis (1975) showed that given any joint survival function with arbitrary dependent competing risks, there exists a different joint survival function in which the competing risks are independent and which leads to exactly the same sub-densities $f_j(t)$ as the former. It brings in the following theorem.

Theorem 2.6.1 (Tsiatis, 1975)**.** *Suppose $T_1,..,T_K$ are dependent potential latent failure times with joint survival function $S(t_1,...,t_K)$; $f_j(t)$ is the corresponding sub-density function of risk j and can be computed by 2.7.1. Then there exists a joint survival function $S^*(t_1,...,t_K)=\prod_{j=1}^{K} S_j^*(t_j)$, where $S_j^*(t)=\exp(-\int_0^t h_j(s)ds)$ and the sub-hazard function $h_j(t)$ derives from the given $f_j(t)$.*

Heckman & Honoré (1989) showed that the identification of the joint survival function $S(t_1,...,t_K)$ may be possible from the cause-specific survival functions when there are covariates in the model within a certain framework.

Theorem 2.6.2 (Heckman & Honoré, 1989)**.** *Suppose the joint survival function of latent failure times T_1,T_2 is given by*

$$S(t_1,t_2)=\mathbf{K}[S_1(t_1|\mathbf{x}),S_2(t_2|\mathbf{x})]$$

where $S(j,t_j|\mathbf{x})=\exp[-g_j(\mathbf{x})H_j(t_j)]$ for $j=1,2$ expressed in terms of the integrated hazard function $H_j(t_j)$ and $g_j(\mathbf{x})$, which is usually specified as $\exp(\mathbf{x}\eta_\mathbf{j})$ (where $\eta_\mathbf{j}$ is the correlation coefficient). $\mathbf{K}$ *is a distribution function. Then under the assumptions:*

1. $\mathbf{K}$ *is continuously differentiable and strictly increasing in* $[0,1]\times[0,1]$
2. *the support of* $g_1(\mathbf{x}),g_2(\mathbf{x})$ *is* $(0,\infty)\times(0,\infty)$

3. $H_1(1) = 1$, $H_2(1) = 1$, $g_1(x_0) = 1$ *and* $g_2(x_0) = 1$ *for some fixed point* x_0 *in the support of* $\mathbf{X}$

4. $H_j(t) \geq 0$ *and* $H_j^{'}(t) = dH_j(t)/dt > 0$ *for all* t,

the joint survival function $S(t_1,t_2)$ *is identified from the identified minimum of* (T_1,T_2) *and the cause. In other words,* H_1, H_2, g_1, g_2 *and* $\mathbf{K}$ *are identified from the observed data* (T,C).

By theorem 2.6.2, the joint survival function $S(t_1,t_2)$ can be determined by the set of sub-survival functions $S_j(t_j)$.

The identifiability can also be regained under specific parametric forms, which has been mentioned by A. P. Basu & Ghosh (1978) and A. P. Basu & Klein (1982). Klein & Moeschberger (2005) furnished a summary of analytic approaches for dealing with competing risks data when the assumption of independence may not be satisfied.

2.7 Remarks

The latent failure time and the cause-specific approaches are related. For instance, the sub-density function of risk j can be obtained from the joint survival function of the latent failure times (Tsiatis, 1975) as

$$f(j,t) = -\frac{\partial S(\mathbf{t})}{\partial t_j}\bigg|_{\mathbf{t}=(t,..,t)} = -\frac{\partial S(t_1,...,t_K)}{\partial t_j}\bigg|_{(t_1,...,t_K)=(t,..,t)} \tag{2.7.1}$$

The sub-density function of risk j can also be computed by integrating the joint density function $g(\cdot)$ of latent failure times $T_1,...,T_K$ over the interval (t,∞) with respect to all but the j^{th} variable:

$$f(j,t) = \int_t^\infty ... \int_t^\infty \int_t^\infty ... \int_t^\infty g(s_1,...,s_K) ds_1 ... ds_{j-1} ds_{j+1} ... ds_K \tag{2.7.2}$$

CHAPTER 3

A DEPENDENT COMPETING RISKS MODEL

3.1 Introduction

In this chapter, we provide the formulation of the proposed dependent competing risks model. In fact, the independent competing risks problem can be handled by this model as well. In competing risks studies, one popular approach is to model the distribution of the observed minimum T and cause C which have been introduced in Prentice et al. (1978) and Prentice & Breslow (1978). An alternative approach is to model the joint distribution of latent failure times, seeing Cox (1959) and Moeschberger & David (1971), which is the way to achieve the new dependent competing risks model. The new model could be motivated via a frailty approach as well as a copula approach. We will first introduce the formulation of the new model and then present the different representations of the model.

Our proposed model is a parametric model using the Weibull distribution for competing risks. It is well known that the Weibull distribution can be used to model a hazard function that is increasing, decreasing or even constant. This makes it popular in studies of parametric models for competing risks. First, we will present the general model with K competing risks. In the latter part, we will present the special case of two competing risks.

3.2 Model Formulation

The traditional approach, via latent failure times, assumes that there is a potential failure time associated with each of the K competing risks to which the system is exposed. More generally, T_j for $j = 1,...,K$ represents the failure time of each risk j. The smallest T_j decides the overall system failure time T and can be observed. Another observed variable is the indicator variable of censoring C_j taking the value 0 if censored and 1 if event occurs from risk j. Thus, $T = \min(T_1,...,T_K) = T_{C_j}$. The vector $\mathbf{T} = (T_1,...,T_K)$ is associated with the joint survival function $S(\mathbf{t}) = S(t_1,...,t_K) = P(T_1 > t_1,...,T_K > t_K)$ and the marginal survival function $S_j(t_j) = P(T_j > t_j)$. In the field of actuarial science and demography, the marginal $S_j(t_j)$ is also called the net survival function. If the joint survival function $S(\mathbf{t})$ of $\mathbf{T}$ is known, then the marginal of each T_j for $j = 1,...,K$ can be derived by the theorem due to Tsiatis (1975) from the book by Crowder (2001).

Theorem 3.2.1 (Tsiatis, 1975)**.** *The cause-specific sub-densities can be calculated directly from the joint survival function of latent failure times as*

$$f(j,t) = -\frac{\partial S(t_1,...,t_K)}{\partial t_j}\bigg|_{t_j=t}, \text{ for } j = 1,...,K \tag{3.2.1}$$

Therefore, the sub-hazards

$$h(j,t) = \frac{f(j,t)}{S(t_1,...,t_K)} = \frac{\partial S(t_1,...,t_K)/\partial t_j}{S(t_1,...,t_K)}\bigg|_{t_j=t} \tag{3.2.2}$$

We propose a parametric model where the latent failure times are assumed to follow Weibull distributions so that the marginal survival function of T_j for $j = 1,...,K$ is given by

$$S_j(t_j) = P(T_j > t_j) = \exp(-\mu_j t_j^{\beta_j}) \tag{3.2.3}$$

where $\mu_j \in (0,\infty)$ is the scale parameter of the Weibull distribution and $\beta_j \in (0,\infty)$ is the shape parameter. The hazard function is increasing for $\beta_j > 1$, decreasing for $\beta_j < 1$ and constant if $\beta_j = 1$.

Under the independent assumption on latent failure times T_j, for $j = 1,...,K$, the joint survival function can be expressed in terms of marginal survival functions $S_j(t_j)$ as

$$\begin{aligned} S(t_1,...,t_K) &= P(T_1 > t_1,...,T_K > t_K) \\ &= P(T_1 > t_1)\cdot ....\cdot P(T_K > t_K) \\ &= \prod_{j=1}^{K} S_j(t_j) \\ &= \exp\left(-\sum_{j=1}^{K} \mu_j t_j^{\beta_j} \right) \end{aligned} \tag{3.2.4}$$

We propose a parametric competing risks model as

$$S(t_1,...,t_K) = \exp\left(-\left[\sum_{j=1}^{K} (\mu_j t_j^{\beta_j})^{1/\delta} \right]^{\delta} \right) \tag{3.2.5}$$

where $\mu_j \in (0,\infty)$, $\beta_j \in (0,\infty)$ and $\delta \in (0,1]$ is the dependence parameter. The boundary case of $\delta = 1$ yields the independent case.

For the model 3.2.5 above, it is easy to show the marginal survival function of each latent failure time T_j for $j = 1,...,K$ is the same as 3.2.3.

Theorem 3.2.2. *If the joint distribution of latent failure times $T_1,...,T_K$ is*

$$S(t_1,...,t_K) = \exp\left(-\left[\sum_{j=1}^{K} (\mu_j t_j^{\beta_j})^{1/\delta} \right]^{\delta} \right)$$

then the marginal distribution of each T_j for $j = 1,..,K$ belongs to a Weibull distribution with scale parameter μ_j and shape parameter β_j.

Proof. Since the joint survival function of $T_1,...,T_K$ is $S(t_1,...,t_K) = P(T_1 > t_1,...,T_K > t_K)$, it can be expressed as $S(t_1,...,t_K) = \int_{t_1}^{\infty} ... \int_{t_K}^{\infty} g(s_1,...,s_K)ds_1...ds_K$, where $g(s_1,...,s_K)$ is the joint density function of $T_1,...,T_K$. So the marginal survival function of T_j can be obtained by

$$\begin{aligned} S_j(t_j) &= \int_0^{\infty} ... \int_0^{\infty} \int_0^{\infty} ... \int_0^{\infty} g(s_1,...,s_K)ds_1...ds_{j-1}ds_{j+1}...ds_K \\ &= P(T_1 > 0,...,T_{j-1} > 0, T_j > t_j, T_{j+1} > 0, ...T_K > 0) \\ &= S(0,...,0,t_j,0,...0) \\ &= \exp(-\mu_j t_j^{\beta_j}) \end{aligned}$$

Thus, the marginal distribution of each latent failure time is a Weibull distribution with scale parameter μ_j and shape parameter β_j. □

The features and properties of the proposed competing risks model 3.2.5 is discussed in the following sections.

3.3 Representation of the Model via a Frailty Approach

In general, frailty models are used to analyze clustered survival data such as repeated measures data and recurrent events data. Duchateau & Janssen (2007) presented detailed discussions on frailty models. In many statistical problems arising from biomedicine, researchers have to handle complex disease processes. For instance, a person suffers from two or more diseases which are interacting and the deterioration of one disease could expedite the others. In such a situation, introduction of the frailty factor in the competing risks model, as explained in Section2.4.2, would be advisable. In fact, this kind of a frailty model, known as the *Shared Frailty Model*, is particularly applicable in establishing the model of dependent competing risks. The key point of such shared frailty approach is that the subjects who are in the same cluster all share the same frailty factor.

For example, people who suffer from the same disease would share a common weakness or frailty factor.

Popular frailty models include the Gamma, positive stable and inverse Gaussian frailty distributions. We consider the representation of proposed dependent competing risks model via the positive stable frailty structure, which is discussed next. The use of positive stable distributions in multivariate survival data analysis was considered in Hougaard (1986).

Conditionally given a random variable Z, $T_j|Z$ are independent Weibull distribution with conditional marginal survival functions,

$$S_j(t_j|Z) = \exp\left(-Z\cdot\mu_j^{1/\delta}t_j^{\beta_j/\delta}\right) \tag{3.3.1}$$

where the frailty factor Z is assumed to be a positive random variable associated with a positive stable distribution with parameter $\delta \in (0,1)$. Note that the positive stable density function of frailty factor Z mentioned by Duchateau & Janssen (2007) is given by

$$f(z) = -\frac{1}{\pi z}\sum_{k=1}^{\infty}\frac{\Gamma(k\delta+1)}{k!}(-z^{-\delta})^k\sin(\delta k\pi) \tag{3.3.2}$$

Marginalizing over Z, we can obtain the corresponding marginal survival function of T_j as

$$\begin{aligned} S_j(t_j) &= \int_0^{\infty} S_j(t_j|Z=z)\cdot f(z)dz \\ &= \int_0^{\infty}\exp\left(-z\cdot\mu_j^{1/\delta}t_j^{\beta_j/\delta}\right)\cdot f(z)dz \\ &= \int_0^{\infty}\exp(-z\cdot w)\cdot f(z)dz \\ &= E[\exp(-z\cdot w)] \end{aligned} \tag{3.3.3}$$

where $w = \mu_j^{1/\delta} t_j^{\beta_j/\delta}$. By applying the Laplace transform of z, $E[\exp(-zs)] = \exp(-s^{\delta})$ as in Chen et al. (2002), the expression 3.3.3 can be simplified as

$$\begin{aligned} S_j(t_j) &= \exp(-w^{\delta}) \\ &= \exp[-(\mu_j^{1/\delta} t_j^{\beta_j/\delta})^{\delta}] \\ &= \exp(-\mu_j t_j^{\beta_j}) \end{aligned} \tag{3.3.4}$$

Conditional on the frailty factor, $(T_1, \ldots, T_K)|Z$ are assumed to be independent. The joint survival function of $T_1, ..., T_K$ can be obtained in terms of marginal survival functions as

$$\begin{aligned} S(t_1, ..., t_K) &= \int_0^{\infty} S(t_1, ..., t_K|z) \cdot f(z) dz \\ &= \int_0^{\infty} P(T_1 > t_1, ..., T_K > t_K|z) \cdot f(z) dz \\ &= E[P(T_1 > t_1, ..., T_K > t_K|Z = z)] \\ &= E[(P(T_1 > t_1|Z = z) \cdot \ldots \cdot P(T_K > t_K|Z = z)] \\ &= E[S_1(t_1|Z) ... S_K(t_K|Z)] \end{aligned}$$

Furthermore, using the Laplace transform, the joint survival function is given by

$$\begin{aligned} S(t_1, ..., t_K) &= E\left[\exp(-Z(\sum_{j=1}^{K} \mu_j^{1/\delta} t_j^{\beta_j/\delta}))\right] \\ &= \exp\left(-\left[\sum_{j=1}^{K} \mu_j^{1/\delta} t_j^{\beta_j/\delta}\right]^{\delta}\right) \end{aligned} \tag{3.3.5}$$

Therefore, the proposed dependent competing risks model in 3.2.5 can alternatively be expressed as a representation via the positive stable frailty approach.

3.4 Representation of the Model via a Copula Approach

The copula approach has been reviewed in the previous chapter, Section 2.4.1. It helps to demonstrate the dependent relationship among random variables. By applying such method, the joint distribution of variables can be expressed in terms of the marginal distribution of each variable. In this section, we introduce the dependent competing risks model as based on an Archimedean copula. One of the popular Archimedean copulas is Gumbel copula (Gumbel, 1960) with the copula generator $\psi(s) = \exp(-s^{\delta})$ and copula generator inverse $\psi^{-1}(s) = [-\log(s)]^{1/\delta}$.

Suppose the marginal survival functions of each latent failure time T_j for $j = 1, ...K$ is given by $S_j(t_j)$; the joint survival function via the Gumbel copula is written as

$$\begin{aligned} S(t_1, ..., t_K) &= P(T_1 > t_1, ...T_K > t_K) \\ &= \psi[\psi^{-1}(S_1(t_1)) + ... + \psi^{-1}(S_K(t_K))] \\ &= \exp\left(-\left[(-\log[S_1(t_1)])^{1/\delta} + ... + (-\log[S_K(t_K)])^{1/\delta} \right]^{\delta} \right) \end{aligned} \tag{3.4.1}$$

where $\delta \in (0, 1]$ is the dependence parameter.

Therefore, the corresponding joint survival function of the parametric model, in which each marginal is Weibull distribution with survival function $S_j(t_j) = \exp(-\mu_j t_j^{\beta_j})$, under Gumbel copula is given by

$$\begin{aligned} S(t_1, ..., t_K) &= \exp\left(-\left[(-\log[\exp(-\mu_1 t_1^{\beta_1})])^{1/\delta} + ... + (-\log[\exp(-\mu_K t_K^{\beta_K})])^{1/\delta} \right]^{\delta} \right) \\ &= \exp\left(-\left[(\mu_1 t_1^{\beta_1})])^{1/\delta} + ... + (\mu_K t_K^{\beta_K})])^{1/\delta} \right]^{\delta} \right) \\ &= \exp\left(-\left[\sum_{j=1}^{K} (\mu_j t_j^{\beta_j})])^{1/\delta} \right]^{\delta} \right) \end{aligned} \tag{3.4.2}$$

Thus the proposed dependent competing risks model 3.2.5 can alternatively also be represented via a copula approach.

3.5 Bivariate Weibull Survival Model

The bivariate Weibull competing risks model corresponding 3.2.5 is given by

$$S(t_1,t_2) = \exp\left(-[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta_2})^{1/\delta}]^{\delta} \right) \tag{3.5.1}$$

In the following chapters, we will discuss the properties of proposed competing risks model 3.2.5. In particular, we look at identifiability issues for such model in a situation involving two competing risks.

CHAPTER 4

IDENTIFIABILITY OF PROPOSED DEPENDENT COMPETING RISKS MODEL

4.1 Introduction

Establishing identifiability of statistical models in competing risks research is important in understanding the behavior of models and checking the uniqueness of the obtained estimators of model parameters. Identifiability issues in competing risks model have been explored by Tsiatis (1975), Heckman & Honoré (1989) and Crowder (1991). Tsiatis (1975) showed that given a joint survival function with arbitrary dependence between the competing risks, there exists a different joint survival function in which the risks are independent and which reproduces the same sub-densities $f_j(t)$. Heckman & Honoré (1989) established identifiability of competing risks models under proportional hazards regression, whereas Crowder (1991) considered identifiability under the cause-specific hazards approach. Zheng & Klein (1995) established identifiability of dependent competing risks based on an assumed known copula.

In this chapter, we consider identifiability issues associated with our proposed model, in particular, the model with two competing risks. In the first section, we consider identifiability of the model under equal shape parameter. In the second section, we consider identifiability conditions of the proposed model under different shape parameters and a known dependence parameter. In the third section, we consider conditions under which our proposed model is identifiable.

4.2 Identifiability Issues Under Equality of Shape Parameters

Suppose there exists two competing risks in the system, and the corresponding latent failure times T_1, T_2 follow the proposed bivariate dependent competing risks model with equal shape parameters $\beta_1 = \beta_2 = \beta$ with survival function

$$S(t_1,t_2) = \exp\Big(- \big[(\mu_1 t_1^{\beta})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta})^{\frac{1}{\delta}}\big]^{\delta}\Big) \tag{4.2.1}$$

According to the representation via copula (in Section 3.4), the generator $\psi(x) = \exp(-x^{\delta})$ and generator inverse $\psi^{-1}(u) = (-\log u)^{\frac{1}{\delta}}$. Thus $S(t_1,t_2)$ can also be exhibited as

$$\begin{aligned} S(t_1,t_2) &= \psi\big[\psi^{-1}(S_1(t_1)) + \psi^{-1}(S_2(t_2))\big] \\ &= \exp\Big[- \Big(\big[-\log(\exp(-\mu_1 t_1^{\beta}))\big]^{\frac{1}{\delta}} + \big[-\log(\exp(-\mu_2 t_2^{\beta}))\big]^{\frac{1}{\delta}}\Big)^{\delta}\Big] \end{aligned} \tag{4.2.2}$$

with the marginal survival functions of T_1 and T_2 given by $S_1(t_1) = \exp(-\mu_1 t_1^{\beta})$ and $S_2(t_2) = \exp(-\mu_2 t_2^{\beta})$. We have the following theorem.

Theorem 4.2.1. *Suppose the latent failure times T_1, T_2 follow the joint survival function 4.2.2; then the marginal survival function $S_1(t_1)$ and $S_2(t_2)$ are not identifiable based on $T = \min(T_1, T_2)$ and $C = \arg\min(T_1, T_2)$.*

Proof. We define two new functions $S_1^*(t_1)$ and $S_2^*(t_2)$ following A. Wang et al. (2015) as follows:

$$S_1^*(t_1) = \phi\Big[\int_0^{t_1} \frac{\phi^{-1\prime}(\pi(u))}{\psi^{-1\prime}(\pi(u))} d\psi^{-1}(S_1(u))\Big] \tag{4.2.3}$$

and

$$S_2^*(t_2) = \phi\Big[\int_0^{t_2} \frac{\phi^{-1\prime}(\pi(u))}{\psi^{-1\prime}(\pi(u))} d\psi^{-1}, (S_2(u))\Big] \tag{4.2.4}$$

where

$$\begin{aligned}\pi(u) &= P(T_1 > u, T_2 > u) = S(u,u) \\ &= \psi\big[\psi^{-1}\big(S_1(u)\big) + \psi^{-1}\big(S_2(u)\big)\big] \\ &= \exp\big(-\big[(\mu_1 u^{\beta})^{\frac{1}{\delta}} + (\mu_2 u^{\beta})^{\frac{1}{\delta}}\big]^{\delta}\big) \\ &= \exp\big[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta} u^{\beta}\big] \end{aligned} \tag{4.2.5}$$

Since the generator $\psi(x) = \exp(-x^{\delta})$ with $\psi^{-1}(u) = (-\log u)^{\frac{1}{\delta}}$, it follows that

$$\psi'(x) = -\exp(-x^{\delta})\delta x^{\delta-1}$$

Thus

$$\psi^{-1}\big[\pi(u)\big] = \big[-\log\big(\exp[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta} u^{\beta}]\big)\big]^{\frac{1}{\delta}} = (\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}}) u^{\frac{\beta}{\delta}} \tag{4.2.6}$$

and

$$\psi^{-1}\big[S_1(u)\big] = \big[-\log\big(\exp(-\mu_1 u^{\beta})\big)\big]^{\frac{1}{\delta}} = \mu_1^{\frac{1}{\delta}} u^{\frac{\beta}{\delta}} \tag{4.2.7}$$

According to the fact that $\phi^{-1\prime}(\pi(u)) = \dfrac{1}{\phi'\big(\phi^{-1}[\pi(u)]\big)}$ and $\psi^{-1\prime}(\pi(u)) = \dfrac{1}{\psi'\big(\psi^{-1}[\pi(u)]\big)}$, it is easy to take the inverse of $\psi'\big[\psi^{-1}\big(\pi(u)\big)\big]$ to obtain $\psi^{-1\prime}(\pi(u))$. And $\psi'\big[\psi^{-1}\big(\pi(u)\big)\big]$ can be written as

$$\begin{aligned}\psi'\big[\psi^{-1}\big(\pi(u)\big)\big] &= -\exp\big[-((\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})u^{\frac{\beta}{\delta}})^{\delta}\big]\delta\big[(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})u^{\frac{\beta}{\delta}}\big]^{\delta-1} \\ &= -\exp\big[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta} u^{\beta}\big]\delta(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1} u^{\beta-\frac{\beta}{\delta}} \end{aligned} \tag{4.2.8}$$

As shown in 4.2.7, the $d\psi^{-1}\big(S_1(u)\big)$ is given by

$$d\psi^{-1}\big(S_1(u)\big) = d(\mu_1^{\frac{1}{\delta}}u^{\frac{\beta}{\delta}}) = \frac{\beta}{\delta}\mu_1^{\frac{1}{\delta}}u^{\frac{\beta}{\delta}-1}du \tag{4.2.9}$$

There is another key factor to obtain $S_1^*(t_1)$, which is $\phi^{-1\prime}(\pi(u))$. Now, let's define $\phi = \exp(x)$; then $\phi^{-1} = \log x$ and $\phi' = \exp(x)$, which implies

$$\phi^{-1}\big(\pi(u)\big) = \log\big(\exp[-(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta}u^{\beta}]\big) = -(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta}u^{\beta} \tag{4.2.10}$$

and

$$\phi'\big[\phi^{-1}\big(\pi(u)\big)\big] = \exp\big[-(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta}u^{\beta}\big] \tag{4.2.11}$$

By applying 4.2.11, 4.2.8 and 4.2.9 into the definition of $S_1^*(t_1)$ in 4.2.3, it is obvious that $S_1^*(t_1)$ could be obtained by computing ϕ function of $\int_0^t \frac{\phi^{-1\prime}(\pi(u))}{\psi^{-1\prime}(\pi(u))}d\psi^{-1}(S_1(u))$, which is

$$\begin{aligned}
\int_0^t \frac{\phi^{-1\prime}(\pi(u))}{\psi^{-1\prime}(\pi(u))}d\psi^{-1}(S_1(u)) &= \int_0^t \frac{-\exp[-(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta}u^{\beta}]\delta(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}u^{\beta-\frac{\beta}{\delta}}}{\exp[-(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta}u^{\beta}]}\frac{\beta}{\delta}\mu_1^{\frac{1}{\delta}}u^{\frac{\beta}{\delta}-1}du \\
&= \int_0^t -\beta(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}u^{\beta-1}du \\
&= -(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}u^{\beta}|_0^t \\
&= -(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}t^{\beta}
\end{aligned} \tag{4.2.12}$$

So $S_1^*(t_1)$ can be written as

$$\begin{aligned} S_1^*(t_1) &= \phi\Big[\int_0^{t_1} \frac{\phi^{-1\prime}(\pi(u))}{\psi^{-1\prime}(\pi(u))} d\psi^{-1}(S_1(u))\Big] \\ &= \phi\big[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}} t^{\beta}\big] \\ &= \exp\big[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}} t_1^{\beta}\big] \end{aligned} \tag{4.2.13}$$

Similarly, we can construct $S_2^*(t_2)$ via the definition 4.2.4 which is given by

$$\begin{aligned} S_2^*(t_2) &= \phi\Big[\int_0^{t_2} \frac{\phi^{-1\prime}(\pi(u))}{\psi^{-1\prime}(\pi(u))} d\psi^{-1}(S_2(u))\Big] \\ &= \exp\big[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_2^{\frac{1}{\delta}} t_2^{\beta}\big] \end{aligned} \tag{4.2.14}$$

Moreover, the joint distribution of $T_1,\ T_2$ could also be obtained by

$$\begin{aligned} S^*(t_1,t_2) &= \phi\big[\phi^{-1}(S_1^*(t_1)) + \phi^{-1}(S_2^*(t_2))\big] \\ &= \exp\big[\log(\exp[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}} t_1^{\beta}]) + \log(\exp[-(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_2^{\frac{1}{\delta}} t_2^{\beta}])\big] \\ &= \exp\big(-\big[(\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}} t_1^{\beta} + (\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_2^{\frac{1}{\delta}} t_2^{\beta}\big]\big) \end{aligned} \tag{4.2.15}$$

Therefore, the corresponding crude survival (sub-survival) functions $Q_1(t)$ from $S(t_1,t_2)$ and $Q_1^*(t)$ from $S^*(t_1,t_2)$ are given by

$$\begin{aligned} Q_1(t) &= P(T_1 > t, T_1 < T_2) \\ &= -\int_t^\infty P(T_1 = u, T_1 < T_2) dS_1(u) \\ &= -\int_t^\infty P(T_1 = u, T_2 > T_1) dS_1(u) \\ &= -\int_t^\infty P(T_1 = u, T_2 > u) dS_1(u) \\ &= -\int_t^\infty \left. \frac{dS(t_1,t_2)}{dt_1} \right|_{t_1=t_2=u} du \end{aligned} \tag{4.2.16}$$

where t is an arbitrary point, and

$$\begin{aligned} Q_1^*(t) &= P(T_1^* > t, T_1^* < T_2^*) \\ &= -\int_t^\infty P(T_1^* = u, T_1^* < T_2^*) dS_1^*(u) \\ &= -\int_t^\infty P(T_1^* = u, T_2^* > T_1^*) dS_1^*(u) \\ &= -\int_t^\infty P(T_1^* = u, T_2^* > u) dS_1^*(u) \\ &= -\int_t^\infty \left. \frac{dS^*(t_1,t_2)}{dt_1} \right|_{t_1=t_2=u} du \end{aligned} \tag{4.2.17}$$

where t is an arbitrary point.

Now, we will show $Q_1(t) = Q_1^*(t)$ by showing

$$\left. \frac{dS(t_1,t_2)}{dt_1} \right|_{t_1=t_2=u} = \left. \frac{dS^*(t_1,t_2)}{dt_1} \right|_{t_1=t_2=u}$$

As it is known that

$$\begin{aligned}
\left.\frac{dS(t_1,t_2)}{dt_1}\right|_{t_1=t_2=u} &= \left.\frac{d\exp\Big(-\big[(\mu_1 t_1^{\beta})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{dt_1}\right|_{t_1=t_2=u} \\
&= \left.\frac{d\exp\Big(-\big[(\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}})+(\mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}})\big]^{\delta}\Big)}{dt_1}\right|_{t_1=t_2=u} \\
&= -\exp\Big(-\big[(\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}})+(\mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}})\big]^{\delta}\Big)\delta\big[(\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}})+(\mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}})\big]^{\delta-1} \\
&\quad \times\mu_1^{\frac{1}{\delta}}\frac{\beta}{\delta}t^{\frac{\beta}{\delta}-1}\Big|_{t_1=t_2=u} \\
&= -\exp\Big(-\big[(\mu_1^{\frac{1}{\delta}} u^{\frac{\beta}{\delta}})+(\mu_2^{\frac{1}{\delta}} u^{\frac{\beta}{\delta}})\big]^{\delta}\Big)\delta\big[(\mu_1^{\frac{1}{\delta}} u^{\frac{\beta}{\delta}})+(\mu_2^{\frac{1}{\delta}} u^{\frac{\beta}{\delta}})\big]^{\delta-1}\mu_1^{\frac{1}{\delta}}\frac{\beta}{\delta}u^{\frac{\beta}{\delta}-1} \\
&= -\exp\Big[-\big(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}\big)^{\delta}u^{\beta}\Big]\big(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}\big)^{\delta-1}\mu_1^{\frac{1}{\delta}}\beta u^{\beta-1}
\end{aligned}$$

and

$$\begin{aligned}
\left.\frac{dS^*(t_1,t_2)}{dt_1}\right|_{t_1=t_2=u} &= \left.\frac{d\exp\big(-\big[(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}t_1^{\beta}+(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_2^{\frac{1}{\delta}}t_2^{\beta}\big]\big)}{dt_1}\right|_{t_1=t_2=u} \\
&= -\exp\Big(-\big[(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}t_1^{\beta}+(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_2^{\frac{1}{\delta}}t_2^{\beta}\big]\Big) \\
&\quad \times(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}\beta t^{\beta-1}\Big|_{t_1=t_2=u} \\
&= -\exp\Big(-\big[(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}u^{\beta}+(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_2^{\frac{1}{\delta}}u^{\beta}\big]\Big) \\
&\quad \times(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}\beta u^{\beta-1} \\
&= -\exp\Big[-(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta}u^{\beta}\Big](\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}})^{\delta-1}\mu_1^{\frac{1}{\delta}}\beta u^{\beta-1}
\end{aligned}$$

therefore, it is explicit that $Q_1(t) = Q_1^*(t)$, so the crude survival functions of T_1 are the same.

Similarly, we can derive that

$$\left.\frac{dS(t_1,t_2)}{dt_2}\right|_{t_1=t_2=u} = -\exp\Big[-\big(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}\big)^{\delta}u^{\beta}\Big]\big(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}\big)^{\delta-1}\mu_2^{\frac{1}{\delta}}\beta u^{\beta-1} \tag{4.2.18}$$

and

$$\left.\frac{dS^*(t_1,t_2)}{dt_2}\right|_{t_1=t_2=u} = -\exp\Big[-\big(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}\big)^{\delta}u^{\beta}\Big]\big(\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}\big)^{\delta-1}\mu_2^{\frac{1}{\delta}}\beta u^{\beta-1} \tag{4.2.19}$$

to show that $Q_2(t) = Q_2^*(t)$, where t is an arbitrary point. Thus, the crude survival functions of T_2 are the same. □

We have thus established that there is an alternative joint survival function of latent failure times T_1, T_2 given by $S^*(t_1,t_2)$ in 4.2.15 that leads to the same crude survival function of $T_1,\ T_2$ as the joint survival function $S(t_1,t_2)$ in 4.2.2. This establishes that the proposed bivariate dependent competing risks model under equal shape parameters is not identifiable.

4.3 Identifiability with Unrestricted Shape Parameters and a Known Dependence Parameter

In a general setting, the shape parameters in the Weibull distribution of T_1 and T_2 usually may not be equal. In this section, we explore identifiability of the proposed model without the equality assumption of $\beta_1 = \beta_2$. Suppose the latent failure times T_1 and T_2 follow the joint survival model

$$S(t_1,t_2) = \exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big) \tag{4.3.1}$$

The marginal survival functions of T_1 and T_2 are given by $S_1(t) = \exp(-\mu_1 t^{\beta_1})$, $S_2(c) = \exp(-\mu_2 c^{\beta_2})$, and the corresponding crude survival functions can be written as

$$\begin{aligned} Q_1(t) &= P(T_1 > t, T_1 < T_2) \\ &= -\int_t^\infty P(T_1 = u, T_1 < T_2) dS_1(u) \\ &= -\int_t^\infty P(T_1 = u, T_2 > T_1) dS_1(u) \\ &= -\int_t^\infty P(T_1 = u, T_2 > u) dS_1(u) \\ &= -\int_t^\infty \left. \frac{dS(t_1,t_2)}{dt_1} \right|_{t_1=t_2=u} du \end{aligned} \tag{4.3.2}$$

and

$$\begin{aligned} Q_2(t) &= P(T_2 > t, T_2 < T_1) \\ &= -\int_t^\infty \left. \frac{dS(t_1,t_2)}{dt_2} \right|_{t_1=t_2=u} du \end{aligned} \tag{4.3.3}$$

We establish identifiability of the model in the following theorem.

Theorem 4.3.1. *Suppose the joint survival function of T_1 and T_2 is given by*

$$S(t_1,t_2) = \exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta} \Big)$$

If the dependent parameter δ is known, then the crude survival functions of T_1 and T_2 are identifiable based on $T = \min(T_1,T_2)$ and $C = \arg\min(T_1,T_2)$.

Proof. The proof of the theorem can be expounded by showing that $\dfrac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$ and $\dfrac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$ are simultaneously true for $\forall t > 0$ if and only if $\mu_1 = \mu_1^*, \mu_2 = \mu_2^*, \beta_1 = \beta_1^*, \beta_2 = \beta_2^*$.

Since the crude survival functions of T_1 and T_2 are given by $Q_1(t)$ and $Q_2(t)$ in 4.3.2 and 4.3.3, we have

$$\begin{aligned}
\frac{dQ_1(t)}{dt} &= -\frac{dS(t_1,t_2)}{dt_1}\bigg|_{t_1=t_2=t} \\
&= -\frac{d\exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{dt}\bigg|_{t_1=t_2=t} \\
&= -\exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)\delta\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta-1}\mu_1^{\frac{1}{\delta}}\frac{\beta_1}{\delta}t_1^{\frac{\beta_1}{\delta}-1}\bigg|_{t_1=t_2=t} \\
&= -\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)\big[\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}t^{\frac{\beta_2-\beta_1}{\delta}}\big]^{\delta-1}\beta_1\mu_1^{\frac{1}{\delta}}t^{\beta_1-1} \qquad (4.3.4)
\end{aligned}$$

and

$$\begin{aligned}
\frac{dQ_2(t)}{dt} &= -\frac{dS(t_1,t_2)}{dt_2}\bigg|_{t_1=t_2=t} \\
&= -\frac{d\exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{dt_2}\bigg|_{t_1=t_2=t} \\
&= -\exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)\delta\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta-1}\mu_2^{\frac{1}{\delta}}\frac{\beta_2}{\delta}t^{\frac{\beta_2}{\delta}-1}\bigg|_{t_1=t_2=t} \\
&= -\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)\big[\mu_1^{\frac{1}{\delta}}t^{\frac{\beta_1-\beta_2}{\delta}}+\mu_2^{\frac{1}{\delta}}\big]^{\delta-1}\beta_2\mu_2^{\frac{1}{\delta}}t^{\beta_2-1} \qquad (4.3.5)
\end{aligned}$$

therefore, the expressions of $\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt}$ and $\frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt}$ are written as

$$\begin{aligned}\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} &= \frac{-\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)\big[\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}t^{\frac{\beta_2-\beta_1}{\delta}}\big]^{\delta-1}}{-\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)\big[(\mu_1^*)^{\frac{1}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}t^{\frac{\beta_2^*-\beta_1^*}{\delta}}\big]^{\delta-1}} \\ &\quad\times\frac{\beta_1\mu_1^{\frac{1}{\delta}}t^{\beta_1-1}}{\beta_1^*(\mu_1^*)^{\frac{1}{\delta}}t^{\beta_1^*-1}} \\ &= \frac{\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)}\left[\frac{\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}t^{\frac{\beta_2-\beta_1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}t^{\frac{\beta_2^*-\beta_1^*}{\delta}}}\right]^{\delta-1} \\ &\quad\times\left[\frac{\mu_1}{\mu_1^*}\right]^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}t^{\beta_1-\beta_1^*} \end{aligned} \tag{4.3.6}$$

and

$$\begin{aligned}\frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} &= \frac{-\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)\big[\mu_1^{\frac{1}{\delta}}t^{\frac{\beta_1-\beta_2}{\delta}}+\mu_2^{\frac{1}{\delta}}\big]^{\delta-1}}{-\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)\big[(\mu_1^*)^{\frac{1}{\delta}}t^{\frac{\beta_1^*-\beta_2^*}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}\big]^{\delta-1}} \\ &\quad\times\frac{\beta_2\mu_2^{\frac{1}{\delta}}t^{\beta_2-1}}{\beta_2^*(\mu_2^*)^{\frac{1}{\delta}}t^{\beta_2^*-1}} \\ &= \frac{\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)}\left[\frac{\mu_1^{\frac{1}{\delta}}t^{\frac{\beta_1-\beta_2}{\delta}}+\mu_2^{\frac{1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}t^{\frac{\beta_1^*-\beta_2^*}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}}\right]^{\delta-1} \\ &\quad\times\left[\frac{\mu_2}{\mu_2^*}\right]^{\frac{1}{\delta}}\frac{\beta_2}{\beta_2^*}t^{\beta_2-\beta_2^*} \end{aligned} \tag{4.3.7}$$

In order to show $\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt}=1$ and $\frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt}=1$ are true simultaneous, equivalently, we need to show 4.3.6 and 4.3.7 equal to 1 simultaneous, that is, to show

$$\frac{\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)}\left[\frac{\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}t^{\frac{\beta_2-\beta_1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}t^{\frac{\beta_2^*-\beta_1^*}{\delta}}}\right]^{\delta-1}\left[\frac{\mu_1}{\mu_1^*}\right]^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}t^{\beta_1-\beta_1^*}=1 \tag{4.3.8}$$

and

$$\frac{\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)}\left[\frac{\mu_1^{\frac{1}{\delta}}t^{\frac{\beta_1-\beta_2}{\delta}}+\mu_2^{\frac{1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}t^{\frac{\beta_1^*-\beta_2^*}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}}\right]^{\delta-1}\left[\frac{\mu_2}{\mu_2^*}\right]^{\frac{1}{\delta}}\frac{\beta_2}{\beta_2^*}t^{\beta_2-\beta_2^*}=1 \tag{4.3.9}$$

are true simultaneous.

Case1: When $\mu_1=\mu_1^*,\mu_2=\mu_2^*,\beta_1=\beta_1^*,\beta_2=\beta_2^*$, equation 4.3.8 and equation 4.3.9 are obvious true.

Case2: When $\mu_1\neq\mu_1^*,\mu_2\neq\mu_2^*,\beta_1\neq\beta_1^*,\beta_2\neq\beta_2^*$, take the limit of equation 4.3.6 and equation 4.3.7, then

$$\begin{aligned}
\lim_{t\longrightarrow 0}\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} &= \lim_{t\longrightarrow 0}\frac{\exp\Big(-\big[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big)}{\exp\Big(-\big[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\big]^{\delta}\Big)}\\
&\quad\times\left[\frac{\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}t^{\frac{\beta_2-\beta_1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}t^{\frac{\beta_2^*-\beta_1^*}{\delta}}}\right]^{\delta-1}\left[\frac{\mu_1}{\mu_1^*}\right]^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}t^{\beta_1-\beta_1^*}\\
&= \lim_{t\longrightarrow 0}\left[\frac{\mu_1^{\frac{1}{\delta}}+\mu_2^{\frac{1}{\delta}}t^{\frac{\beta_2-\beta_1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}t^{\frac{\beta_2^*-\beta_1^*}{\delta}}}\right]^{\delta-1}\left[\frac{\mu_1}{\mu_1^*}\right]^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}t^{\beta_1-\beta_1^*}\\
&\propto \lim_{t\longrightarrow 0}\frac{\mu_2^{1-\frac{1}{\delta}}}{(\mu_2^*)^{1-\frac{1}{\delta}}}\frac{\beta_1}{\beta_1^*}\left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}}t^{\beta_1-\beta_1^*+\beta_2-\beta_1-\frac{\beta_2-\beta_1}{\delta}-\beta_2^*+\beta_1^*+\frac{\beta_2^*-\beta_1^*}{\delta}}\\
&= \lim_{t\longrightarrow 0}\left(\frac{\mu_2}{\mu_2^*}\right)^{1-\frac{1}{\delta}}\left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}t^{\beta_2-\beta_2^*-\frac{\beta_2-\beta_1}{\delta}+\frac{\beta_2^*-\beta_1^*}{\delta}}
\end{aligned} \tag{4.3.10}$$

a. When $\beta_2>\beta_1,\beta_2^*>\beta_1^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t\longrightarrow 0}\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt}=\begin{cases}0, & \text{if } \beta_1>\beta_1^*\\ \infty, & \text{if } \beta_1<\beta_1^*\end{cases}$$

b. When $\beta_2>\beta_1,\beta_2^*<\beta_1^*$:

1. $\beta_2 < \beta_2^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \infty$$

2. $\beta_2 > \beta_2^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_2 - \beta_2^* > \frac{\beta_2-\beta_1}{\delta} - \frac{\beta_2^*-\beta_1^*}{\delta} \\ \infty, \\ \text{if } \beta_2 - \beta_2^* < \frac{\beta_2-\beta_1}{\delta} - \frac{\beta_2^*-\beta_1^*}{\delta} \\ \left(\frac{\mu_2}{\mu_2^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}} \frac{\beta_1}{\beta_1^*}, \\ \text{if } \beta_2 - \beta_2^* = \frac{\beta_2-\beta_1}{\delta} - \frac{\beta_2^*-\beta_1^*}{\delta} \end{cases}$$

c. When $\beta_2 < \beta_1, \beta_2^* > \beta_1^*$:

1. $\beta_2 > \beta_2^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 0$$

2. $\beta_2 < \beta_2^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_2 - \beta_2^* > \frac{\beta_2-\beta_1}{\delta} - \frac{\beta_2^*-\beta_1^*}{\delta} \\ \infty, \\ \text{if } \beta_2 - \beta_2^* < \frac{\beta_2-\beta_1}{\delta} - \frac{\beta_2^*-\beta_1^*}{\delta} \\ \left(\frac{\mu_2}{\mu_2^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}} \frac{\beta_1}{\beta_1^*}, \\ \text{if } \beta_2 - \beta_2^* = \frac{\beta_2-\beta_1}{\delta} - \frac{\beta_2^*-\beta_1^*}{\delta} \end{cases}$$

d. When $\beta_2 < \beta_1, \beta_2^* < \beta_1^*$:

1. $\beta_2 < \beta_2^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_2 - \beta_2^* + \frac{\beta_2^*-\beta_1^*}{\delta} > \frac{\beta_2-\beta_1}{\delta} \\ \infty, \\ \text{if } \beta_2 - \beta_2^* + \frac{\beta_2^*-\beta_1^*}{\delta} < \frac{\beta_2-\beta_1}{\delta} \\ \left(\frac{\mu_2}{\mu_2^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}} \frac{\beta_1}{\beta_1^*}, \\ \text{if } \beta_2 - \beta_2^* + \frac{\beta_2^*-\beta_1^*}{\delta} = \frac{\beta_2-\beta_1}{\delta} \end{cases}$$

2. $\beta_2 > \beta_2^*$, equation 4.3.10 $\Longrightarrow$

$$\lim_{t\longrightarrow 0}\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_2-\beta_2^*-\frac{\beta_2-\beta_1}{\delta} > \frac{\beta_2^*-\beta_1^*}{\delta} \\ \infty, \\ \text{if } \beta_2-\beta_2^*-\frac{\beta_2-\beta_1}{\delta} < \frac{\beta_2^*-\beta_1^*}{\delta} \\ \left(\frac{\mu_2}{\mu_2^*}\right)^{1-\frac{1}{\delta}}\left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}, \\ \text{if } \beta_2-\beta_2^*-\frac{\beta_2-\beta_1}{\delta} = \frac{\beta_2^*-\beta_1^*}{\delta} \end{cases}$$

and

$$\begin{aligned}
\lim_{t\longrightarrow 0}\frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} &= \lim_{t\longrightarrow 0}\frac{\exp\left(-\left[(\mu_1 t^{\beta_1})^{\frac{1}{\delta}}+(\mu_2 t^{\beta_2})^{\frac{1}{\delta}}\right]^{\delta}\right)}{\exp\left(-\left[(\mu_1^* t^{\beta_1^*})^{\frac{1}{\delta}}+(\mu_2^* t^{\beta_2^*})^{\frac{1}{\delta}}\right]^{\delta}\right)} \\
&\quad\times\left[\frac{\mu_1^{\frac{1}{\delta}}t^{\frac{\beta_1-\beta_2}{\delta}}+\mu_2^{\frac{1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}t^{\frac{\beta_1^*-\beta_2^*}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}}\right]^{\delta-1}\left[\frac{\mu_2}{\mu_2^*}\right]^{\frac{1}{\delta}}\frac{\beta_2}{\beta_2^*}t^{\beta_2-\beta_2^*} \\
&= \lim_{t\longrightarrow 0}\left[\frac{\mu_1^{\frac{1}{\delta}}t^{\frac{\beta_1-\beta_2}{\delta}}+\mu_2^{\frac{1}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}}t^{\frac{\beta_1^*-\beta_2^*}{\delta}}+(\mu_2^*)^{\frac{1}{\delta}}}\right]^{\delta-1}\left[\frac{\mu_2}{\mu_2^*}\right]^{\frac{1}{\delta}}\frac{\beta_2}{\beta_2^*}t^{\beta_2-\beta_2^*} \\
&\propto \lim_{t\longrightarrow 0}\frac{\mu_1^{1-\frac{1}{\delta}}}{(\mu_1^*)^{1-\frac{1}{\delta}}}\frac{\beta_2}{\beta_2^*}\left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}}t^{\beta_2-\beta_2^*+\beta_1-\beta_2-\frac{\beta_1-\beta_2}{\delta}-\beta_1^*+\beta_2^*+\frac{\beta_1^*-\beta_2^*}{\delta}} \\
&= \lim_{t\longrightarrow 0}\left(\frac{\mu_1}{\mu_1^*}\right)^{1-\frac{1}{\delta}}\left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}}\frac{\beta_2}{\beta_2^*}t^{\beta_1-\beta_1^*-\frac{\beta_1-\beta_2}{\delta}+\frac{\beta_1^*-\beta_2^*}{\delta}} \qquad (4.3.11)
\end{aligned}$$

a. When $\beta_2 > \beta_1, \beta_2^* > \beta_1^*$,

1. $\beta_1 < \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_1 - \beta_1^* + \frac{\beta_1^* - \beta_2^*}{\delta} > \frac{\beta_1 - \beta_2}{\delta} \\ \infty, \\ \text{if } \beta_1 - \beta_1^* + \frac{\beta_1^* - \beta_2^*}{\delta} < \frac{\beta_1 - \beta_2}{\delta} \\ \left(\frac{\mu_1}{\mu_1^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}} \frac{\beta_2}{\beta_2^*}, \\ \text{if } \beta_1 - \beta_1^* + \frac{\beta_1^* - \beta_2^*}{\delta} = \frac{\beta_1 - \beta_2}{\delta} \end{cases}$$

2. $\beta_1 > \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_1 - \beta_1^* - \frac{\beta_1 - \beta_2}{\delta} > \frac{\beta_1^* - \beta_2^*}{\delta} \\ \infty, \\ \text{if } \beta_1 - \beta_1^* - \frac{\beta_1 - \beta_2}{\delta} < \frac{\beta_1^* - \beta_2^*}{\delta} \\ \left(\frac{\mu_1}{\mu_1^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}} \frac{\beta_2}{\beta_2^*}, \\ \text{if } \beta_1 - \beta_1^* - \frac{\beta_1 - \beta_2}{\delta} = \frac{\beta_1^* - \beta_2^*}{\delta} \end{cases}$$

b. When $\beta_2 > \beta_1, \beta_2^* < \beta_1^*$,

1. $\beta_1 > \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 0$$

2. $\beta_1 < \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_1 - \beta_1^* > \frac{\beta_1-\beta_2}{\delta} - \frac{\beta_1^*-\beta_2^*}{\delta} \\ \infty, \\ \text{if } \beta_1 - \beta_1^* < \frac{\beta_1-\beta_2}{\delta} - \frac{\beta_1^*-\beta_2^*}{\delta} \\ \left(\frac{\mu_1}{\mu_1^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}} \frac{\beta_2}{\beta_2^*}, \\ \text{if } \beta_1 - \beta_1^* = \frac{\beta_1-\beta_2}{\delta} - \frac{\beta_1^*-\beta_2^*}{\delta} \end{cases}$$

c. When $\beta_2 < \beta_1, \beta_2^* > \beta_1^*$,

1. $\beta_1 < \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \infty$$

2. $\beta_1 > \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, \\ \text{if } \beta_1 - \beta_1^* > \frac{\beta_1-\beta_2}{\delta} - \frac{\beta_1^*-\beta_2^*}{\delta} \\ \infty, \\ \text{if } \beta_1 - \beta_1^* < \frac{\beta_1-\beta_2}{\delta} - \frac{\beta_1^*-\beta_2^*}{\delta} \\ \left(\frac{\mu_1}{\mu_1^*}\right)^{1-\frac{1}{\delta}} \left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}} \frac{\beta_2}{\beta_2^*}, \\ \text{if } \beta_1 - \beta_1^* = \frac{\beta_1-\beta_2}{\delta} - \frac{\beta_1^*-\beta_2^*}{\delta} \end{cases}$$

d. When $\beta_2 < \beta_1, \beta_2^* < \beta_1^*$, equation 4.3.11 $\Longrightarrow$

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = \begin{cases} 0, & \text{if } \beta_2 > \beta_2^* \\ \infty, & \text{if } \beta_2 < \beta_2^* \end{cases}$$

For both sub-case a, when $\beta_2 > \beta_1, \beta_2^* > \beta_1^*$, it could be true that

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$$

however,

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} \neq 1$$

So, equation 4.3.8 and equation 4.3.9 cannot hold simultaneously; it is a contradiction!

For both sub-case b, when $\beta_2 > \beta_1, \beta_2^* < \beta_1^*$, suppose it is true simultaneously that

$$\lim_{t \longrightarrow 0} \frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$$

and

$$\lim_{t \longrightarrow 0} \frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$$

which implies

$$\left(\frac{\mu_2}{\mu_2^*}\right)^{1-\frac{1}{\delta}}\left(\frac{\mu_1}{\mu_1^*}\right)^{\frac{1}{\delta}}\frac{\beta_1}{\beta_1^*}=1,$$
$$\left(\frac{\mu_1}{\mu_1^*}\right)^{1-\frac{1}{\delta}}\left(\frac{\mu_2}{\mu_2^*}\right)^{\frac{1}{\delta}}\frac{\beta_2}{\beta_2^*}=1,$$
$$\beta_2>\beta_2^*,$$
$$\beta_1<\beta_1^*,$$
$$\beta_1-\beta_1^*=\frac{\beta_1-\beta_2}{\delta}-\frac{\beta_1^*-\beta_2^*}{\delta}, \tag{4.3.12}$$
$$\beta_2-\beta_2^*=\frac{\beta_2-\beta_1}{\delta}-\frac{\beta_2^*-\beta_1^*}{\delta} \tag{4.3.13}$$

From condition 4.3.12 and condition 4.3.13 , we can gain that

$$\beta_2-\beta_2^*=\beta_1^*-\beta_1 \tag{4.3.14}$$

Combine condition 4.3.14 and condition 4.3.12, it is easy to obtain

$$(1-\frac{1}{\delta})(\beta_2-\beta_2^*)=\frac{1}{\delta}(\beta_1^*-\beta_1)$$

so $1-\frac{1}{\delta}=\frac{1}{\delta}$, which implies $\delta=2$; contradictions!

For both sub-case c, when $\beta_2<\beta_1,\beta_2^*>\beta_1^*$, we could get the same contradiction with sub-case b.

For both sub-case d, when $\beta_2<\beta_1,\beta_2^*<\beta_1^*$, it is impossible to make equation 4.3.8 and equation 4.3.9 hold simultaneously.

Therefore, when $\mu_1\neq\mu_1^*,\mu_2\neq\mu_2^*,\beta_1\neq\beta_1^*,\beta_2\neq\beta_2^*$, equation 4.3.8 and equation 4.3.9 cannot hold simultaneously.

So $\frac{dQ_1(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_1^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$ and $\frac{dQ_2(t|\mu_1,\mu_2,\beta_1,\beta_2)/dt}{dQ_2^*(t|\mu_1^*,\mu_2^*,\beta_1^*,\beta_2^*)/dt} = 1$ are true simultaneous for $\forall t > 0$ if and only if $\mu_1 = \mu_1^*, \mu_2 = \mu_2^*, \beta_1 = \beta_1^*, \beta_2 = \beta_2^*$. □

As a consequence, for a known dependent parameter, the crude survival functions are identifiable, and hence the corresponding joint survival model 4.3.1 is identifiable.

4.4 Identifiability Issues Under Unknown Dependence Parameter and a Regression Framework

In general, the dependence parameter δ may not be known, and the identifiability of the model with unknown δ is explored in this section. Consider the potential latent failure times of two risks which are denoted by T_1 and T_2 and satisfy with the proposed bivariate competing risks model

$$S(t_1,t_2) = \exp\Big(-\big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\big]^{\delta}\Big) \tag{4.4.1}$$

which can also be represented through a copula approach as

$$S(t_1,t_2) = \psi\Big[\psi^{-1}\big(S_1(t_1)\big) + \psi^{-1}\big(S_2(t_2)\big)\Big] \tag{4.4.2}$$

where generator $\psi = \exp(-x^{\delta})$ and generator inverse $\psi^{-1} = (-\log x)^{\frac{1}{\delta}}$, and the marginal survival functions are $S_1(t_1) = \exp(-\mu_1 t_1^{\beta_1})$ and $S_2(t_2) = \exp(-\mu_2 t_2^{\beta_2})$.

An alternative representation of the proposed bivariate model 4.4.1 is via a frailty approach. The process of constructing the model is exhibited as follows. As remarked in Oakes (1989), that Archimedean copula models arise spontaneous from bivariate frailty models, that is, T_1 and T_2 are

conditionally independent given an unobserved frailty factor Z. $T_1|Z$ and $T_2|Z$ follow the Weibull distribution with survival functions

$$S_1(t_1|Z) = \exp\big(-\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta_1}{\delta}} Z\big) \tag{4.4.3}$$

$$S_2(t_2|Z) = \exp\big(-\mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta_2}{\delta}} Z\big) \tag{4.4.4}$$

where Z follows a positive stable distribution with parameter $\delta \in (0,1)$.

We consider a competing risks regression framework for a given covariate vector $\mathbf{X}$ under a proportional hazards framework as

$$\log \mu_1^{\frac{1}{\delta}} = \eta_{0t_1} + X\eta_{t_1} \tag{4.4.5}$$

$$\log \mu_2^{\frac{1}{\delta}} = \eta_{0t_2} + X\eta_{t_2} \tag{4.4.6}$$

This results in

$$\begin{aligned} h_1(t_1|\beta_1,\mu_1,\delta,z) &= -\frac{\partial S_1(t_1|z)/\partial t_1}{S_1(t_1|z)} \\ &= \mu_1^{\frac{1}{\delta}} \frac{\beta_1}{\delta} t_1^{\frac{\beta_1}{\delta}-1} z \\ &= \underbrace{\exp(\mathbf{X}\eta_{\mathbf{t_1}})}\cdot \underbrace{\exp(\eta_{0t_1})\frac{\beta_1}{\delta} t_1^{\frac{\beta_1}{\delta}-1}}\cdot z \\ &= g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)\cdot z \end{aligned} \tag{4.4.7}$$

where $h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)$ is the baseline hazard function of latent failure time T_1.

Similarly, we have

$$h_2(t_2|\beta_2,\mu_2,\delta,z) = g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot h_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\cdot z \tag{4.4.8}$$

The corresponding cumulative hazard functions could be written as

$$H_1(t_1|z) = g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)\cdot z \tag{4.4.9}$$

$$H_2(t_2|z) = g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\cdot z \tag{4.4.10}$$

where $H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)$ and $H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)$ are the baseline cumulative hazard functions. Note that $\mathbf{X}$ is the covariates vector which is shared by T_1 and T_2. The frailty factor Z is a common random variable which is also shared by T_1 and T_2, such that T_1 and T_2 are conditionally independent. According to Abbring & Van den Berg (2003), such extended model is called mixed proportional hazards competing risks model as well.

Therefore, the joint survival function of T_1 and T_2 under the known covariates vector is given by

$$\begin{aligned} S(t_1,t_2|\mathbf{X}) &= S(t_1,t_2|Z,\mathbf{X}) \\ &= E\big[P(T_1>t_1,T_2>t_2)|Z,\mathbf{X}\big] \\ &= E\big[P(T_1>t_1|Z)\cdot P(T_2>t_2|Z,\mathbf{X})\big] \\ &= E\big[S_1(t_1|Z,\mathbf{X})\cdot S_2(t_2|Z,\mathbf{X})\big] \end{aligned}$$

Based on the relationship between survival function and cumulative hazard function that $S(t) = \exp[-H(t)]$ and 4.4.10, 4.4.10, the aforementioned joint distribution of T_1 and T_2 can be written as

$$\begin{aligned} S(t_1,t_2|\mathbf{X}) &= E\Big[\exp\big(-g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)\cdot Z\big)\cdot\exp\big(-g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\cdot Z\big)\Big] \\ &= E\left[\exp\Big(-\big(g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_C(\mathbf{X}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0c_2},\delta)\big)\cdot Z\Big)\right] \end{aligned}$$

By applying the Laplace transform of Z, which is $\psi(s) = E[\exp(-sZ)]$, the aforementioned joint distribution could be simplified as

$$\begin{aligned} S(t_1,t_2|\mathbf{X}) &= \psi\Big[g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) + g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0c_2},\delta)\Big] \\ &\Big\downarrow \quad \textit{since } \psi^{-1}(s) = -\frac{\log S}{Z} \\ &= \psi\Big[\psi^{-1}\big(S_1(t_1|\mathbf{X})\big) + \psi^{-1}\big(S_2(t_2|\mathbf{X})\big)\Big] \end{aligned}$$

As a consequence, the following lemma can be established.

Lemma 4.4.1. *Suppose the latent failure times T_1 and T_2 follow the mixed proportional hazards competing risk model under a given covariate $\mathbf{X}$ in the representation via a frailty approach 4.4.7, 4.4.8, that is, $h_1(t_1|\beta_1,\mu_1,\delta,z) = g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)\cdot z$, $h_2(t_2|\beta_2,\mu_2,\delta,z) = g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot h_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\cdot z$. Then under the Laplace transform $\psi(s) = E[\exp(-sZ)]$, the joint survival function of T_1 and T_2 is given by*

$$S(t_1,t_2|\mathbf{X}) = \psi\Big[\psi^{-1}\big(S_1(t_1|\mathbf{X})\big) + \psi^{-1}\big(S_2(t_2|\mathbf{X})\big)\Big] \tag{4.4.11}$$

where ψ^{-1} is the inverse function of ψ, and $S_1(t_1|\mathbf{X})$, $S_2(t_2|\mathbf{X})$ are corresponding marginal survival functions given by

$$S_1(t_1|\mathbf{X}) = \psi[g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)] \tag{4.4.12}$$

and

$$S_2(t_2|\mathbf{X}) = \psi[g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0c_2},\delta)] \tag{4.4.13}$$

Note that $H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)$ and $H_{0_2}(t_2|\beta_2,\eta_{0c_2},\delta)$ are the baseline cumulative hazard functions which are given by $H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) = \exp(\eta_{0t_1})t_1^{\frac{\beta_1}{\delta}}$ and $H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta) = \exp(\eta_{0t_2})t_2^{\frac{\beta_2}{\delta}}\cdot\ g_1(\mathbf{X}\eta_{\mathbf{t_1}})$

and $g_2(\mathbf{X}\eta_{\mathbf{t_2}})$ *are regression functions which do not involve latent failure times and are given in 4.4.7 and 4.4.8.*

It is obvious from the previous lemma that (1) the regression functions $g_1(\mathbf{X}\eta_{\mathbf{t_1}}) = \exp(\mathbf{X}\eta_{\mathbf{t_1}})$ and $g_2(\mathbf{X}\eta_{\mathbf{t_2}}) = \exp(\mathbf{X}\eta_{\mathbf{t_2}})$ are strictly convex functions; (2) the baseline cumulative hazard functions $H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) = \exp(\eta_{0t_1})t_1^{\frac{\beta_1}{\delta}}$ and $H_{0_2}(t_2|\beta_2,\eta_{0c_2},\delta) = \exp(\eta_{0t_2})t_2^{\frac{\beta_2}{\delta}}$ are differentiable with respect to t_1 and t_2 respectively.

Furthermore, with an established set of conditions, the model 4.4.7 and 4.4.8 are identifiable. Such conditions will be given by the following theorem.

Theorem 4.4.2. *Suppose the latent failure times* T_1 *and* T_2 *follow the mixed proportional hazards competing risk model under a given covariate* $\mathbf{X}$ *in the representation via a frailty approach 4.4.7,4.4.8, which are given by*

$$h_1(t_1|\beta_1,\mu_1,\delta,z) = g_1(\mathbf{X}\eta_{\mathbf{t_1}})\cdot h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)\cdot z$$

and

$$h_2(t_2|\beta_2,\mu_2,\delta,z) = g_2(\mathbf{X}\eta_{\mathbf{t_2}})\cdot h_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\cdot z$$

Associated with the frailty distribution which has the Laplace transform $\psi(s) = E[\exp(-sZ)]$ *and the following assumptions:*

1. $E(Z) < \infty$
2. *the Laplace transforms* ψ *and* ϕ *belong to the same class which satisfy with* $\phi^{-1\prime}(s)/\psi^{-1\prime}(s)$ *is a strictly monotone function of s*
3. *the baseline cumulative hazard functions are equal, that is,* $H_{0_1}(t_0|\beta_1,\eta_{0t_1},\delta) = H_{0_2}(t_0|\beta_2,\eta_{0t_2},\delta)$ *at some fixed time point* t_0

then the competing risk model 4.4.7 and 4.4.8 are identifiable through the observed data $T = \min(T_1, T_2)$, $C = \arg\min(T_1, T_2)$ *and the known covariate vector* $\mathbf{X}$. *Thus, the corresponding joint survival function in 4.4.11 is identifiable.*

Proof. In order to show the identifiability, it is necessary to obtain the crude survival functions of the model. Since it is known that the crude survival functions can be given by

$$\begin{aligned} Q_1(t) &= P(T_1 > t, T_1 < T_2) \\ &= -\int_t^\infty P(T_1 = u, T_1 < T_2) dS_1(u) \\ &= -\int_t^\infty P(T_1 = u, T_2 > T_1) dS_1(u) \\ &= -\int_t^\infty P(T_1 = u, T_2 > u) dS_1(u) \\ &= -\int_t^\infty \left. \frac{dS(t_1, t_2)}{dt_1} \right|_{t_1 = t_2 = u} du \end{aligned}$$

Similarly,

$$\begin{aligned} Q_2(t) &= P(T_2 > t, T_2 < T_1) \\ &= -\int_t^\infty \left. \frac{dS(t_1, t_2)}{dt_2} \right|_{t_1 = t_2 = u} du \end{aligned}$$

so the corresponding differential of crude survival functions of T_1 under the given covariate vector $\mathbf{X}$ can be written as

$$\begin{aligned} \frac{dQ_1(t|\mathbf{X} = \mathbf{X_1})}{dt} &= -\left. \frac{dS(t_1, t_2|\mathbf{X} = \mathbf{X_1})}{dt_1} \right|_{t_1 = t_2 = t} \\ &= -\psi' \left[g_1(\mathbf{X_1}\eta_{\mathbf{t_1}}) \cdot H_{0_1}(t_1|\beta_1, \eta_{0t_1}, \delta) + g_2(\mathbf{X_1}\eta_{\mathbf{t_2}}) \cdot H_{0_2}(t_2|\beta_2, \eta_{0c_2}, \delta) \right] \\ &\quad \times g_1(\mathbf{X_1}\eta_{\mathbf{t_1}}) h_{0_1}(t_1|\beta_1, \eta_{0t_1}, \delta) \end{aligned} \tag{4.4.14}$$

and

$$\begin{aligned}\frac{dQ_1(t|\mathbf{X}=\mathbf{X_0})}{dt} &= -\left.\frac{dS(t_1,t_2|\mathbf{X}=\mathbf{X_0})}{dt_1}\right|_{t_1=t_2=t} \\ &= -\psi'\left[g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(\mathbf{X_0}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\right] \\ &\quad\times g_1(\mathbf{X_0}\eta_{\mathbf{t}})h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) \qquad (4.4.15)\end{aligned}$$

Therefore, the ratio of two aforementioned equations are given by

$$\begin{aligned}\frac{dQ_1(t|\mathbf{X}=\mathbf{X_1})/dt}{dQ_1(t|\mathbf{X}=\mathbf{X_0})/dt} &= \frac{-\psi'\left[g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(\mathbf{X_1}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0c_2},\delta)\right]}{-\psi'\left[g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(\mathbf{X_0}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\right]} \\ &\quad\times\frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)}{g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)} \\ &= \frac{\psi'\left[g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(\mathbf{X_1}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\right]}{\psi'\left[g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(\mathbf{X_0}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\right]} \\ &\quad\times\frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})}{g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})} \qquad (4.4.16)\end{aligned}$$

Now, letting $t \longrightarrow 0$, then

$$\begin{aligned}H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) &= \exp(\eta_{0t_1})t_1^{\frac{\beta_1}{\delta}} \longrightarrow 0 \\ H_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta) &= \exp(\eta_{0t_2})t_2^{\frac{\beta_2}{\delta}} \longrightarrow 0\end{aligned}$$

so that the ratio 4.4.16 approaches to

$$\frac{dQ_1(t|\mathbf{X}=\mathbf{X_1})/dt}{dQ_1(t|\mathbf{X}=\mathbf{X_0})/dt} \longrightarrow \frac{\psi'(0)\cdot g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})}{\psi'(0)\cdot g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})} \longrightarrow \frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})}{g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})}$$

Since the function $g(\cdot)$ is given by $g_1(\mathbf{X}\eta_{\mathbf{t_1}}) = \exp(\mathbf{X}\eta_{\mathbf{t_1}})$ which is identified, so the ratio $\frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})}{g_1(\mathbf{X_0}\eta_{\mathbf{t_1}})}$ is identified as well. Similarly, $\frac{g_2(\mathbf{X_1}\eta_{\mathbf{t_1}})}{g_2(\mathbf{X_0}\eta_{\mathbf{t_1}})}$ can be identified. And it is known that

$$S_1(t_1|Z) = \exp(-\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta_1}{\delta}} Z) = \psi\left[g_1(\mathbf{X}\eta_{\mathbf{t_1}}) \cdot H_{0_1}(t_1|\beta_1, \eta_{0t_1}, \delta)\right] = S_1(t_1|\mathbf{X})$$

for any $\mathbf{X_1} \neq \mathbf{X_2}$; then it is easy to achieve

$$\psi^{-1}\left[S_1(t_1|\mathbf{X_1})\right] = g_1(\mathbf{X_1}\eta_{\mathbf{t_1}}) \cdot H_{1_0}(t_1|\beta_1, \eta_{0t_1}, \delta) \tag{4.4.17}$$

and

$$\psi^{-1}\left[S_1(t_1|\mathbf{X_2})\right] = g_1(\mathbf{X_2}\eta_{\mathbf{t_1}}) \cdot H_{1_0}(t_1|\beta_1, \eta_{0t_1}, \delta) \tag{4.4.18}$$

Since the model can be expressed via a frailty approach 3.3 as well as via a copula approach 3.4, the Laplace transform ψ can also be written as a copula ψ. In order to show 4.4.11 is identifiable by the observed data set $\min(T_1, T_2)$ and $argmin(T_1, T_2)$, a proof by contradiction will be introduced here. Suppose the model is not identifiable; that is, given a covariate vector $\mathbf{X}$ the ψ is not identifiable only based on the distribution of $\min(T_1, T_2)$. In other words, $S(t_1, t_2|\mathbf{X})$ is not identifiable. Then there exists another Laplace transform ϕ which is not the same as ψ belongs to the same class with ψ, such that $S^*(t_1, t_2|\mathbf{X}) = \phi\left[\phi^{-1}\big(S_1(t_1)\big) + \phi^{-1}\big(S_2(t_2)\big)\right]$ and $S^*(t_1, t_2|\mathbf{X}) = S(t_1, t_2|\mathbf{X}) = \psi\left[\psi^{-1}\big(S_1(t_1)\big) + \psi^{-1}\big(S_2(t_2)\big)\right]$ Similar to 4.4.17 and 4.4.18, for any $\mathbf{X_1} \neq \mathbf{X_2}$ there exists $S_1^*(t_1|\mathbf{X_1})$, $H_{0_1}^*(t_1|\beta_1, \eta_{0t_1}, \delta)$ and $S_1^*(t_1|\mathbf{X_2})$, $H_{0_1}^*(t_1|\beta_1, \eta_{0t_1}, \delta)$ such that

$$\phi^{-1}\left[S_1^*(t_1|\mathbf{X_1})\right] = g_1(\mathbf{X_1}\eta_{\mathbf{t_1}}) \cdot H_{0_1}^*(t_1|\beta_1, \eta_{0t_1}, \delta)$$

and

$$\phi^{-1}\left[S_1^*(t_1|\mathbf{X_2})\right] = g_1(\mathbf{X_2}\eta_{\mathbf{t_1}})\cdot H_{0_1}^*(t_1|\beta_1,\eta_{0t_1},\delta)$$

Since $S^*(t_1,t_2|\mathbf{X}) = S(t_1,t_2|\mathbf{X})$, the corresponding crude survival functions $Q_1(t) = Q_1^*(t)$ and $Q_2(t) = Q_2^*(t)$.

According to Rivest & Wells (2001), the formula of $S_1^*(t_1|\mathbf{X})$ is given by

$$S_1^*(t_1|\mathbf{X}) = \phi\left[\int_0^{t_1} \frac{\phi^{-1\prime}\big(\pi(u)\big)}{\psi^{-1\prime}\big(\pi(u)\big)} d\psi^{-1}\big(S_1(u|\mathbf{X})\big)\right]$$

After taking derivate of $\phi^{-1}\big(S_1^*(t_1|\mathbf{X})\big)$ with respect to t_1, we will get

$$\frac{d\phi^{-1}\big(S_1^*(t_1|\mathbf{X})\big)}{dt_1} = \frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X})\big)} d\psi^{-1}\big(S_1(t_1|\mathbf{X})\big)/dt_1$$

and then we can obtain for $\mathbf{X} = \mathbf{X_1}$ and $\mathbf{X} = \mathbf{X_2}$ respectively, so that

$$\frac{d\phi^{-1}\big(S_1^*(t_1|\mathbf{X_1})\big)}{dt_1} = \frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)} d\psi^{-1}\big(S_1(t_1|\mathbf{X_1})\big)/dt_1 \tag{4.4.19}$$

and

$$\frac{d\phi^{-1}\big(S_1^*(t_1|\mathbf{X_2})\big)}{dt_1} = \frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)} d\psi^{-1}\big(S_1(t_1|\mathbf{X_2})\big)/dt_1 \tag{4.4.20}$$

Therefore, the ratio of equation 4.4.19 and 4.4.20 is

$$\begin{aligned}
\frac{4.4.19}{4.4.20} &= \frac{d\phi^{-1}\big(S_1^*(t_1|\mathbf{X_1})\big)/dt_1}{d\phi^{-1}\big(S_1^*(t_1|\mathbf{X_2})\big)/dt_1} \\
&= \frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)/\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)}{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)/\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)} \cdot \frac{d\psi^{-1}\big(S_1(t_1|\mathbf{X_1})\big)/dt_1}{d\psi^{-1}\big(S_1(t_1|\mathbf{X_2})\big)/dt_1} \\
&= \frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)\cdot\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)\cdot\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)} \cdot \frac{d\psi^{-1}\big(S_1(t_1|\mathbf{X_1})\big)/dt_1}{d\psi^{-1}\big(S_1(t_1|\mathbf{X_2})\big)/dt_1}
\end{aligned} \tag{4.4.21}$$

Since we know the fact that

$$\frac{d\psi^{-1}\left[S_1(t_1|\mathbf{X_1})\right]/dt_1}{d\psi^{-1}\left[S_1(t_1|\mathbf{X_2})\right]/dt_1} = \frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})\cdot h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)}{g_1(\mathbf{X_2}\eta_{\mathbf{t_1}})\cdot h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)} = \frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})}{g_1(\mathbf{X_2}\eta_{\mathbf{t_1}})} \tag{4.4.22}$$

and

$$\frac{d\phi^{-1}\left[S_1^*(t_1|\mathbf{X_1})\right]/dt_1}{d\phi^{-1}\left[S_1^*(t_1|\mathbf{X_2})\right]/dt_1} = \frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})\cdot h_{0_1}^*(t_1|\beta_1,\eta_{0t_1},\delta)}{g_1(\mathbf{X_2}\eta_{\mathbf{t_1}})\cdot h_{0_1}^*(t_1|\beta_1,\eta_{0t},\delta)} = \frac{g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})}{g_1(\mathbf{X_2}\eta_{\mathbf{t_1}})} \tag{4.4.23}$$

then from equations 4.4.22 and 4.4.23 we can obtain

$$\frac{d\psi^{-1}\left[S_1(t_1|\mathbf{X_1})\right]/dt_1}{d\psi^{-1}\left[S_1(t_1|\mathbf{X_2})\right]/dt_1} = \frac{d\phi^{-1}\left[S_1(t_1|\mathbf{X_1})\right]/dt_1}{d\phi^{-1}\left[S_1(t_1|\mathbf{X_2})\right]/dt_1} \tag{4.4.24}$$

Comparing equations 4.4.21 and 4.4.24, we obtain the equation below:

$$\frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)\cdot\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)\cdot\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)} = 1$$

and also obtain the transformed equation:

$$\frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_1})\big)} = \frac{\phi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)}{\psi^{-1\prime}\big(\pi(t_1|\mathbf{X_2})\big)} \tag{4.4.25}$$

From the assumption 2 in **Theorem 4.4.2** we know that $\phi^{-1\prime}(s)/\psi^{-1\prime}(s)$ is a strictly monotone function of s; in order to make 4.4.25 true, we only need $\pi(t_1|\mathbf{X_1}) = \pi(t_1|\mathbf{X_2})$, which is

$$\begin{aligned}&\psi\Big[g_1(\mathbf{X_1}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) + g_2(\mathbf{X_1}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta)\Big]\\ =\ &\psi\Big[g_1(\mathbf{X_2}\eta_{\mathbf{t_1}})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) + g_2(\mathbf{X_2}\eta_{\mathbf{t_2}})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta)\Big]\end{aligned} \tag{4.4.26}$$

Case1: When $\mathbf{X_i}$ is a scalar, suppose there exists x_3 such that $x_1 < x_3 < x_2$, and $x_3\eta_t$ is within the interval $(x_1\eta_{t_1},\ x_2\eta_{t_1})$ or $(x_2\eta_{t_1},\ x_1\eta_{t_1})$ where $\eta_{t_1} \neq 0$. Similarly, $x_3\eta_{t_2}$ is within the interval

$(x_1\eta_{t_2},\ x_2\eta_{t_2})$ or $(x_2\eta_{t_2},\ x_1\eta_{t_2})$. Since ψ is a monotone function and x_1, x_2 are arbitrary, equivalently, we can obtain the following equation from condition 4.4.26:

$$\begin{aligned} & g_1(x_1\eta_{t_1})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_1\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \\ =\ & g_1(x_3\eta_{t_1})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_3\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \end{aligned} \tag{4.4.27}$$

And since $x_3 \neq x_1$ and $x_3 \neq x_2$, similarly, we have

$$\begin{aligned} & g_1(x_2\eta_{t_1})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_2\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \\ =\ & g_1(x_3\eta_{t_1})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_3\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \end{aligned} \tag{4.4.28}$$

Combining equations 4.4.27 and 4.4.28 we can obtain the new equation:

$$\begin{aligned} & 2\Big[g_1(x_3\eta_{t_1})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_3\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta)\Big] \\ =\ & g_1(x_1\eta_t)\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_1\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \\ & +g_1(x_2\eta_{t_1})\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+g_2(x_2\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \end{aligned}$$

Moreover, take out the common factor from the right side to yield

$$\begin{aligned} & 2g_1(x_3\eta_t)\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)+2g_2(x_3\eta_{t_1})\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \\ =\ & \Big[g_1(x_1\eta_{t_1})+g_1(x_2\eta_{t_1})\Big]\cdot H_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta) \\ & +\Big[g_2(x_1\eta_{t_1})+g_2(x_2\eta_{t_1})\Big]\cdot H_{0_2}(t_1|\beta_2,\eta_{0t_2},\delta) \end{aligned}$$

By the assumption 3 in **Theorem 4.4.2**, $H_{0_1}(t_0|\beta_1,\eta_{0t_1},\delta) = H_{0_2}(t_0|\beta_2,\eta_{0t_2},\delta)$ at some fixed point t_0, so the aforementioned equation can be rewritten as

$$2g_1(x_3\eta_{t_1}) + 2g_2(x_3\eta_{t_1}) = \Big[g_1(x_1\eta_{t_1}) + g_1(x_2\eta_{t_1})\Big] + \Big[g_2(x_1\eta_{t_1}) + g_2(x_2\eta_{t_1})\Big] \qquad (4.4.29)$$

Since $g_1(x\eta_{t_1}) = \exp(x\eta_{t_1})$ and $g_2(x\eta_{t_2}) = \exp(x\eta_{t_2})$, which are strictly convex function, and $x_1 < x_3 < x_2$, so

$$2g_1(x_3\eta_{t_1}) < g_1(x_1\eta_{t_1}) + g_1(x_2\eta_{t_1})$$

and

$$2g_2(x_3\eta_{t_1}) < g_2(x_1\eta_{t_1}) + g_2(x_2\eta_{t_1})$$

Therefore, the sum of these two inequalities can be given by

$$2g_1(x_3\eta_{t_1}) + 2g_2(x_3\eta_{t_1}) < \Big[g_1(x_1\eta_{t_1}) + g_1(x_2\eta_{t_1})\Big] + \Big[g_2(x_1\eta_{t_1}) + g_2(x_2\eta_{t_1})\Big]$$

which is a contradiction compared with the result in 4.4.29.

Case2: When $\mathbf{X_i}$ is not a scalar, $\mathbf{X_i} = (x_{i1}, x_{i2}, ..., x_{ij}..., x_{ik})$, which is a covariate vector with k components, and $\eta_{\mathbf{t_1}} = (\eta_{t_1 1}, \eta_{t_1 2}, ...\eta_{t_1 j}, ..., \eta_{t_1 k})$ and $\eta_{\mathbf{t_2}} = (\eta_{t_2 1}, \eta_{t_2 2}, ...\eta_{t_2 j}, ..., \eta_{t_2 k})$. We can have three covariate vectors with choosing one of the components to be different and others to be the same, that is,

$$\mathbf{X_1} = (x_{11}, x_{12}, ..., x_{1j}, ..., x_{1k}) = (x_1, x_2, ..., x_{1j}, ..., x_k)$$

$$\mathbf{X_2} = (x_{21}, x_{22}, ..., x_{2j}, ..., x_{2k}) = (x_1, x_2, ..., x_{2j}, ..., x_k)$$

$$\mathbf{X_3} = (x_{31}, x_{32}, ..., x_{3j}, ..., x_{3k}) = (x_1, x_2, ..., x_{3j}, ..., x_k)$$

And let's suppose $x_{1j} < x_{3j} < x_{2j}$; then $\mathbf{X_3}\eta_{\mathbf{t_1}}$ is within the interval

$$(\mathbf{X_1}\eta_{\mathbf{t_1}}, \mathbf{X_2}\eta_{\mathbf{t_1}}) \quad \text{or} \quad (\mathbf{X_2}\eta_{\mathbf{t_1}}, \mathbf{X_1}\eta_{\mathbf{t_1}})$$

where $\eta_{\mathbf{t_1}} \neq 0$ and also $\eta_{t_1 j} \neq 0$. Similarly, we have $\mathbf{X_3}\eta_{\mathbf{t_2}}$ is within the interval

$$(\mathbf{X_1}\eta_{\mathbf{t_2}}, \mathbf{X_2}\eta_{\mathbf{t_2}}) \quad \textit{or} \quad (\mathbf{X_2}\eta_{\mathbf{t_2}}, \mathbf{X_1}\eta_{\mathbf{t_2}})$$

where $\eta_{\mathbf{t_2}} \neq 0$ and $\eta_{t_2 j} \neq 0$. Then we can reach the same conclusion under Case 1, which leads to a contradiction.

Therefore, equation 4.4.26 cannot hold, so that the copula ψ can be uniquely determined by the distribution of $min(T_1, T_2)$. Furthermore, model 4.4.11 $S(t_1, t_2|\mathbf{X}) = \psi\Big[\psi^{-1}\big(S_1(t_1|\mathbf{X})\big) + \psi^{-1}\big(S_2(t_2|\mathbf{X})\big)\Big]$ is identifiable. □

4.5 Remarks

Heckman & Honoré (1989) have explored a general class of models:

$$S(t_1, t_2|\mathbf{X}) = \mathbf{K}[S_1(t_1|\mathbf{X}), S_2(t_2|\mathbf{X})] \tag{4.5.1}$$

where $\mathbf{K}$ is the joint cumulative distribution function on $[0,1]^2$. To identify this model, the support of $(g_1(\mathbf{X}\eta_{\mathbf{t_1}}), g_2(\mathbf{X}\eta_{\mathbf{t_2}}))$ needs to be $(0,\infty) \times (0,\infty)$. Further, Abbring & Van den Berg (2003) have studied a more specific model that is a mixed proportional hazards competing risks survival model. Such a model is related to model 4.5.1 if $\mathbf{K}(u,v) = \int_0^\infty \int_0^\infty u^{t_1} v^{t_2} dG(t_1, t_2)$. Sufficient conditions for identifiability of this model include that the support of $(g_1(\mathbf{X}\eta_{\mathbf{t_1}}), g_2(\mathbf{X}\eta_{\mathbf{t_2}}))$ is a non-empty open set. In contrast to the semi-parametric models considered in these two articles, we established

identifiability conditions for the proposed parametric model under a regression framework. Rivest & Wells (2001) have also marked a similar condition as assumption 3 in **Theorem 4.4.2**.

CHAPTER 5

NUMERICAL STUDIES

5.1 Introduction

In this section, we report on results of simulation studies assessing the performance of Bayesian and maximum likelihood (MLE) estimators. The Bayesian estimates are approximated by Markov chain sampling analysis of the proposed models. The MLEs are obtained by numerical optimization of the likelihood function or, in some cases, of the profile likelihood. We conclude this section with an application of the proposed model on real data from a Tamoxifen trial, reported in Pintilie (2006). We consider analysis of the Tamoxifen data under the dependent competing risks model as well as under a regression framework. See Ibrahim et al. (2005) for Bayesian methods in time-to-event data and Zacks (1971) for maximum likelihood estimation.

5.2 Data Generating Model

The simulation studies primarily consider the case of two competing risks and consider different sample sizes of $n = 100$, $n = 500$ and $n = 1000$, respectively. Each generated data set contains $\{(t_i, d_{ji}), i = 1, ..., n;\ j = 1, 2\}$ with 10 iterations, which is denoted by $\{(t_{ik}, d_{jik}), i = 1, ..., n;\ k = 1, ..., 10,\ j = 1, 2\}$, where $t_{ik} = \min(t_{1ik}, t_{2ik})$ and n is the sample size.

The data for the independent case are generated by the following steps: Let $T_{1ik} \sim \text{Weibull}(\mu_1, \beta_1)$ and $T_{2ik} \sim \text{Weibull}(\mu_2, \beta_2)$, and set $T_{ik} = \min(T_{1ik}, T_{2ik})$; the indicator d_{jik} is defined as

$$\begin{aligned} d_{1ik} &= \begin{cases} 1 & \text{if } t_{1ik} \leq t_{2ik} \\ 0 & \text{otherwise} \end{cases} \\ d_{2ik} &= 1 - d_{1ik} \end{aligned}$$

We consider two representations for data generation in the dependent cases. The first method is introduced by L. Lee (1979) through the following steps: Let $U \sim \text{Unif}(0,1)$ and $V = V_{11} + M_\delta V_{12}$ be independent, where $M_\delta \sim \text{Bern}(\delta)$, $V_{11} \sim \text{Exp}(1)$ and $V_{12} \sim \text{Exp}(1)$ are independent.

Lemma 5.2.1. *Suppose the random variables T_1 and T_2 have joint distribution $S(t_1,t_2) = \exp\big(-\big[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta 2})^{1/\delta}\big]^\delta\big)$. Then they can be represented as $T_1 = (U)^{\delta/\beta_1}(V/\mu_1)^{1/\beta_1}$, $T_2 = ([(1-U)]^{\delta/\beta_2}(V/\mu_2)^{1/\beta_2}$, where $U \sim Unif(0,1)$, $M_\delta \sim Bern(\delta)$, $V_{11} \sim Exp(1)$ and $V_{12} \sim Exp(1)$ are independent, and $V = V_{11} + M_\delta V_{12}$.*

Proof. Consider the random variables $X_1 = (\mu_1 T_1^{\beta_1})^{1/\delta}$ and $X_2 = (\mu_2 T_2^{\beta_2})^{1/\delta}$; then the joint distribution of them is given by

$$S(x_1,x_2) = \exp[-(x_1+x_2)^\delta] \tag{5.2.1}$$

By taking derivative we can obtain the joint density function, which can be written as

$$f(x_1,x_2) = \big[(1-\delta)\delta(x_1+x_2)^{\delta-2} + \delta^2(x_1+x_2)^{2\delta-2}\big] \cdot \exp[-(x_1+x_2)^\delta] \tag{5.2.2}$$

Since $T_1 = (U)^{\delta/\beta_1}(V/\mu_1)^{1/\beta_1}$,$T_2 = ([(1-U)]^{\delta/\beta_2}(V/\mu_2)^{1/\beta_2}$, solving for U and V by replacing $T_1,\ T_2$ with $X_1,\ X_2$, then

$$\begin{aligned} U &= \frac{(\mu_1 T_1^{\beta_1})^{1/\delta}}{(\mu_1 T_1^{\beta_1})^{1/\delta} + (\mu_2 T_2^{\beta_2})^{1/\delta}} = \frac{X_1}{X_1 + X_2}, \\ V &= \left((\mu_1 T_1^{\beta_1})^{1/\delta} + (\mu_2 T_2^{\beta_2})^{1/\delta}\right)^{\delta} = (X_1 + X_2)^{\delta} \end{aligned}$$

having the Jacobian $J = \delta(X_1 + X_2)^{\delta-2}$.

Because $V = V_{11} + M_\delta V_{12}$ where $M_\delta \sim \text{Bern}(\delta)$, $V_{11} \sim \text{Exp}(1)$ and $V_{12} \sim \text{Exp}(1)$ are independent, then

$$V = \begin{cases} V_{11} & \text{if } M = 0 \\ V_{11} + V_{12} & \text{if } M = 1 \end{cases}$$

with $f(V|M=0) \sim \text{Exp}(1) = \exp(-v)$ and $f(V|M=1) \sim \text{Gamma}(2,1) = v\exp(-v)$.

Thus, $f(v) = f(V|M=0)\cdot P(M=0) + f(V|M=1)\cdot P(M=1) = (1-\delta)\exp(-v) + \delta v\exp(-v)$, combining with independent $U \sim \text{Unif}(0,1)$, then the joint distribution of V and U is given by

$$\begin{aligned} f(u,v) &= f(u)\cdot f(v) \\ &= (1-\delta)\exp(-v) + \delta v\exp(-v) \end{aligned} \tag{5.2.3}$$

The joint density of X_1 and X_2 can be written as

$$\begin{aligned} f(x_1,x_2) &= \left[(1-\delta)\exp[-(x_1+x_2)^{\delta}] + \delta(x_1+x_2)^{\delta}\exp[-(x_1+x_2)^{\delta}]\right]\cdot\delta(x_1+x_2)^{\delta-2} \\ &= \left[(1-\delta)\delta(x_1+x_2)^{\delta-2} + \delta^2(x_1+x_2)^{2\delta-2}\right]\cdot\exp[-(x_1+x_2)^{\delta}] \end{aligned} \tag{5.2.4}$$

which is the same as 5.2.2. □

An alternative, second representation of (T_1, T_2) from the proposed model is as follows:

$$\begin{aligned} T_1 &= (U)^{\delta/\beta_1}(V/\mu_1)^{1/\beta_1} \\ T_2 &= (1-U)^{\delta/\beta_2}(V/\mu_2)^{1/\beta_2} \end{aligned}$$

where $U \sim \text{Unif}(0,1)$ and $V = (1-M_\delta)V_{11} + M_\delta V_{12}$ are independent, $M_\delta \sim \text{Bern}(\delta)$, $V_{11} \sim \text{Exp}(1)$ and $V_{12} \sim \text{Gamma}(2,1)$ are independent.

Lemma 5.2.2. *Suppose the random variables T_1 and T_2 have joint distribution $S(t_1,t_2) = \exp\big(-\big[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta 2})^{1/\delta}\big]^{\delta}\big)$; then it can be represented in terms of independent random variables: $T_1 = (U)^{\delta/\beta_1}(V/\mu_1)^{1/\beta_1}, T_2 = ([(1-U)]^{\delta/\beta_2}(V/\mu_2)^{1/\beta_2}$, where $U \sim Unif(0,1)$ and $V = (1-M_\delta)V_{11} + M_\delta V_{12}$ are independent, $M_\delta \sim Bern(\delta)$, $V_{11} \sim Exp(1)$ and $V_{12} \sim Gamma(2,1)$ are independent.*

Proof. Similar to the proof in Lemma 5.2.1, the joint density of X_1 and X_2 is given by equation 5.2.2. Since $V = (1-M_\delta)V_{11} + M_\delta V_{12}$ where $M_\delta \sim \text{Bern}(\delta)$, $V_{11} \sim \text{Exp}(1)$ and $V_{12} \sim \text{Gamma}(2,1)$ are independent, then

$$V = \begin{cases} V_{11} & \text{if } M = 0 \\ V_{12} & \text{if } M = 1 \end{cases}$$

with $f(V|M=0) \sim \text{Exp}(1) = \exp(-v)$ and $f(V|M=1) \sim \text{Gamma}(2,1) = v\exp(-v)$. Thus, $f(v) = f(V|M=0)\cdot P(M=0) + f(V|M=1)\cdot P(M=1) = (1-\delta)\exp(-v) + \delta v\exp(-v)$, and combining with independent $U \sim \text{Unif}(0,1)$, then the joint distribution of V and U is given by

$$\begin{aligned} f(u,v) &= f(u)\cdot f(v) \\ &= (1-\delta)\exp(-v) + \delta v\exp(-v) \end{aligned} \tag{5.2.5}$$

The joint of X_1 and X_2 can be written as

$$\begin{aligned} f(x_1,x_2) &= \left[(1-\delta)\exp[-(x_1+x_2)^{\delta}]+\delta(x_1+x_2)^{\delta}\exp[-(x_1+x_2)^{\delta}]\right]\cdot\delta(x_1+x_2)^{\delta-2} \\ &= \left[(1-\delta)\delta(x_1+x_2)^{\delta-2}+\delta^2(x_1+x_2)^{2\delta-2}\right]\cdot\exp[-(x_1+x_2)^{\delta}] \end{aligned} \tag{5.2.6}$$

which is the same as 5.2.2. □

5.3 Simulation Study I

5.3.1 Bayesian Inference

Bayesian analysis of the proposed competing risks model is performed using Markov chain sampling. S. Basu et al. (2003) considered Markov chain sampling for Bayesian analysis of competing risks, whereas the book by Ibrahim et al. (2005) discusses Bayesian methods in time-to-event data.

Bayesian inference usually involves specification of appropriate prior distributions for the parameters involved in the model and derivation of the posterior distributions for the parameters conditional on the data. The Bayesian estimate of a parameter is usually the mean or median of the posterior distribution. As mentioned earlier, for the i^{th} subject, the observed data is denoted by (t_i, d_{1i}, d_{2i}), where t_i is the observed time to event, d_{1i} is the censoring indicator of risk 1 taking the value 0 if censored and 1 if event occurs from risk1, and d_{2i} is the censoring indicator of risk2 taking the value 0 if censored and 1 if event occurs from risk2. Let $\mathbf{D} = \{(t_i, d_{ji}), i = 1, ..., \ j = 1,2\}$. The parameters specified by the model can be collectively written as $\Omega = (\mu_1, \mu_2, \beta_1, \beta_2, \delta)$. Then, the joint probability function of the data given the parameters Ω can be written as

$$P(\mathbf{D}|\Omega) = \prod_{i=1}^{n}\left([f_1(t_i)]^{d_{1i}}\cdot[f_2(t_i)]^{d_{2i}}\cdot[S(t_1,t_2)]^{1-d_{1i}-d_{2i}}\right)$$

where $f_1(t_1;\mu_1,\beta_1)$ and $f_2(t_2;\mu_2,\beta_2)$ are the sub-density functions, $S(t_1,t_2;\mu_1,\mu_2,\beta_1,\beta_2,\delta)$ is the joint survival function. Thus, the posterior distribution of Ω can be formulated by

$$P(\Omega|\mathbf{D}) \propto P(\mathbf{D}|\Omega)\times\pi(\Omega) \tag{5.3.1}$$

where $\pi(\Omega)$ is the joint prior probability distribution of the parameters that are involved in the model.

The prior distribution of each parameter is given by

$$\begin{aligned}
\mu_1 &\sim \text{Gamma}(1,1)\\
\mu_1 &\sim \text{Gamma}(1,1)\\
\beta_1 &\sim \text{Unif}(0.5,2.5)\\
\beta_2 &\sim \text{Unif}(0.5,2.5)\\
\delta &\sim \pi\cdot\text{Unif}(0,1)+(1-\pi)\cdot\text{degenerate}(1)\\
\pi &\sim \text{Unif}(0,1)
\end{aligned}$$

We consider 10 replicated data sets generated from

$$S(t_1,t_2)=\exp\big(-\big[(\mu_1 t_1^{\beta_1})^{1/\delta}+(\mu_2 t_2^{\beta 2})^{1/\delta}\big]^{\delta}\big) \tag{5.3.2}$$

In the data generating model, $\mu_1=0.8,\ \mu_2=0.5,\ \beta_1=1.1,\ \beta_2=2$. We further considered the independent case ($\delta=1$) and dependent cases ($\delta=0.8;\ \delta=0.3$) in the simulation study.The sample

size of each study contains 1000 observations. The Bayesian inference results reported in Table 5.1 are the posterior mean of the parameters based on 10,000 Markov chain samples after a burn-in of 5,000. We consider 10 replicated data sets, and the results reported are the median, minimum and maximum over these 10 replications. According to the Table 5.1, most of the estimations are close to their true value in the data generating model. The independent case and dependent case can be suggested in right direction.

Table 5.1: Simulation Results of Bayesian Estimation

Analysis	Data Generating Model					
Model	$\delta = 1$		$\delta = 0.8$		$\delta = 0.3$	
Parameter	Estimated Values: median (min,max)					
μ_1	0.811	(0.784, 0.839)	0.782	(0.724, 0.845)	0.790	(0.760, 0.837)
μ_2	0.527	(0.495, 0.619)	0.439	(0.385, 0.576)	0.505	(0.341, 0.635)
β_1	1.12	(1.057, 1.192)	1.099	(1.018, 1.134)	1.096	(1.043, 1.235)
β_2	1.978	(1.89, 2.056)	2.128	(1.983, 2.185)	2.007	(1.718, 2.325)
δ	0.965	(0.866, 0.981)	0.906	(0.766, 0.991)	0.319	(0.164, 0.447)
$P(\delta = 1)$	0.792	(0.574, 0.862)	0.651	(0.379, 0.906)	0	(0, 0)

Note: The estimation of each parameter is given by the minimum, the maximum and the median of 10 replications.

The posterior density of δ for each study is shown in Figure 5.1. According to the figure, for data generating model with $\delta = 1$ and $\delta = 0.8$, most of the iterations obtain a flat likelihood over δ such that the obtained Bayesian estimation of dependence parameter may not be close to the true value. For the data generating model with $\delta = 0.3$, this situation is not significant.

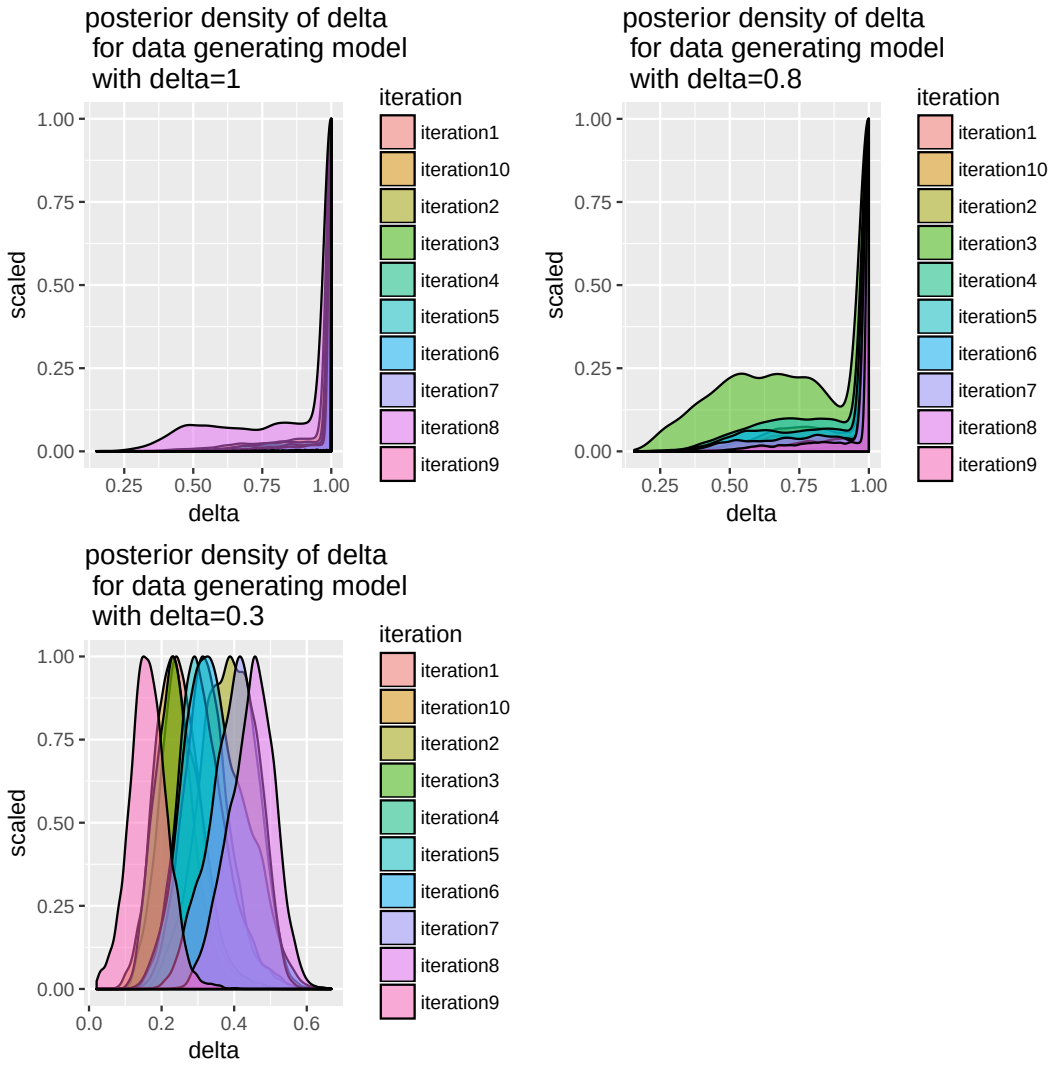


Figure 5.1: The plot of posterior density of δ for data generating model with $\delta = 1, 0.8, 0.3$

5.3.2 Maximum Likelihood Estimator

Let $f_1(t_1;\mu_1,\beta_1)$ and $f_2(t_2;\mu_2,\beta_2)$ and $S(t_1,t_2;\mu_1,\mu_2,\beta_1,\beta_2,\delta)$ denote the sub-density functions of the joint survival function of latent failure times $T_1,\ T_2$. The log-likelihood can then be written as

$$\begin{aligned}
\log L_T(\mu_1,\mu_2,\beta_1,\beta_2,\delta) &= \sum_{i=1}^{n} \log L_i(\mu_1,\mu_2,\beta_1,\beta_2,\delta) \\
&= \sum_{i=1}^{n} \log \left([f_1(t_i)]^{d_{1i}} \cdot [f_2(t_i)]^{d_{2i}} \cdot [S(t_1,t_2)]^{1-d_{1i}-d_{2i}} \right) \\
&= \sum_{i=1}^{n} \left(d_{1i} \log f_1(t_i) + d_{2i} \log f_2(t_i) + (1-d_{1i}-d_{2i}) \log S(t_1,t_2) \right)
\end{aligned}$$

The maximum likelihood estimates are given by

$$(\hat{\mu_1},\hat{\mu_2},\hat{\beta_1},\hat{\beta_2},\hat{\delta}) = \arg \max_{\mu_1,\mu_2,\beta_1,\beta_2,\delta} \log L_T(\mu_1,\mu_2,\beta_1,\beta_2,\delta)$$

We in fact consider a profile likelihood approach by first fixing the dependence parameter δ and maximizing over the remaining parameters, followed by maximization over δ. In particular, for given δ, let $(\hat{\mu_1}(\delta),\hat{\mu_2}(\delta),\hat{\beta_1}(\delta),\hat{\beta_2}(\delta)) = \arg\max_{\mu_1,\mu_2,\beta_1,\beta_2} \log L_T(\mu_1,\mu_2,\beta_1,\beta_2,\delta)$. We then maximize the profile likelihood of δ to obtain

$$(\hat{\mu_1},\hat{\mu_2},\hat{\beta_1},\hat{\beta_2},\hat{\delta}) = \arg\max_{\delta} \log L_T(\hat{\mu_1}(\delta),\hat{\mu_2}(\delta),\hat{\beta_1}(\delta),\hat{\beta_2}(\delta),\delta)$$

The results of the simulation studies based on 10 replications are shown in the following subsections.

5.3.2.1 Independent Case

The analysis results are based on data generating model with $\mu_1 = 0.8$, $\mu_2 = 0.5$, $\beta_1 = 1.1$ and $\beta_2 = 2$, summarized in Table 5.2

Table 5.2: Simulation Results Under Independent Case

Analysis Model	Data Generating Model with $\delta = 1$		
	sample size $n = 100$	sample size $n = 500$	sample size $n = 1000$
Parameter	Estimated Values: median (min,max)		
μ_1	0.860 (0.742, 1.160)	0.781 (0.725, 1.084)	0.847 (0.802, 1.128)
μ_2	0.501 (0.389, 1.111)	0.549 (0.471, 1.018)	0.517 (0.471, 1.166)
β_1	1.137 (0.942, 1.380)	1.116 (1.004, 1.377)	1.125 (1.035, 1.239)
β_2	2.205 (1.813, 2.558)	1.915 (1.534, 2.055)	1.954 (1.440, 2.137)
δ	1 (0.28, 1)	0.99 (0.25, 1)	1 (0.18, 1)
-log(likelihood)	114.3	615.9	1229.3

Note: The estimator of each parameter is given by the minimum, the maximum and the median over 10 iterations.

The profile log-likelihood estimators are given in Table 5.2. We note that most of the estimates are close to their true data-generating values except for the case of δ. The main reason for such results is that the profile likelihood over δ is flat, which makes the estimation to be hard. The plots of δ over profile log-likelihood at different sample sizes are give in Figure 5.2

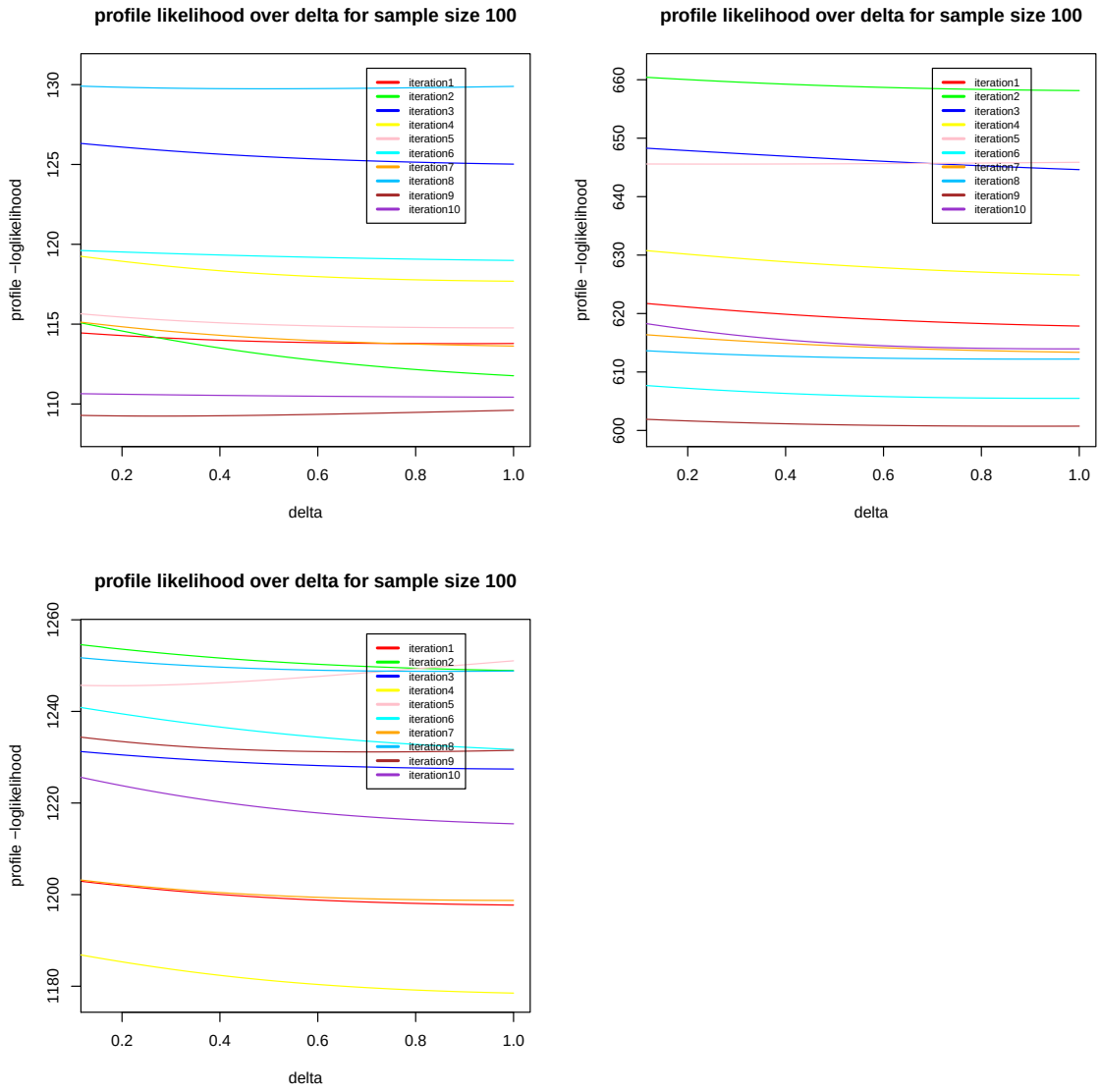


Figure 5.2: The plot of profile likelihood over delta when true delta value is 1

5.3.2.2 Dependent Case

The analysis results for the data generating model with $\mu_1 = 0.8$, $\mu_2 = 0.5$, $\beta_1 = 1.1$ and $\beta_2 = 2$, summarized in Table 5.3 and Table 5.4

1. For $\delta = 0.8$, the results are summarized in Table5.3.

Table 5.3: Simulation Results Under Dependent Case with Dependent Parameter Equal to 0.8

Analysis Model	Data Generating Model with $\delta = 0.8$		
	sample size $n = 100$	sample size $n = 500$	sample size $n = 1000$
Parameter	Estimated Values: median (min,max)		
μ_1	0.805 (0.595, 1.155)	0.771 (0.726, 0.937)	0.787 (0.753, 0.935)
μ_2	0.587 (0.341, 1.136)	0.508 (0.355, 0.841)	0.491 (0.359, 0.729)
β_1	1.058 (0.938, 1.325)	1.106 (0.992, 1.258)	1.131 (1.002, 1.205)
β_2	1.902 (1.428, 2.694)	2.007 (1.698, 2.171)	2.035 (1.759, 2.293)
δ	0.915 (0.13,1)	0.805 (0.36,1)	0.83 (0.47,1)
-log(likelihood)	128.0	651.9	1285.7

Note: The estimation of each parameter is given by the minimum, the maximum and the median over 10 iterations.

The profile log-likelihood estimators are given in the Table 5.3. The estimations of all the parameters μ_1, μ_2, β_1 and β_2 are closer to the true values. Comparatively speaking, the larger sample size data will give the more accurate maximum likelihood estimators based on the analysis model. One reason could be that the profile likelihood value over δ is more flat with small sample size data according to the plot of profile log-likelihood over δ at different sample sizes, which is given in Figure 5.3. So it is hard to obtain an accurate estimate for most of the iteration studies.

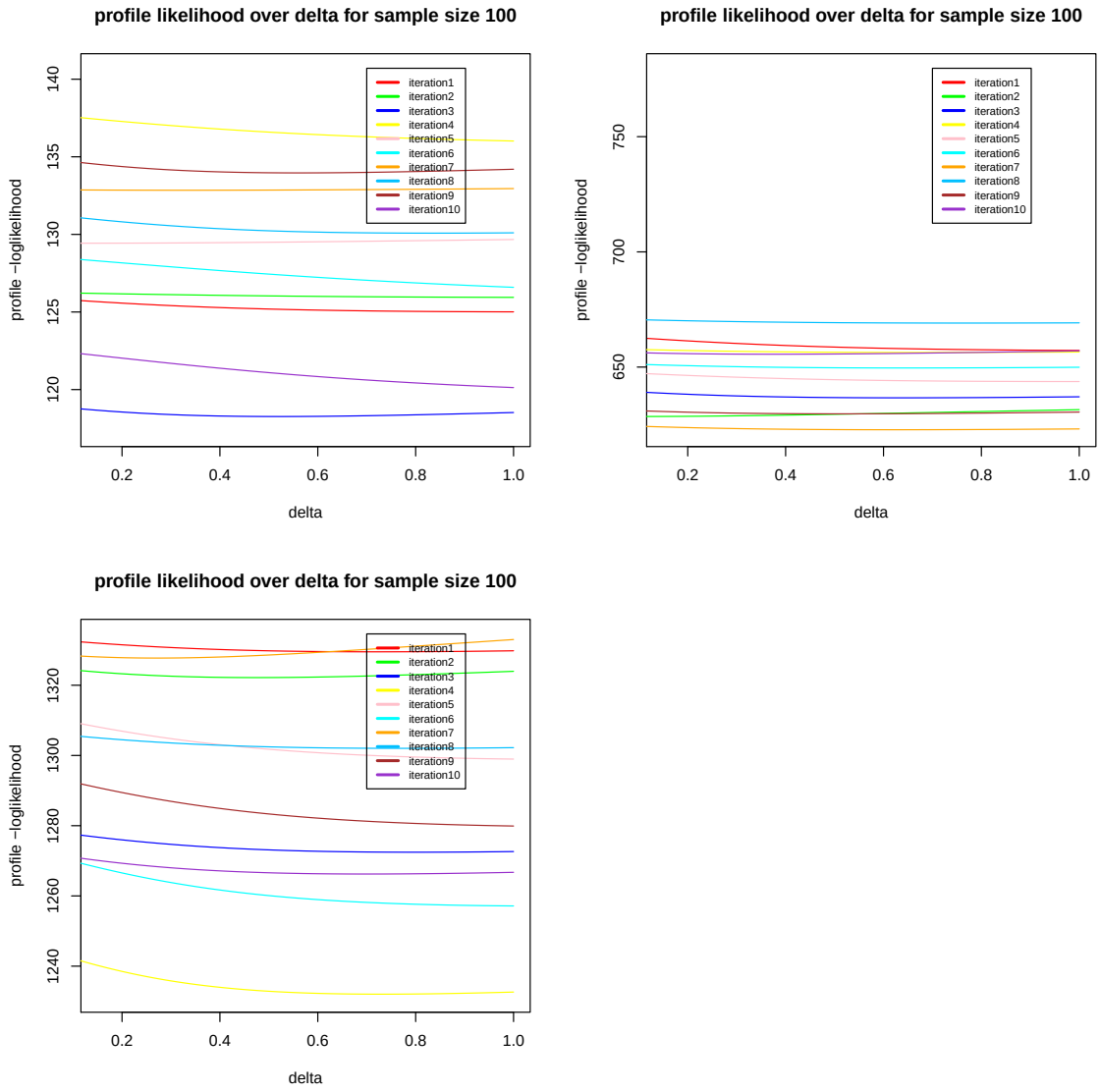


Figure 5.3: The plot of profile likelihood over delta when true delta value is 0.8

Next, we simulate the failure times for the dependent case with $\delta = 0.3$.

2. For $\delta = 0.3$, the results are summarized in Table 5.4.

Table 5.4: Simulation Results Under Dependent Case with Dependent Parameter Equal to 0.3

Analysis	Data Generating Model with $\delta = 0.3$		
Model	sample size $n = 100$	sample size $n = 500$	sample size $n = 1000$
Parameter	Estimated Values: median (min,max)		
μ_1	0.856 (0.596, 1.034)	0.816 (0.745, 0.874)	0.815 (0.722, 0.847)
μ_2	0.508 (0.071, 0.988)	0.437 (0.306, 0.648)	0.514 (0.425, 0.669)
β_1	1.115 (1.028, 1.296)	1.068 (1.000, 1.144)	1.141 (1.044, 1.189)
β_2	1.977 (1.311, 3.456)	2.194 (1.736, 2.409)	1.989 (1.693, 2.202)
δ	0.38 (0.06,1)	0.36 (0.17,0.55)	0.27 (0.17,0.36)
-log(likelihood)	125.7	634.7	1284.9

Note: The estimation of each parameter is given by the minimum, the maximum and the median over 10 iterations.

The profile log-likelihood estimators are given in the Table 5.4. The estimations of all the parameters μ_1, μ_2, β_1 and β_2 are closer to the true values. In particular, along with the growth of sample size, the profile maximum likelihood estimators are closer to the true values in data generating model. The main reason for such results is that the profile likelihood value over δ is more flat with small sample size data. It is hard to obtain an accurate estimator for most of the iteration studies. The plots of profile log-likelihood over δ at different sample sizes are given in Figure5.4.

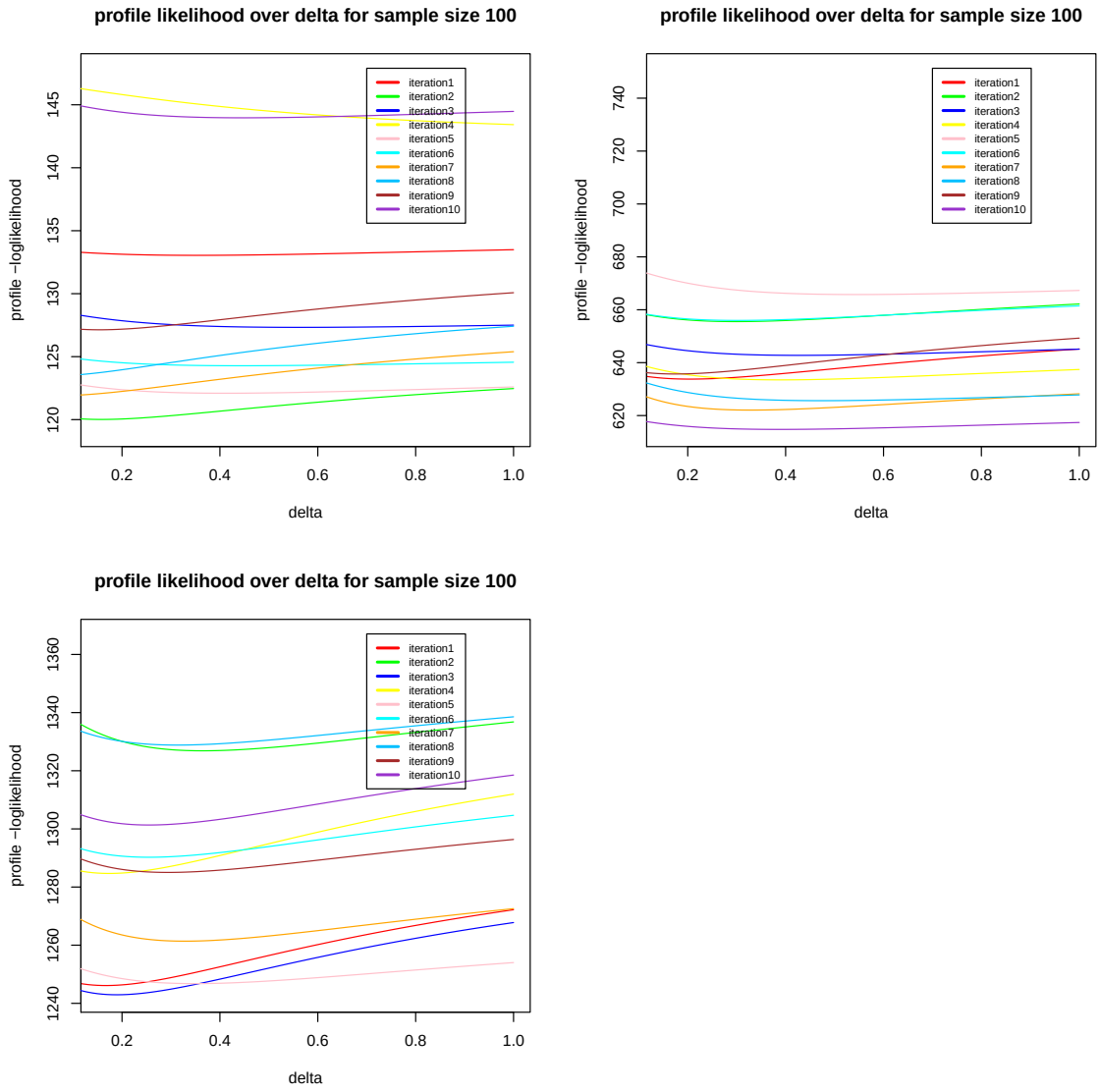


Figure 5.4: The plot of profile likelihood over delta when true delta value is 0.3

The summary of simulation study based on sample size $n = 1000$ with different delta values for profile maximum likelihood estimator is given in the Table 5.5.

Table 5.5: Simulation Results Under Sample Size 1000

Analysis Model	Data Generating Model with Sample Size $N = 1000$, $\mu_1 = 0.8, \mu_2 = 0.5, \beta_1 = 1.1, \beta_2 = 2$				
	$\delta = 0.1$	$\delta = 0.2$	$\delta = 0.3$	$\delta = 0.4$	$\delta = 0.5$
Parameter	Estimated Values: median (min,max)				
μ_1	0.77 (0.75, 0.84)	0.81 (0.77, 0.84)	0.80 (0.76, 0.83)	0.81 (0.73, 0.83)	0.81 (0.72, 0.89)
μ_2	0.50 (0.44, 0.55)	0.53 (0.35, 0.60)	0.52 (0.40, 0.65)	0.47 (0.37, 0.56)	0.52 (0.17, 0.79)
β_1	1.13 (1.02, 1.23)	1.11 (1.03, 1.16)	1.11 (0.98 1.16)	1.12 (1.08, 1.24)	1.08 (1.02, 1.15)
β_2	1.97 (1.59, 2.86)	1.97 (1.93, 2.13)	1.99 (1.79, 2.36)	1.96 (1.81, 2.16)	2.06 (1.93, 2.27)
δ	0.47 (0.19, 1)	0.10 (0.08, 0.13)	0.19 (0.14, 0.33)	0.30 (0.17,0.39)	0.43 (0.31, 0.57)
	$\delta = 0.6$	$\delta = 0.7$	$\delta = 0.8$	$\delta = 0.9$	$\delta = 1$
Parameter	Estimated Values: median (min,max)				
μ_1	0.74 (0.70, 0.87)	0.81 (0.76, 0.92)	0.81 (0.73, 0.95)	0.80 (0.72, 1.01)	0.80 (0.74, 0.96)
μ_2	0.46 (0.25, 0.66)	0.50 (0.28, 0.83)	0.52 (0.33, 0.78)	0.46 (0.42, 1.00)	0.52 (0.44, 0.94)
β_1	1.09 (1.02, 1.13)	1.08 (1.03, 1.23)	1.09 (0.97, 1.21)	1.10 (1.06, 1.31)	1.11 (1.06, 1,19)
β_2	2.14 (1.74, 2.42)	1.98 (1.63, 2.38)	2.03 (1.63, 2.26)	2.05 (1.64, 2.15)	1.95 (1.70, 2.12)
δ	0.57 (0.39, 1)	0.73 (0.31, 1)	0.77 (0.51,1)	0.91 (0.29, 1)	1 (0.47, 1)

Note: The estimation of each parameter is given by the minimum, the maximum and the median of 10 iterations.

5.4 Simulation Study II: Regression Framework

We established identifiability of the proposed model in Theorem 4.4.2 under regression frameworks in 4.4.7 and 4.4.8. In this section, we illustrate the identifiability issues in a simulation

study. We consider the proposed model 5.3.2: $S(t_1,t_2) = \exp\big(-\big[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta 2})^{1/\delta}\big]^{\delta}\big)$ with regression models given by

$$\begin{aligned} \log(\mu_1^{1/\delta}) &= \eta_{10} + \eta_{11}x_1 + \eta_{12}x_2 \\ \log(\mu_2^{1/\delta}) &= \eta_{20} + \eta_{21}x_1 + \eta_{22}x_2 \end{aligned} \tag{5.4.1}$$

In the data generating model, we consider a continuous and a binary x variables:

$$\begin{aligned} x_1 &\sim \mathrm{N}(0,\sigma^2) \\ x_2 &\sim \mathrm{Bern}(p) \end{aligned}$$

In particular, we have considered $x_2 \sim \mathrm{Bern}(0.5)$ and $x_1 \sim \mathrm{N}(0,0.1),\ \mathrm{N}(0,0.5),\ \mathrm{N}(0,1)$ respectively. For each data set with sample size $N = 1000$, we obtain the result under analysis model via the profile maximum likelihood estimator, which is summarized in Tables $5.6 - 5.8$, and the estimation of each parameter is given by the minimum, maximum and median over 10 iterations. According to the tables with different data generating models below, most of the estimation are close to their true value, which brings the identifiability to be achieved in these studies. The shape parameter β can also be estimated accurately via frequentist way. Although the estimation of coefficient of covariates may not be close to their true values for some of the cases, it is because the given MLE is the median (minimum, maximum) of 10 iterations. The larger number of repetitions will give more accuracy.

Table 5.6: Simulation Result 1 Under Model with Regression on μ

Data Generating Model with $x_1 \sim \mathrm{N}(0, 0.1)$, $x_2 \sim \mathrm{Bern}(0.5)$

$\eta_{10} = -2.5$, $\eta_{11} = 0.5$, $\eta_{12} = 1.3$, $\eta_{20} = -2$, $\eta_{21} = 0.4$, $\eta_{22} = 1.8$, $\beta_1 = 1.1$, $\beta_2 = 2$

Analysis	$\delta = 0.1$		$\delta = 0.2$		$\delta = 0.3$		$\delta = 0.4$		$\delta = 0.5$	
Model	Estimated Values: median (min,max)									
η_{10}	-2.3	(-3.5, -1.5)	-2.5	(-3.1, -2.0)	-2.5	(-3.4, -2.3)	-2.5	(-3.3, -1.9)	-2.6	(-4.1, -2.1)
η_{11}	0.87	(-1.6, 1.35)	0.45	(-0.9, 1.15)	0.47	(-1.3, 1.32)	0.52	(-1.7, 1.91)	0.44	(-1.3, 2.28)
η_{12}	0.97	(0.09, 3.97)	1.21	(0.62, 1.82)	1.19	(0.97, 2.26)	1.32	(0.67, 1.93)	1.52	(1.08, 2.75)
η_{20}	-1.7	(-2.9, -1.0)	-2.0	(-2.6, -1.7)	-2.0	(-2.6, -1.8)	-2.1	(-2.7, -1.6)	-2.0	(-3.4, -1.7)
η_{21}	0.39	(-0.7, 1.84)	0.62	(-1.8, 1.37)	0.36	(-1.7, 0.98)	0.54	(-1.0, 1.49)	0.43	(-0.8, 1.01)
η_{22}	1.52	(0.64, 4.27)	1.53	(1.36, 2.47)	1.83	(1.67, 2.60)	1.82	(1.20, 2.65)	1.83	(1.36, 3.40)
β_1	1.06	(1.01, 1.13)	1.09	(1.03, 1.17)	1.12	(1.07, 1.22)	1.13	(1.03, 1.28)	1.14	(0.98, 1.40)
β_2	2.05	(1.89, 2.22)	2.01	(1.91, 2.25)	1.99	(1.88, 2.27)	2.05	(1.99, 2.45)	1.98	(1.82, 2.11)
δ	0.11	(0.08, 0.14)	0.20	(0.17, 0.28)	0.29	(0.23,0.38)	0.40	(0.27, 0.59)	0.47	(0.25, 0.65)
Analysis	$\delta = 0.6$		$\delta = 0.7$		$\delta = 0.8$		$\delta = 0.9$		$\delta = 1$	
Model	Estimated Values: median (min,max)									
η_{10}	-2.7	(-3.8, -2.0)	-2.4	(-2.8, -2.0)	-2.7	(-3.7, -2.1)	-2.4	(-3.8, -2.3)	-2.5	(-4.7, -2.3)
η_{11}	0.51	(-1.0, 1.90)	0.65	(-0.8, 1.56)	0.47	(-1.2, 0.86)	0.47	(-0.2, 1.65)	0.45	(-0.5, 1.61)
η_{12}	1.55	(1.08, 2.48)	1.25	(0.86, 1.51)	1.41	(0.82, 2.47)	1.26	(1.08, 2.82)	1.34	(1.11, 3.35)
η_{20}	-2.1	(-3.0, -1.7)	-2.0	(-2.3, -1.8)	-2.1	(-3.1, -1.7)	-2.0	(-3.1, -1.8)	-2.1	(-4.1, -2.0)
η_{21}	0.59	(-0.2, 1.06)	0.38	(-0.3, 1.77)	0.40	(-0.5, 0.81)	0.39	(-0.3, 0.98)	0.37	(-0.0, 0.77)
η_{22}	2.00	(1.35, 3.00)	1.83	(1.47, 2.18)	1.98	(1.51, 3.11)	1.80	(1.54, 3.16)	1.87	(1.77, 4.16)
β_1	1.11	(1.00, 1.28)	1.10	(1.04, 1.23)	1.14	(1.01, 1.32)	1.10	(0.97, 1.29)	1.12	(1.02, 1.42)
β_2	1.95	(1.77, 2.14)	2.03	(1.90, 2.21)	2.00	(1.81, 2.12)	2.01	(1.84, 2.12)	1.97	(1.82, 2.12)
δ	0.62	(0.37, 0.85)	0.73	(0.58, 0.89)	0.76	(0.46, 1)	0.90	(0.51, 1)	1	(0.39, 1)

Table 5.7: Simulation Result 2 Under Model with Regression on μ

Data Generating Model with $x_1 \sim \mathrm{N}(0, 0.5)$, $x_2 \sim \mathrm{Bern}(0.5)$ $\eta_{10} = -2.5$, $\eta_{11} = 0.5$, $\eta_{12} = 1.3$, $\eta_{20} = -2$, $\eta_{21} = 0.4$, $\eta_{22} = 1.8$, $\beta_1 = 1.1$, $\beta_2 = 2$										
Analysis	$\delta = 0.1$		$\delta = 0.2$		$\delta = 0.3$		$\delta = 0.4$		$\delta = 0.5$	
Model	Estimated Values: median (min,max)									
η_{10}	-2.5	(-4.0, -1.8)	-2.5	(-2.8, -2.3)	-2.4	(-3.3, -2.1)	-2.4	(-3.0, -2.0)	-2.6	(-3.5, -1.9)
η_{11}	0.44	(-0.5, 1.54)	0.59	(-0.0, 1.30)	0.48	(-0.2, 0.88)	0.60	(0.25, 0.88)	0.48	(0.32, 0.87)
η_{12}	0.99	(0.01, 2.82)	1.24	(0.82, 1.63)	1.22	(0.93, 2.00)	1.17	(0.99, 2.03)	1.50	(1.03, 2.34)
η_{20}	-2.1	(-3.3, -0.9)	-2.0	(-2.3, -1.6)	-1.9	(-2.6, -1.6)	-1.9	(-2.5, -1.6)	-2.0	(-2.9, -1.6)
η_{21}	0.47	(-0.9, 1.50)	0.35	(-0.1, 0.74)	0.47	(0.03, 0.77)	0.41	(0.07, 0.56)	0.39	(0.23, 0.82)
η_{22}	1.77	(0.14, 3.17)	1.76	(1.20, 1.95)	1.70	(1.30, 2.41)	1.78	(1.29, 2.52)	1.90	(1.28, 2.76)
β_1	1.09	(1.04, 1.22)	1.12	(1.05, 1.14)	1.09	(1.00, 1.22)	1.08	(0.97, 1.17)	1.12	(0.98, 1.28)
β_2	2.01	(1.92, 2.15)	2.00	(1.92, 2.23)	2.01	(1.84, 2.06)	1.96	(1.90, 2.24)	1.96	(1.90, 2.20)
δ	0.10	(0.08, 0.12)	0.20	(0.18, 0.24)	0.30	(0.20,0.35)	0.39	(0.3, 0.63)	0.51	(0.33, 0.73)
Analysis	$\delta = 0.6$		$\delta = 0.7$		$\delta = 0.8$		$\delta = 0.9$		$\delta = 1$	
Model	Estimated Values: median (min,max)									
η_{10}	-2.4	(-3.1, -2.0)	-2.6	(-6.1, -1.9)	-2.6	(-3.2, -2.0)	-2.4	(-3.5, -2.2)	-2.6	(-3.6, -2.6)
η_{11}	0.47	(0.14, 0.72)	0.41	(0.18, 0.99)	0.56	(0.22, 0.78)	0.51	(0.21, 0.82)	0.50	(0.29, 0.83)
η_{12}	1.28	(0.76, 1.85)	1.37	(0.73, 4.76)	1.43	(0.79, 1.89)	1.28	(1.00, 2.17)	1.38	(1.22, 2.35)
η_{20}	-2.0	(-2.4, -1.8)	-2.1	(-5.0, -1.6)	-2.1	(-2.7, -1.7)	-1.9	(-2.6, -1.7)	-2.1	(-2.8, -1.9)
η_{21}	0.33	(0.15, 0.58)	0.37	(0.18, 0.94)	0.42	(0.21, 0.58)	0.39	(0.14, 0.62)	0.37	(0.25, 0.72)
η_{22}	1.77	(1.53, 2.24)	1.86	(1.29, 5.14)	1.90	(1.35, 2.51)	1.71	(1.50, 2.49)	1.91	(1.68, 2.80)
β_1	1.08	(0.95, 1.21)	1.12	(0.89, 1.52)	1.11	(0.95, 1.32)	1.10	(1.05, 1.33)	1.14	(1.06, 1.27)
β_2	2.05	(1.90, 2.18)	1.97	(1.83, 2.10)	2.04	(1.82, 2.08)	2.00	(1.91, 2.13)	2.01	(1.86, 2.04)
δ	0.65	(0.45, 0.78)	0.72	(0.23, 1)	0.80	(0.52, 1)	0.93	(0.63, 1)	1	(0.59, 1)

Table 5.8: Simulation Result 3 Under Model with Regression on μ

Data Generating Model with $x_1 \sim \mathrm{N}(0,1)$, $x_2 \sim \mathrm{Bern}(0.5)$ $\eta_{10} = -2.5$, $\eta_{11} = 0.5$, $\eta_{12} = 1.3$, $\eta_{20} = -2$, $\eta_{21} = 0.4$, $\eta_{22} = 1.8$, $\beta_1 = 1.1$, $\beta_2 = 2$										
Analysis	$\delta = 0.1$		$\delta = 0.2$		$\delta = 0.3$		$\delta = 0.4$		$\delta = 0.5$	
Model	Estimated Values: median (min,max)									
η_{10}	-2.4	(-3.2, -2.0)	-2.4	(-3.1, -2.1)	-2.5	(-3.4, -2.3)	-2.6	(-3.9, -2.2)	-2.4	(-3.0, -2.1)
η_{11}	0.53	(-0.1, 0.99)	0.48	(0.14, 0.78)	0.42	(0.22, 0.74)	0.48	(0.38, 0.89)	0.53	(0.37, 0.67)
η_{12}	1.20	(0.37, 1.66)	1.24	(0.06, 1.89)	1.32	(1.10, 2.65)	1.34	(0.72, 2.75)	1.31	(0.90, 1.84)
η_{20}	-2.1	(-2.8, -1.6)	-2.0	(-2.5, -1.7)	-2.0	(-2.6, -1.8)	-2.1	(-3.2, -1.7)	-1.9	(-2.3, -1.7)
η_{21}	0.36	(0.04, 0.76)	0.43	(0.11, 0.72)	0.37	(0.18, 0.78)	0.38	(0.18, 0.75)	0.40	(0.27, 0.54)
η_{22}	1.75	(1.08, 2.17)	1.79	(0.77, 2.16)	1.83	(1.51, 3.03)	1.90	(1.42, 3.26)	1.77	(1.44, 2.28)
β_1	1.09	(1.02, 1.16)	1.08	(1.04, 1.14)	1.10	(1.03, 1.28)	1.12	(0.96, 1.34)	1.12	(1.00, 1.25)
β_2	2.07	(1.88, 2.31)	2.02	(1.85, 2.12)	2.03	(1.91, 2.13)	1.99	(1.87, 2.18)	1.99	(1.87, 2.17)
δ	0.11	(0.09, 0.14)	0.20	(0.15, 0.23)	0.30	(0.19,0.36)	0.38	(0.23,0.49)	0.50	(0.35, 0.62)
Analysis	$\delta = 0.6$		$\delta = 0.7$		$\delta = 0.8$		$\delta = 0.9$		$\delta = 1$	
Model	Estimated Values: median (min,max)									
η_{10}	-2.3	(-2.8, -2.0)	-2.6	(-3.9, -1.9)	-2.4	(-4.8, -2.1)	-2.5	(-3.5, -2.2)	-2.5	(-3.8, -2.5)
η_{11}	0.47	(0.28, 0.60)	0.54	(0.34, 0.84)	0.50	(0.33, 1.06)	0.50	(0.37, 0.79)	0.51	(0.36, 0.88)
η_{12}	1.26	(0.75, 1.50)	1.33	(0.79, 2.51)	1.23	(0.82, 3.65)	1.34	(0.99, 2.35)	1.36	(1.18, 2.31)
η_{20}	-1.9	(-2.1, -1.7)	-2.1	(-3.4, -1.6)	-1.9	(-3.9, -1.7)	-2.1	(-2.7, -1.8)	-2.0	(-3.0, -1.8)
η_{21}	0.36	(0.26, 0.50)	0.40	(0.28, 0.85)	0.39	(0.30, 0.94)	0.43	(0.25, 0.76)	0.41	(0.38, 0.65)
η_{22}	1.81	(1.39, 2.03)	1.91	(1.29, 3.51)	1.83	(1.41, 3.94)	1.79	(1.51, 2.73)	1.84	(1.65, 2.85)
β_1	1.08	(1.01, 1.21)	1.09	(0.93, 1.27)	1.08	(0.98, 1.46)	1.15	(1.00, 1.30)	1.15	(1.06, 1.30)
β_2	2.04	(1.94, 2.20)	1.93	(1.81, 2.24)	2.06	(1.86, 2.13)	1.98	(1.83, 2.02)	1.98	(1.83, 2.08)
δ	0.61	(0.54, 0.79)	0.68	(0.34, 0.99)	0.83	(0.37, 1)	0.89	(0.56, 1)	1	(0.55, 1)

5.5 Simulation Study III: Regression on the Dependence Parameter δ

In this section, we will use the simulation study to illuminate the identifiability issues under the proposed model 5.3.2: $S(t_1,t_2) = \exp\big(-\big[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta 2})^{1/\delta}\big]^{\delta}\big)$ with regression on the dependence parameter δ given by

$$\text{logit}\,[\delta(x)] = \eta_0 + \eta_1 x_1 + \eta_2 x_2 \tag{5.5.1}$$

As before, we consider continuous and binary x variables. In the data generating model, the x variables are generated from

$$\begin{aligned} x_1 &\sim \text{N}(0,\sigma^2) \\ x_2 &\sim \text{Bern}(p) \end{aligned}$$

In this simulation study, we have simulated the data sets with covariates $x_1 \sim \text{N}(0,0.5)$ and $x_2 \sim \text{Bern}(0.5)$. The sample size is $N = 1000$; we obtain the result under analysis model via the profile maximum likelihood estimator, which is summarized in Table 5.9, and the estimation of each parameter is given by the minimum, maximum and median over 10 iterations. Most of the estimation of shape (β) and location (μ) parameters are close to their true values, which are used in data generating model. The MLE of coefficients bring in much larger error, especially in intercept. Setting regression on δ makes each individual obtain one dependence. It leads the estimation to getting a bigger variance.

Table 5.9: Simulation Results Under Model with Regression on δ

Analysis Model	Data Generating Model with $x_1 \sim N(0, 0.5)$ and $x_2 \sim Bern(0.5)$ $\eta_0 = -0.5,\ \eta_1 = 3.5,\ \eta_2 = 1,\ \beta_1 = 2,\ \beta_2 = 1.3$			
	$\mu_1 = \mu_2 = 0.3$		$\mu_1 = 0.8,\ \mu_2 = 0.5$	
Parameter	Estimated Values median, (min,max)			
η_0	-0.370	(-0.898, -0.071)	-0.639	(-1.051, 0.115)
η_1	3.501	(2.998, 4.111)	3.344	(3.163 5.005)
η_2	0.939	(0.510, 1.226)	1.011	(0.456, 1.622)
μ_1	0.301	(0.256, 0.315)	0.810	(0.752 0.875)
$\mu 2$	0.308	(0.257, 0.312)	0.493	(0.436, 0.558)
β_1	2.010	(1.952, 2.117)	1.994	(1.873 2.093)
β_2	1.317	(1.194, 1.370)	1.302	(1.169, 1.403)

5.6 Applications on Tamoxifen Trial Data

Tamoxifen is the most common hormone treatment for female breast cancer. It is currently used for the treatment of both early and advanced estrogen receptor-positive breast cancer in pre-menopausal and post-menopausal women. According to FDA, Tamoxifen can prevent breast cancer in women at high risk of developing the disease. In December 1992, a multicenter randomized clinical trial for patients with node-negative breast cancer began accruing subjects. This trial lasted almost 10 years from 1992 to the summer of 2002. A total of 769 women were randomized in the study; there are 383 women in the Tamoxifen-alone arm and 386 in the combined radiation and

Tamoxifen arm. In our study, we only include those patients who accrued at a single contributor institution, i.e., there are 321 patients in the Tamoxifen-alone arm and 320 in the other arm.

The original design of Tamoxifen trial was for an equivalence study with disease-free survival as the main endpoint. The events recorded were local relapse, axillary relapse, distant relapse, second malignancy of any type, and death. The time of the first occurrence of each type of event was documented. For instance, if a patient experienced distant relapse at 1 year, another distant relapse at 2 years and a local relapse at 3 years, then the only events recorded are the distant relapse at 1 year and the local relapse at 3 years. Because distant relapse at 2 years is the second relapse of the same type, so it was not recorded. There is a censoring variable for each type of event to indicate whether the event occurred or not. For each patient in the study, there are seven demographic covariates which contain age, trt (randomized treatment), hgb (haemoglobin), hrlevel (hormone receptor level), hist (histology), pathsize (size of the tumor) and nodediss (whether axillary node dissection was done). The clinical aspects of the study and more details of the original analysis can be found in Fyles et al. (2004).

5.6.1 Data Description

The data set we analyzed involved 641 women at age 50 or older with early breast cancer. We considered T_1 as the time of first relapse; that is, $T_1 = \min(T_{11}, T_{12}, T_{13}, T_{14})$ where T_{11} is time to local relapse, T_{12} is time to axillary relapse, T_{13} is time to distant relapse and T_{14} is time to second malignancy of any type. The competing risk of first relapse that we consider is death and T_2 is time to death without any of these events. The observed time is $T = \min(T_1, T_2)$ and censoring indicator d_i is 1 when T_i is observed and 0 for censored. The survival time (in years) and cause of death information are classified in Table 5.10.

Table 5.10: Six Hundred Forty-One Patients with Early Breast Cancer

Type of Event	Number of Events
Death	12
Relapse	126
Censored	503

The Kaplan-Meier plots for cause-specific survival from each risk, relapse and death, are shown in Figure 5.5.

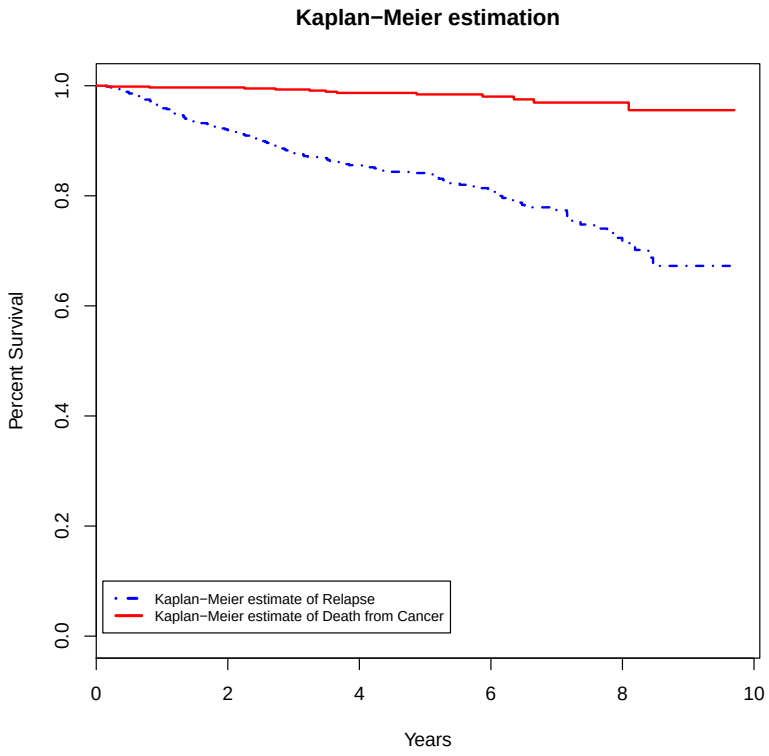


Figure 5.5: The Kaplan-Meier plot for Tamoxifen trial data

5.6.2 Dependent Competing Risks Model

For dependent competing risks model (DCR model), we used the complete cause of death information which included cancer, other causes and censored. The two competing risks factors are cancer and other causes. The dependent parametric competing risks model we considered can be written as

$$S(t_1,t_2) = \exp\big(-\big[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta 2})^{1/\delta}\big]^{\delta}\big)$$

Based on the MLE estimators, the plot of cumulative incidence function for Tamoxifen trial data is shown in Figure 5.6 and the corresponding estimators are given by Table 5.11.

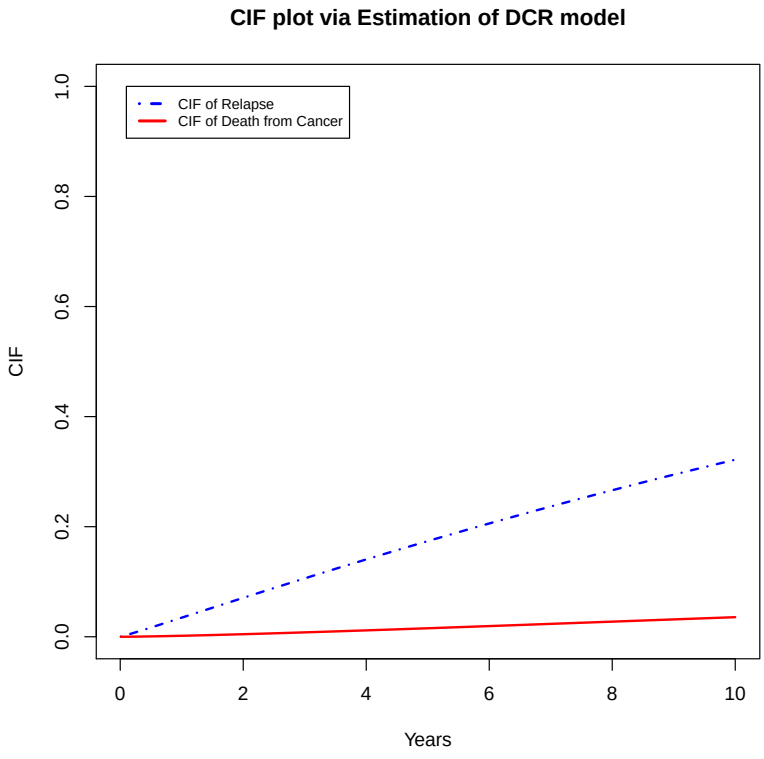


Figure 5.6: The estimated survival plot with MLE estimator for Tamoxifen trial data

Table 5.11: Maximum Likelihood Estimates under DCR Model for Tamoxifen Trial Data

Parameter	μ_1	μ_2	$\beta 1$	β_2	δ	-logLikelihood
mean	0.037	0.034	1.074	1.084	0.03	614.05

In Figure 5.6, the dot blue line is always above the red solid line. It is obvious that patients have higher chance to relapse than death from the cancer according to this Tamoxifen trail data.

We also consider a parallel Bayesian analysis of the data with the following prior distributions: $\mu_1 \sim \text{Gam}(1,1)$, $\mu_2 \sim \text{Gam}(1,1)$, $\beta_1 \sim \text{Gam}(0.01, 2.5)$, $\beta_2 \sim \text{Gam}(0.01, 2.5)$ and $\delta \sim \pi \cdot \text{Unif}(0,1) + (1-\pi) \cdot \text{degenerate}(1)$ with $\pi \sim \text{Unif}(0,1)$. We ran 60000 MCMC iterations and the posterior means of $\mu_1, \mu_2, \beta_1, \beta 2$ and δ which are computed from the last 50000 iterations, shown in Table 5.12. The estimates from maximum likelihood and Bayesian analysis are similar besides the estimate of δ. The posterior mean deviance is 1232, based on which we obtain posterior mean -log(Likelihood) $= 616$.

Table 5.12: Posterior Estimation Under DCR Model for Tamocifen Trial Data

Parameter	μ_1	μ_2	$\beta 1$	β_2	δ
mean	0.037	0.006	1.048	1.222	0.813
median	0.036	0.004	1.046	1.195	1
95% CI	(0.026, 0.051)	(8e-4, 0.022)	(0.886, 1.219)	(0.742, 1.835)	(0, 1)

The posterior density of δ is given by Figure 5.7. It shows that the likelihood may be flat due to density of δ is flat, which could lead the estimation to be hardly obtained accurately.

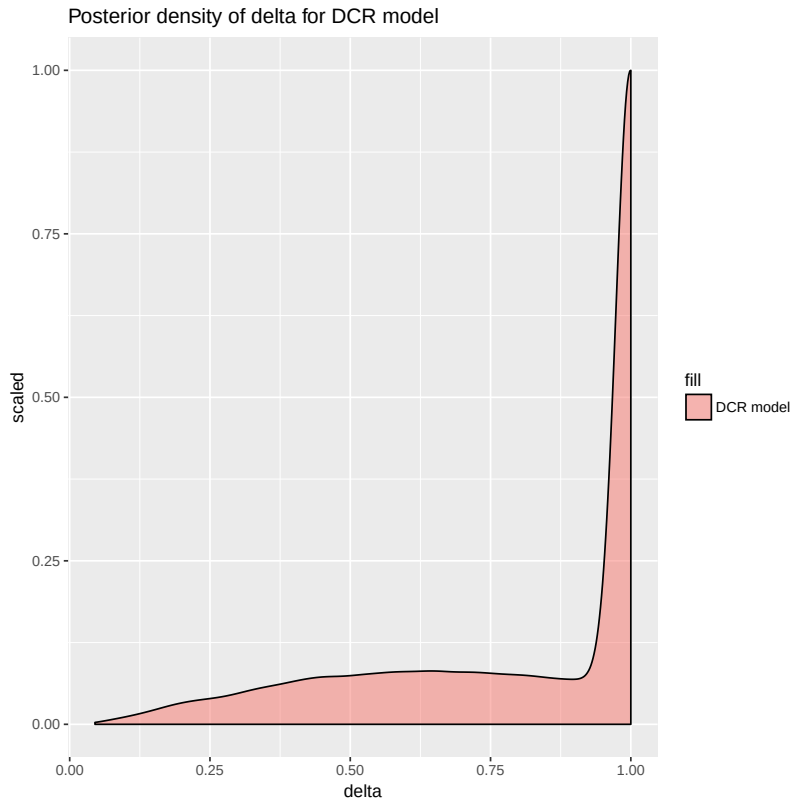


Figure 5.7: The posterior density of δ of DCR model for Tamocifen trial data

5.6.3 Dependent Competing Risks Regression Model

We next analyzed the same data in a regression framework including covariate information. The dependent competing risks model with regression on location parameter (DCRR model) has been considered, which is given by $S(t_1,t_2) = \exp\left(-\left[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta 2})^{1/\delta}\right]^{\delta}\right)$, where $\log(\mu_j^{1/\delta}) = X\eta$ and $j = 1,2$. The covariates that have been used for analysis contain age, pathsize (size of the tumor), hgb (haemoglobin), trt (wether randomized treatment is tamoxifen alone), hrlevelPOS (whether hormone receptor level is positive), dissection (whether axillary node dissection was done) and hist (histology) with five categories. Therefore, the dependent competing risks regres-

sion model involves 10 covariates: age, pathsize, hgb, trt, hrlevelPOS, dissection, DUL, LOB, MED and MIX. The MLE estimates are given by Table 5.13.

Table 5.13: Maximum Likelihood Estimation Under DCRR Model for Tamoxifen Trial Data

Parameter	$\beta 1$	β_2	δ	intercept1	age1	pathsize1	hgb1
mean	1.018	1.040	0.04	-0.05	-0.047	-0.033	-0.555
Parameter	trt1	hrlevelPOS1	dissection1	DUL1	LOB1	MED1	MIX1
mean	-0.404	-0.832	0.264	0.607	0.053	0.785	0.578
Parameter	intercept2	age2	pathsize2	hgb2	trt2	hrlevelPOS2	dissection2
mean	0.497	-0.059	0.492	-0.607	0.085	0.902	0.587
Parameter	DUL2	LOB2	MED2	MIX2	-logL		
mean	0.590	0.612	0.608	1.462	-336.5		

We again consider a parallel Bayesian analysis with prior distributions as $\eta_1^{'}s \sim \text{Normal}(0,5)$, $\eta_2^{'}s \sim \text{Normal}(0,5)$, $\beta_1 \sim \text{Gam}(0.01, 2.5)$, $\beta_2 \sim \text{Gam}(0.01, 2.5)$ and $\delta \sim \pi \cdot \text{Unif}(0,1) + (1-\pi) \cdot \text{degenerate}(1)$ with $\pi \sim \text{Unif}(0,1)$. After specifying the prior distributions, we ran 60000 MCMC iterations and looked at posterior means calculated from the last 50000 iterations of the respective parameters involved in the model. The posterior estimates of $\beta_1,\ \beta_2,\ \delta$ and $\eta^{'}s$ are shown in Table 5.14.

Table 5.14: Posterior Estimation Under DCRR Model for Tamoxifen Trial Data

Parameter	β1	β_2	δ	intercept1	age1
mean	1.078	1.148	0.204	-0.584	-0.073
95% CI	(0.92, 1.24)	(0.93, 1.40)	(0.08, 0.37)	(-4.67, 3.51)	(-0.17, 0.02)
Parameter	pathsize1	hgb1	trt1	hrlevelPOS1	dissection1
mean	2.535	-0.111	1.621	-2.78	-1.27
95% CI	(1.16, 4.78)	(-0.28, -0.01)	(0.06, 3.57)	(-5.26, -0.69)	(-3.19, 0.70)
Parameter	DUL1	LOB1	MED1	MIX1	intercept2
mean	0.450	0.829	-0.198	0.107	-1.334
95% CI	(-1.84, 2.49)	(-2.07, 3.85)	(-3.89, 3.55)	(-2.19, 2.35)	(-5.49, 2.81)
Parameter	age2	pathsize2	hgb2	trt2	hrlevelPOS2
mean	-0.022	2.297	-0.171	0.648	0.226
95% CI	(-0.13, 0.09)	(0.71, 4.54)	(-0.34, -0.07)	(-1.26, 2.77)	(-2.91, 3.56)
Parameter	dissection2	DUL2	LOB2	MED2	MIX2
mean	-0.530	0.381	-1.405	-0.379	-0.421
95% CI	(-2.82, 1.92)	(-2.07, 2.84)	(-5.14, 1.98)	(-4.59, 3.59)	(-3.02, 2.12)

According to the posterior mean of deviance, which is -660, we can compute -log(Likelihood) $= -330$. From the Table 5.14, the credible intervals corresponding to pathsize, hgb trt and hrlevelPOS do not contain zero and hence can be termed as having "significant" influence on the nonterminal event of relapse and only variable pathsize and hgb are "significant" for the terminal event of death. More specifically, the size of tumor and the haemoglobin level of patient will affect both relapse and death. The larger size of tumor leads to worse survival and the higher haemoglobin

brings better survival. The treatment effect influences relapse. Since the correlation of treatment with Tamoxifen only is positive, which means the worse survival, the patients who use Tamoxifen and radiation together as treatment have lower chance to suffer relapse. The hormone receptor level impacts the relapse as well and the positive level will bring the patients to better relapse-free survival. With advances in treatment, early stage breast cancer is less aggressive with longer survival, which may explain the non-significant effects for the terminal event of death.

Figure 5.8 gives comparison between two treatment groups; there is a significant difference between blue solid line and green dot line. That is, the treatment affects the chance to relapse; there is a high probability to suffer relapse if the patient used Tamoxifen only as the treatment. Additionally, Figure 5.9 shows comparison between two treatment groups with different hormone receptor levels. The positive hormone receptor level will lead to a better relapse-free survival and using Tamoxifen combined with radiation as treatment could even achieve a much better relapse-free survival.

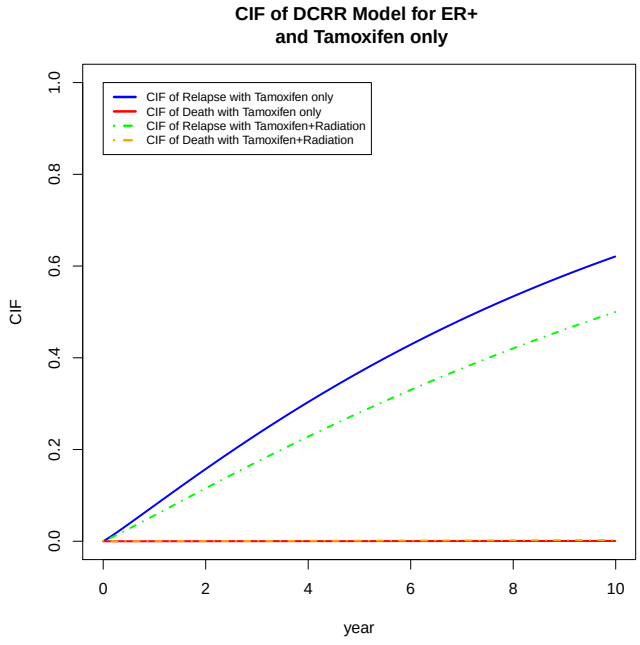


Figure 5.8: CIF plot under DCRR model for Tamoxifen trial data by treatment

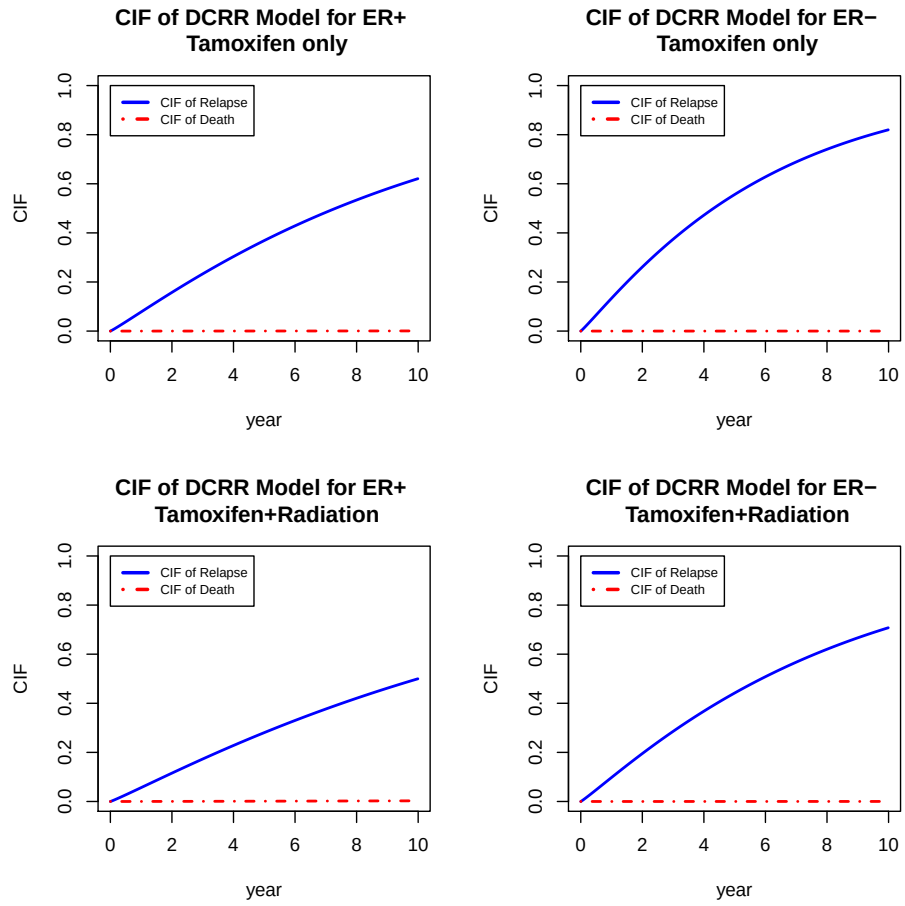


Figure 5.9: CIF plot under DCRR model for Tamoxifen trial data by treatment and hormone receptor level

CHAPTER 6

SEMI-COMPETING RISKS AND APPLICATION

6.1 Introduction

The terminology of semi-competing risks was introduced in Fine et al. (2001) to describe the situation in which the data arise when the observed event time of some non-terminal event is subject to some terminal event. For instance, we can observe both relapse (non-terminal) and death (terminal) from the same cancer. And it is known by Fine et al. (2001) that a terminal event censors a non-terminal event but not vice versa. Methods in the literature have been developed to deal with this type of data usually belong to one of two groups. The first strategy focuses on setting up conditional models for the hazard functions of the non-terminal and terminal events via illness-death model, such as Ye et al. (2007), Xu et al. (2010), K. H. Lee et al. (2015) and K. H. Lee et al. (2017). The other approach considers models which describe the joint distribution of T_1 and T_2 and either model the dependence between T_1 and T_2 via a copula, such as Day et al. (1997), Fine et al. (2001), W. Wang (2003), Ghosh (2006), Peng & Fine (2007), Hsieh et al. (2008) and H. Fu et al. (2013); or make the dependence arbitrary, such as Cook & Lawless (1997) and Ghosh & Lin (2000). In addition, Jiang et al. (2005) have extended the approach to truncated semi-competing risks.

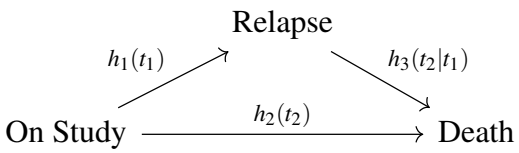
In this chapter, we propose a parametric joint distribution model to the non-terminal and terminal events via a copula. In particular, we will show the identifiability issues associated with this proposed model on semi-competing risks data. We also illustrate the difference between competing risks and semi-competing risks analyses in real data applications.

6.2 Semi-Competing Risks Models

Semi-competing risks data are mainly to study participants being able to experience a non-terminal event (such as relapse), a terminal event (such as death), and possibly both. Let T_1 and T_2 be the failure times and possibly dependent with the joint survival function $S(t_1,t_2) = P(T_1 > t_1, T_2 > t_2)$. The random variable T_2 may censor T_1, but not vice versa; as an illustration, T_2 is death from any cause and T_1 is relapse. The marginal survival function of T_1 is given by $S_1(t_1) = P(T_1 > t_1)$ and the corresponding marginal survival function of T_2 can be written as $S_2(t_2) = S(0,t_2) = P(T_2 > t_2)$. Let C be a censoring time independent of both T_1 and T_2 (for example, administrative loss to follow-up). We observe the variables $Y_1 = \min(T_1, T_2, C)$, $d_1 = I(T_1 \leq \min(T_2, C))$, $Y_2 = \min(T_2, C)$ and $d_2 = I(T_2 \leq C)$, where $I(\cdot)$ is the indicator function. In particular, (Y_1, Y_2) are observed on the restricted domain $Y_1 \leq Y_2$ and when terminal event occurs first, we have $Y_1 = Y_2$. Thus, the observed data are n independent and identically distributed replications of (Y_1, Y_2, d_1, d_2), which is denoted by $\{(Y_{1i}, Y_{2i}, d_{1i}, d_{2i}), i = 1, \ldots, n\}$.

6.2.1 The Illness-Death Model

One of the approaches to handle semi-competing risks data is considering the illness-death model, which is a special case of the multi-state modeling framework, such as Xu et al. (2010), K. H. Lee et al. (2015) and K. H. Lee et al. (2017)).



Illustrate the situation with three nodes labeled *on study* for the beginning of study, *relapse* for the occurrence of nonterminal event, *death* for the occurrence of terminal event. The hazard or transition rates are defined as:

$$h_1(t_1) = \lim_{\triangle \to 0} P(t_1 \leq T_1 < t_1 + \triangle | T_1 \geq t_1, T_2 \geq t_1)/\triangle, \qquad t_1 > 0 \tag{6.2.1}$$

$$h_2(t_2) = \lim_{\triangle \to 0} P(t_2 \leq T_2 < t_2 + \triangle | T_1 \geq t_2, T_2 \geq t_2)/\triangle, \qquad t_2 > 0 \tag{6.2.2}$$

$$h_3(t_2|t_1) = \lim_{\triangle \to 0} P(t_2 \leq T_2 < t_2 + \triangle | T_1 = t_1, T_2 \geq t_2)/\triangle, \qquad t_2 > t_1 > 0 \tag{6.2.3}$$

When either terminal or non-terminal event occurs first, equations 6.2.1 and 6.2.2, which are the cause-specific hazard function or crude hazard function for the competing risks part of the model, will be observed. The equation 6.2.3 defines the hazard rate of terminal event which follows the occurrence of the non-terminal event. Generally speaking, $h_3(t_2|t_1)$ depends on both event times t_1 and t_2. According to a Markov process, the hazard rate $h_3(t_2|t_1) = h_3(t_2)$ only depends on t_2. Alternatively, in a semi-Markov model, it is specified that the hazard rate $h_3(t_2|t_1) = h_3(t_2 - t_1)$ depends only on the difference between t_1 and t_2. An illness-death model via a copula framework has been considered in by Xu et al. (2010). K. H. Lee et al. (2015) considered a Bayesian approach and K. H. Lee et al. (2017) considered an accelerated failure time model in this setting.

6.2.2 Joint Distribution Model with Dependence via a Frailty

An alternative approach to semi-competing risks data considers model for the joint analysis of event times based on frailties. Fine et al. (2001) models the correlated events via a Gamma frailty,

which can be referred to Clayton (1978) and Hougaard (1986), defined on the upper wedge, given by

$$S(t_1,t_2) = [S_1(t_1)^{1-\theta} + S_2(t_2)^{1-\theta} - 1]^{1/(1-\theta)} \tag{6.2.1}$$

where $S_1(\cdot)$ and $S_2(\cdot)$ are survival functions. θ is the shared frailty parameter, when $\theta = 1$, T_1 and T_2 are independent on the upper wedge and $S_1(t_1) = P(T_1 > t_1 | T_2 > t_1)$. Since $S(t_1,t_2)$ may not follow the model 6.2.1 on the lower wedge, θ may not be related to Kendall's τ and hence may not satisfy

$$\tau = (\theta - 1)/(\theta + 1) \tag{6.2.2}$$

The joint density on the upper wedge can be written as

$$f_{t_1 \le t_2}(t_1,t_2) = \theta \left[S_1(t_1)^{1-\theta} + S_2(t_2)^{1-\theta} - 1\right]^{\frac{2\theta-1}{\theta-1}} S_1(t_1)^{-\theta} S_2(t_2)^{-\theta} f_1(t_1) f_2(t_2) \tag{6.2.3}$$

where $f_j(t_j) = -\partial S(t_1,t_2)/\partial t_j$ for $j = 1,2$. The joint density $f_{t_2<t_1}(t_1,t_2)$ on the lower wedge is unspecified but satisfies the condition

$$S(t_1,t_2) = \int_{t_2}^{\infty} \int_{t_1}^{v} f_{t_1 \le t_2}(u,v) du dv \; + \; \int_{t_2}^{\infty} \int_{v}^{\infty} f_{t_2<t_1}(u,v) du dv \tag{6.2.4}$$

Both Genest et al. (1995) and Shih & Louis (1995) have shown that maximization of a pseudo-likelihood derived from consistent estimators of $S_1(t_1)$ and $S_2(t_2)$ gives a consistent estimator of θ. However, for semi-competing risks data, Fine et al. (2001) have pointed out that a consistent estimator for $S_1(\cdot)$ may not exist without an estimator for θ and provided a robust estimator $\hat{\theta}$ based on the joint density on upper wedge.

6.2.3 A Dependent Parametric Semi-Competing Risks Model

According to Day et al. (1997) and Fine et al. (2001), the joint survival function $S(t_1,t_2)$ is only identifiable when $t_1 \leq t_2$. In this section, we examine the identifiability of a parametric model on semi-competing risks data. In the setup we only specify the model on the upper wedge, that is, the region $T_1 \leq T_2$. The dependence structure can be formulated via the Gumbel copula which also originates in a positive stable frailty (refer to Chapter 3). That is, for $\delta \in (0,1]$ and $0 \leq t_1 \leq t_2 \leq \infty$, the joint survival function $S(t_1,t_2)$ can be written as

$$S(t_1,t_2) = \begin{cases} \exp\big(-[(\log S_1(t_1))^{1/\delta} + (\log S_2(t_2))^{1/\delta}]^{\delta}\big), & \delta \in (0,1) \\ S_1(t_1)\cdot S_2(t_2), & \delta = 1 \end{cases} \tag{6.2.1}$$

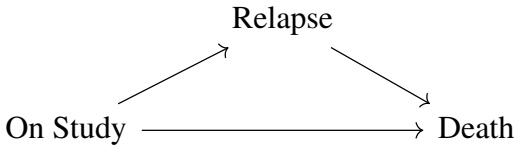
when $\delta = 1$, T_1 and T_2 are independent on the upper wedge and $S_1(t_1) = P(T_1 > t_1 | T_2 > t_1)$. δ represents the dependence between T_1 and T_2. More specifically, the joint survival function $S(t_1,t_2)$ with Weibull marginal is given by

$$\begin{aligned} S(t_1,t_2) &= \psi\big[\psi^{-1}(S_1(t_1)) + \psi^{-1}(S_2(t_2))\big] \\ &= \exp\Big[-\Big(\big[-\log(\exp(-\mu_1 t_1^{\beta_1}))\big]^{\frac{1}{\delta}} + \big[-\log(\exp(-\mu_2 t_2^{\beta_2}))\big]^{\frac{1}{\delta}}\Big)^{\delta}\Big] \\ &= \exp\Big(-\Big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\Big]^{\delta}\Big) \end{aligned} \tag{6.2.2}$$

where $\psi(\cdot)$ is the generator function and $\psi^{-1}(\cdot)$ is the generator inverse function of Gumbel copula.

6.2.3.1 Likelihood Function

Suppose T_1 and T_2 are the failure times to non-terminal event and terminal event respectively. The joint survival function of T_1 and T_2 can be written in terms of $S(t_1,t_2)$, which is given by equation 6.2.2. Since the semi-competing risk problem can be described as below,



the observed data set$\{(Y_{1i},d_{1i},Y_{2i},d_{2i}),i=1,\ldots,n\}$ can be shown in the Table 6.1.

Table 6.1: Observed Data Information

Observed Data Set		
	(Y_1,Y_2)	(d_1,d_2)
Relapse and censored prior to death	(T_{1i},C_i)	$(1,0)$
Dead following relapse	(T_{1i},T_{2i})	$(1,1)$
Dead without relapse	(T_{2i},T_{2i})	$(0,1)$
Censored prior to relapse or death	(C_i,C_i)	$(0,0)$

To obtain the likelihood function, we have the following four cases where independent censoring is assumed.

Case1: $d_1 = 1, d_2 = 0$

$$\begin{aligned} P(Y_1 = T_1, Y_2 = C) &= Pr(T_1 = t_1, T_2 > t_1) \cdot Pr(T_2 > c | T_1 = t_1, T_2 > t_1) \\ &= Pr(T_1 = t_1, T_2 > t_1, T_2 > c) \\ &= Pr(T_1 = t_1, T_2 > c) \\ &= \left. \frac{-\partial S(t_1, t_2)}{\partial t_1} \right|_{t_2 = c} \end{aligned} \tag{6.2.1}$$

Case2: $d_1 = 1, d_2 = 1$

$$\begin{aligned} P(Y_1 = T_1, Y_2 = T_2) &= Pr(T_1 = t_1, T_2 > t_1) \cdot Pr(T_2 = t_2 | T_1 = t_1, T_2 > t_1) \\ &= Pr(T_1 = t_1, T_2 > t_1, T_2 = t_2) \\ &= Pr(T_1 = t_1, T_2 = t_2) \\ &= f(t_1, t_2) \end{aligned} \tag{6.2.2}$$

Case3: $d_1 = 0, d_2 = 1$

$$\begin{aligned} P(Y_1 = T_2, Y_2 = T_2) &= Pr(T_1 > t_2, T_2 = t_2) \\ &= \left. \frac{-\partial S(t_1, t_2)}{\partial t_2} \right|_{t_1 = t_2} \end{aligned} \tag{6.2.3}$$

Case4: $d_1 = 0, d_2 = 0$

$$\begin{aligned} P(Y_1 = C, Y_2 = C) &= Pr(T_1 > c, T_2 > c) \\ &= S(t_1, t_2)\big|_{t_1 = t_2 = c} \end{aligned} \tag{6.2.4}$$

According to W. Wang (2003) and H. Fu et al. (2013), the joint likelihood function can be given by:

$$\begin{aligned}
\mathbf{L} = & \prod_{i=1}^{n} P(Y_{1i} = T_{1i}, Y_{2i} = C_i)^{d_{1i}(1-d_{2i})} \cdot P(Y_{1i} = T_{1i}, Y_{2i} = T_{2i})^{d_{1i}d_{2i}} \\
& \times P(Y_{1i} = T_{2i}, Y_{2i} = T_{2i})^{(1-d_{1i})d_{2i}} \cdot P(Y_{1i} = C_i, Y_{2i} = C_i)^{(1-d_{1i})(1-d_{2i})} \\
= & \prod_{i=1}^{n} f(t_{1i}, t_{2i})^{d_{1i}d_{2i}} \cdot S_1^{'}(t_{1i}, t_{2i})^{d_{1i}(1-d_{2i})} \cdot S_2^{'}(t_i, t_i)^{(1-d_{1i})d_{2i}} \cdot S(t_i, t_i)^{(1-d_{1i})(1-d_{2i})} \quad (6.2.5)
\end{aligned}$$

where (t_{1i}, t_{2i}) is the observed event times from case 1 and case 2; instead of using (t_{1i}, t_{2i}), we use t_i as the observed failure time for case3 and case4 since only a signal time point is available in these two cases. $f(\cdot)$ is the joint density function and $S(\cdot)$ is the joint survival function of (T_1, T_2), and the notation $S_j^{'}(\cdot)$ for $j = 1, 2$ is given by

$$S_j^{'}(t_1, t_2) = \frac{-\partial S(t_1, t_2)}{\partial t_j}$$

6.3 Identifiability Issues Under Equality of Shape Parameters

The nonparametric estimator for the predictive hazard ratio with bivariate right-censored data has been introduced by Oakes (1989), from which the dependence parameter can be obtained. Hougaard (1986) has also proposed a way to estimate the dependence parameter in which to estimate the integrated hazards in the marginal distributions by using the generalization of Nelson's method first and then restricting the marginal distribution to MLE of dependence parameter. According to Fine et al. (2001), in the fully nonparametric setting, $S_2(t_2)$ is fully identified from semi-competing risks data, and $S(t_1, t_2)$ is completely identified only in the upper wedge of the support of (T_1, T_2), that is, the region $(0 < t_1 \leq t_2)$. Moreover, $S_1(t_1)$ is not identified without additional assumptions. Jiang et al. (2003) has showed the identifiability of a semiparametric model

with a parameterized association (via the Clayton copula) and unspecified marginal distributions on semi-competing risks data.

In Theorem 4.2.1 in Section 4.2 we showed that our proposed parametric model with the restriction of equality of shape parameters is not identifiable in the setting of competing risks. Semi-competing risks data provide an additional level of information (time to terminal event following the non-terminal event) or that when observe T_1 we additionally observe $\min(T_2, C)$ and $d_2 = I(T_2 \leq C)$. The joint density $f(t_1, t_2)$ on the upper wedge which is denoted by $P(T_1 = t_1, T_2 = t_2, T_2 \geq T_1)$ can be obtained via

$$P(T_2 = t_2 | T_1 = t_1, T_2 \geq T_1) = \frac{P(T_1 = t_1, T_2 = t_2, T_2 \geq T_1)}{P(T_1 = t_1, T_2 \geq T_1)}$$

where $P(T_1 = t_1, T_2 \geq T_1)$ and $P(T_2 = t_2 | T_1 = t_1, T_2 \geq T_1)$ can be obtained from the observed data directly. As long as we can show $f(t_1, t_2)$ is identifiable for $\forall\, 0 \leq t_1 \leq t_2 \leq \infty$, then the identifiability property is established according to the semi-competing risks data.

Theorem 6.3.1. *Suppose the latent failure times (T_1, T_2) follow the joint survival function 6.2.2; then the joint density function $f(t_1, t_2)$ is identifiable from the observed semi-competing data (Y_1, Y_2, d_1, d_2).*

Proof. Since the $S(t_1, t_2)$ is defined on the upper wedge for semi-competing risks data, $P(T_1 = t_1, T_2 = t_2, T_2 \geq T_1) = P(T_1 = t_1, T_2 = t_2) = f(t_1, t_2)$, to show $f(t_1, t_2)$ is identifiable on the upper wedge, which is equivalent to show that

$$f(t_1, t_2 | \theta) = f(t_1, t_2 | \theta^*) \quad or \quad \frac{f(t_1, t_2 | \theta)}{f(t_1, t_2 | \theta^*)} = 1 \tag{6.3.1}$$

for $\forall\ 0 \le t_1 \le t_2 \le \infty$ if and only if $\theta = \theta^*$. According to 6.2.2, the joint survival function under same shape parameter structure is given by

$$S(t_1,t_2) = \exp\Big(- \Big[(\mu_1 t_1^{\beta})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta})^{\frac{1}{\delta}} \Big]^{\delta} \Big)$$

The joint density function $f(t_1,t_2)$ can be attained by taking the second derivative of $S(t_1,t_2)$ with respect to $t_1,\ t_2$:

$$\begin{aligned} f(t_1,t_2) =& \frac{\partial^2 (1 - S(t_1,t_2))}{\partial t_1 \partial t_2} \\ =& \exp\Big(- \big[\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}} \big]^{\delta} \Big) \cdot \big[\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}} \big]^{\delta-2} \cdot (\mu_1\mu_2)^{\frac{1}{\delta}} \cdot \beta^2 \\ & \times (t_1 t_2)^{\frac{\beta}{\delta}-1} \cdot \left(\frac{\delta - 1}{\delta} - \big[\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}} \big]^{\delta} \right) \end{aligned} \tag{6.3.2}$$

Case1: When $\mu_1 = \mu_1^*, \mu_2 = \mu_2^*, \beta = \beta^*$, equation 6.3.1 is obviously satisfied.

Case2: When $\mu_1 \neq \mu_1^*, \mu_2 \neq \mu_2^*, \beta \neq \beta^*$, assume that equation 6.3.1 is true for $\forall\ 0 \le t_1 \le t_2 \le \infty$, then take the limit of equation 6.3.1 with respect to $t_2 \to 0$, that is:

$$\begin{aligned} \lim_{t_2 \longrightarrow 0} \frac{f(t_1,t_2|\mu_1,\mu_2,\beta)}{f(t_1,t_2|\mu_1^*,\mu_2^*,\beta^*)} =& \lim_{t_2 \longrightarrow 0} \frac{\exp\Big(- \big[\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}} \big]^{\delta} \Big)}{\exp\Big(- \big[(\mu_1^*)^{\frac{1}{\delta}} t_1^{\frac{\beta^*}{\delta}} + (\mu_2^*)^{\frac{1}{\delta}} t_2^{\frac{\beta^*}{\delta}} \big]^{\delta} \Big)} \cdot \frac{(t_1 t_2)^{\frac{\beta}{\delta}-1}}{(t_1 t_2)^{\frac{\beta^*}{\delta}-1}} \\ & \times \frac{\big[\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}} \big]^{\delta-2}}{\big[(\mu_1^*)^{\frac{1}{\delta}} t_1^{\frac{\beta^*}{\delta}} + (\mu_2^*)^{\frac{1}{\delta}} t_2^{\frac{\beta^*}{\delta}} \big]^{\delta-2}} \cdot \frac{(\mu_1\mu_2)^{\frac{1}{\delta}} \cdot \beta^2}{(\mu_1^*\mu_2^*)^{\frac{1}{\delta}} \cdot (\beta^*)^2} \\ & \times \frac{\left(\frac{\delta-1}{\delta} - \big[\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}} \big]^{\delta} \right)}{\left(\frac{\delta-1}{\delta} - \big[(\mu_1^*)^{\frac{1}{\delta}} t_1^{\frac{\beta^*}{\delta}} + (\mu_2^*)^{\frac{1}{\delta}} t_2^{\frac{\beta^*}{\delta}} \big]^{\delta} \right)} \end{aligned} \tag{6.3.3}$$

since $T_1 \leq T_2$, when $t_2 \to 0$, it is obvious $t_1 \to 0$ as well. Thus, the limit of equation 6.3.1 can also be written as

$$\begin{aligned}\lim_{t_2 \longrightarrow 0} \frac{f(t_1,t_2|\mu_1,\mu_2,\beta)}{f(t_1,t_2|\mu_1^*,\mu_2^*,\beta^*)} &= \lim_{t_2 \longrightarrow 0} \left[\frac{\mu_1^{\frac{1}{\delta}} t_1^{\frac{\beta}{\delta}} + \mu_2^{\frac{1}{\delta}} t_2^{\frac{\beta}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}} t_1^{\frac{\beta^*}{\delta}} + (\mu_2^*)^{\frac{1}{\delta}} t_2^{\frac{\beta^*}{\delta}}} \right]^{\delta-2} \cdot \left(\frac{\mu_1\mu_2}{\mu_1^*\mu_2^*} \right)^{\frac{1}{\delta}} \\ &\quad \times \cdot \left(\frac{\beta}{\beta^*} \right)^2 \cdot (t_1 t_2)^{\frac{\beta-\beta^*}{\delta}} \\ &= \lim_{t_2 \longrightarrow 0} \left[\frac{\mu_1^{\frac{1}{\delta}} + \mu_2^{\frac{1}{\delta}} \left(\frac{t_2}{t_1}\right)^{\frac{\beta}{\delta}}}{(\mu_1^*)^{\frac{1}{\delta}} + (\mu_2^*)^{\frac{1}{\delta}} \left(\frac{t_2}{t_1}\right)^{\frac{\beta^*}{\delta}}} \right]^{\delta-2} \cdot \left(\frac{\mu_1\mu_2}{\mu_1^*\mu_2^*} \right)^{\frac{1}{\delta}} \\ &\quad \times \cdot \left(\frac{\beta}{\beta^*} \right)^2 \cdot \left(\frac{t_2}{t_1} \right)^{\frac{\beta-\beta^*}{\delta}} \cdot t_1^{\beta-\beta^*} \end{aligned} \tag{6.3.4}$$

because $\beta \neq \beta^*$ and $\dfrac{t_2}{t_1} \geq 1$; the limit of equation 6.3.1 can be simplified as

$$\lim_{t_2 \longrightarrow 0} \frac{f(t_1,t_2|\mu_1,\mu_2,\beta)}{f(t_1,t_2|\mu_1^*,\mu_2^*,\beta^*)} = \begin{cases} 0, & \text{if } \beta > \beta^* \\ \infty, & \text{if } \beta < \beta^* \end{cases} \tag{6.3.5}$$

which is a contradiction!

When $\mu_1 \neq \mu_1^*, \mu_2 \neq \mu_2^*, \beta \neq \beta^*$, the equation 6.3.1 is not true for $\forall\ 0 \leq t_1 \leq t_2 \leq \infty$.

Therefore, the equation 6.3.1 holds for $\forall\ 0 \leq t_1 \leq t_2 \leq \infty$ if and only if $\mu_1 = \mu_1^*, \mu_2 = \mu_2^*, \beta = \beta^*$.

□

6.4 Identifiability Issues Under Others Cases

Identifiability issues under unequal shape parameters and a known dependence parameter for a competing risks model has been shown in Chapter 4 as well as the identifiability issues un-

der different shape parameters when the dependence parameter is not known. In this section, we are showing these identifiability issues can be established under semi-competing risks data simultaneously. Based on the competing risk data we can obtain the crude survival functions $Q_1(t_1) = P(T_1 > t_1, T_1 \leq T_2)$ and $Q_2(t_2) = (T_2 > t_2, T_2 \leq T_1)$; by contrast in semi-competing risks scenario, we can observe $P(T_1 > t_1, T_2 > t_2, T_1 \leq T_2)$ and $P(T_2 > t_2, T_2 \leq T_1)$. In the other words, as long as $t_2 \longrightarrow \infty$ both crude survival functions $Q_1(t_1)$ and $Q_2(t_2)$ that are defined in competing risks can be maintained under semi-competing risks data. Since the proposed model is a parametric one, therefore the identifiability issues under different shape parameters and either a known dependence or an unknown dependence parameter can both be set up in semi-competing risks field.

6.5 Simulation Studies

The simulation studies only contain two competing risks with sample sizes $n = 1000$. The generated data set contains $\{(t_{1i}, d_{1i}, t_{2i}, d_{2i}), i = 1, ..., n\}$ with 10 repetitions, which is denoted by $\{(t_{1ik}, d_{1ik}, t_{2ik}, d_{2ik}), i = 1, ..., n;\ k = 1, ..., 10\}$, where t_{1ik}, t_{2ik} are the latent failure times to non-terminal and terminal events respectively and d_{jik} represents the indicator for censoring. According to Lemma 5.2.1, we generate the data as following: $T_1 = (U)^{\delta/\beta_1}(V/\mu_1)^{1/\beta_1}$,$T_2 = ([(1 - U)]^{\delta/\beta_2}(V/\mu_2)^{1/\beta_2}$, where $U \sim \text{Unif}(0,1)$, $M_\delta \sim \text{Bern}(\delta)$, $V_{11} \sim \text{Exp}(1)$ and $V_{12} \sim \text{Exp}(1)$ are independent, and $V = V_{11} + M_\delta V_{12}$. In this simulation study, there is no censoring involved. The results of simulation study based on 10 replications are shown in the following subsections.

6.5.1 Model with Equal Shape Parameters

Simulation results from analysis model with same shape parameters

$$S(t_1,t_2) = \exp\Big(-\Big[(\mu_1 t_1^{\beta})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta})^{\frac{1}{\delta}}\Big]^{\delta}\Big)$$

based on data generating model with $\mu_1 = 0.8, \mu_2 = 0.5, \beta_1 = 2, \beta_2 = 2$ and dependence parameters respectively equal to 0.3, 0.8 and 1 are summarized in Table 6.2.

Table 6.2: Simulation Results Under Models with Equal Shape Parameters

Analysis	Data Generating Model with $\mu_1 = 0.8, \mu_2 = 0.5, \beta_1 = 2, \beta_2 = 2$		
Model	$\delta = 0.3$	$\delta = 0.8$	$\delta = 1$
Parameter	Estimated Values: median (min,max)		
μ_1	0.760 (0.744, 0.822)	0.691 (0.638, 0.744)	0.759 (0.709, 0.785)
μ_2	0.486 (0.472, 0.537)	0.482 (0.457, 0.536)	0.505 (0.479, 0.549)
β_1	1.953 (1.905, 2.024)	1.952 (1.889, 2.091)	1.901 (1.842, 1.984)
β_2	1.979 (1.904, 2.028)	2.028 (1.951, 2.111)	1.971 (1.913, 2.074)
δ	0.34 (0.31,0.36)	0.845 (0.83,0.97)	1 (1,1)
-log(likelihood)	1151.6	1762.3	1793.5

Note: The estimation of each parameter is given by the minimum, the maximum and the median over 10 iterations.

According to Table 6.2, the dependent and independent case can be identified via the analysis model with same shape parameter, although there is some acceptable error with the MLE of shape and location parameters. The range of each parameter is narrow and close to its true value. The plot

of profile likelihood of delta is given by Figure 6.1, which indicates distinct identifiable minimum -(likelihood) to identify dependent and independent cases.

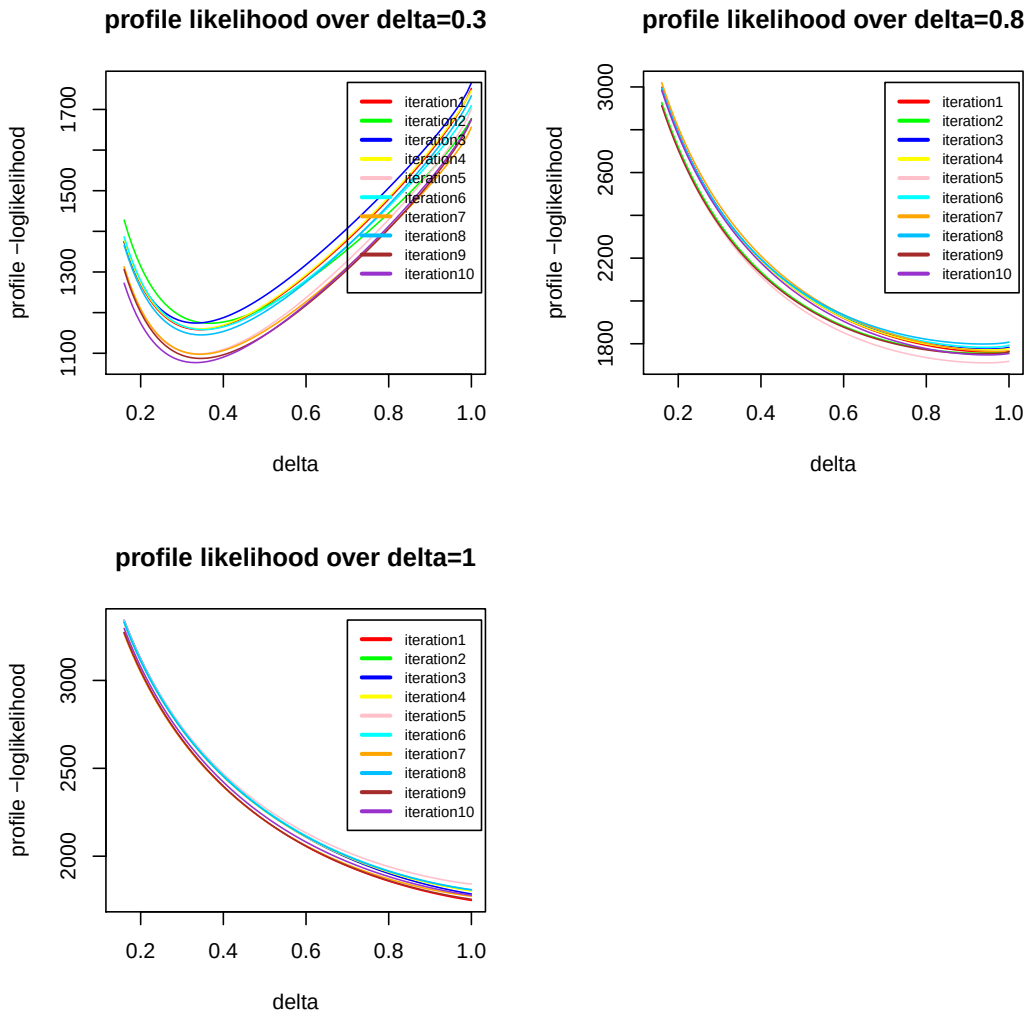


Figure 6.1: The profile likelihood plot for model with equal shape parameter

6.5.2 Model with Unequal Shape Parameters

Simulation results from analysis model with different shape parameters $S(t_1,t_2) = \exp\Big(-\Big[(\mu_1 t_1^{\beta_1})^{\frac{1}{\delta}} + (\mu_2 t_2^{\beta_2})^{\frac{1}{\delta}}\Big]^{\delta}\Big)$ based on data generating model with $\mu_1 = 0.8, \mu_2 = 0.5, \beta_1 = 1.1, \beta_2 = 2$ and dependence parameters respectively equal to 0.3, 0.8 and 1 are summarized in Table 6.3.

Table 6.3: Simulation Results Under Models with Unequal Shape Parameters

Analysis Model	Data Generating Model with unequal shape parameter		
	$\delta = 0.3$	$\delta = 0.8$	$\delta = 1$
Parameter	Estimated Values: median (min,max)		
μ_1	0.704 (0.653, 0.737)	0.786 (0.759, 0.809)	0.743 (0.716, 0.784)
μ_2	0.506 (0.465, 0.539)	0.493 (0.450, 0.510)	0.506 (0.469, 0.526)
β_1	1.057 (1.021, 1.100)	1.026 (1.006, 1.053)	1.066 (1.032, 1.103)
β_2	1.934 (1.893, 2.009)	2.005 (1.978, 2.052)	2.020 (1.964, 2.079)
δ	0.33 (0.31,0.35)	0.78 (0.77,0.82)	1 (1,1)
-log(likelihood)	1340.5	1877.9	1891.4

Note: The estimation of each parameter is given by the minimum, the maximum and the median over 10 iterations.

According to Table 6.3, the dependent and independent case can be identified via the analysis model with unequal shape parameters. The range of each parameter is close to its true value, although there is some acceptable error, especially associated with the MLE of μ_1, β_1. The plot of profile likelihood of delta is given by Figure 6.2, which indicates distinct identifiable minimum -(likelihood) to identify dependent and independent cases.

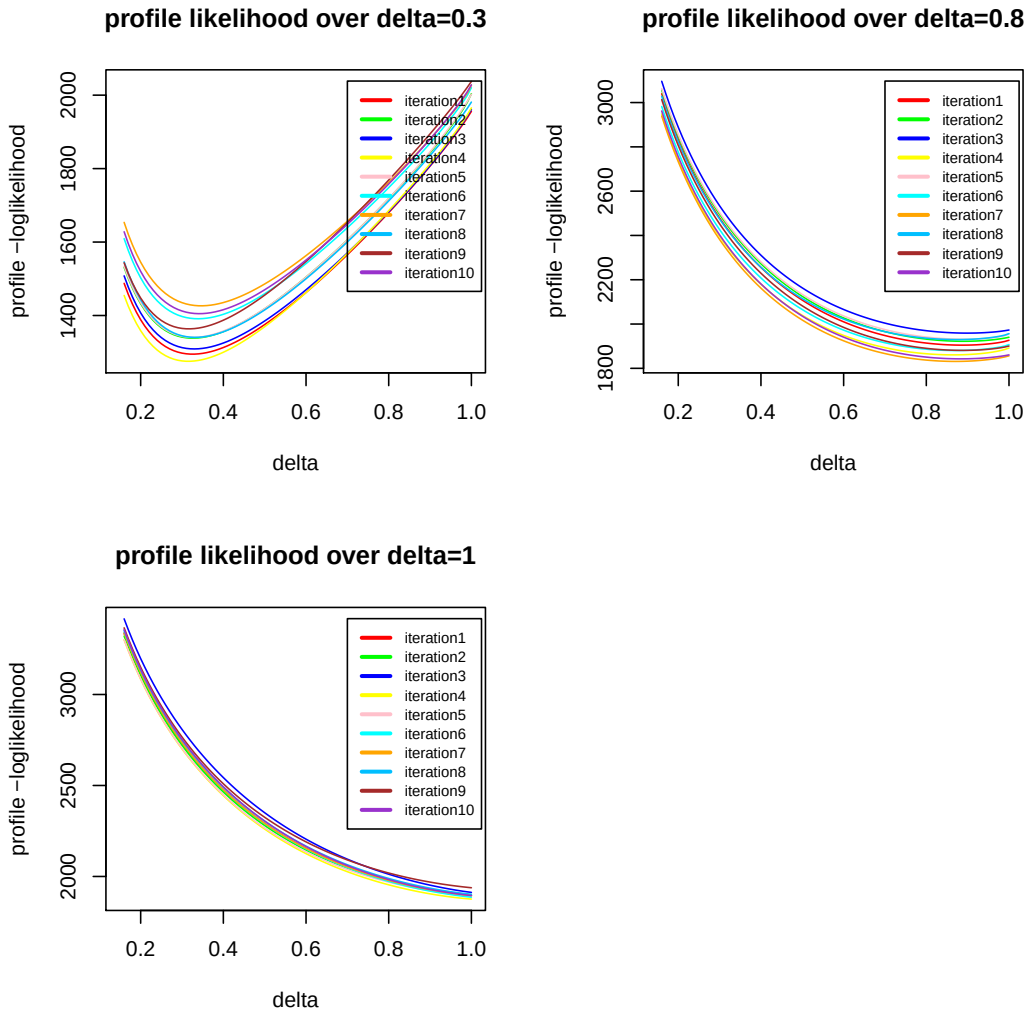


Figure 6.2: The profile likelihood plot for model with unequal shape parameter

6.6 Applications on Semi-Competing Risks

It is known that tuberculous pericarditis is associated with high morbidity and mortality even if antituberculosis therapy is administered. Mayosi et al. (2014) have evaluated the effects of adjunctive glucocorticoid therapy and Mycobacterium indicus pranii immunotherapy in patients with tuberculous pericarditis through a 2-by-2 factorial design. The data coordinating center for

this study was the Population Health Research Institute at Hamilton Health Sciences and McMaster University, Canada. The data we used for the analysis in this section is a subset of the full data set in Mayosi et al. (2014). It is well known that tuberculous pericarditis is a common cause of pericardial effusion, cardiac tamponade, and constrictive pericarditis in sub-Saharan Africa and parts of Asia. This trial data was collected in eight countries in Africa. It was conducted from January 2009 through February 2014 at 19 hospitals. All together, 1400 patients were enrolled for the comparison of prednisolone with placebo; 694 of them were assigned to receive placebo and 706 were in the group to receive prednisolone. In total, there were 1250 patients enrolled for comparison of Mycobacterium indicus pranii with placebo: half of the patients were assigned to the group to receive Mycobacterium indicus pranii and others were assigned to receive placebo.

6.6.1 Data Description

There are 443 eligible patients in the data subset we analyzed. The two treatments prednisolone (yes/no) and Mycobacterium indicus pranii (yes/no) were administered in a 2×2 factorial design in which the patients were randomized. The original hypotheses were there would be an overall benefit of adjunctive prednisolone for patients, and that intradermal Mycobacterium indicus pranii could be effective in suppressing inflammation and its sequela in patients with tuberculous pericarditis. The analysis in Mayosi et al. (2014) considered the first occurrence of any event without considering the aspect of competing risks. We consider a competing risk analysis of these data. We consider the first of tamponade and constriction as the non-terminal event time T_1 and time to death as latent failure time T_2. The corresponding censoring indicators are d_1 and d_2, where $d_j = 0$ means censoring and $d_j = 1$ for otherwise. The summary of demographic variables from this data is given in Table 6.4. The observed data set $\{(Y_{1i}, d_{1i}, Y_{2i}, d_{2i}), i = 1, \ldots, n\}$ is summarized in the Table 6.5.

Table 6.4: Characteristics of the Patients for Mycobacterium Data

Variable	Description
Age(yr)	38.36 ± 13.36
Gender	180(F)/263(M)
Weight(kg)	58.82 ± 13.01
Treatment = pred	212(Y)/231(N)
Treatment = myco	207(Y)/236(N)

Table 6.5: Mycobacterium Data Information

Observed Data Set			
	(Y_1, Y_2)	(d_1, d_2)	N
Constric/Tampo and censored prior to death	(T_{1i}, C_i)	$(1,0)$	28
Dead following Constric/Tampo	(T_{1i}, T_{2i})	$(1,1)$	20
Dead without Constric/Tampo	(T_{2i}, T_{2i})	$(0,1)$	71
Censored prior to Constric/Tampo or death	(C_i, C_i)	$(0,0)$	325

6.6.2 Data Analysis

We analyze the data using both a semi-competing risks model (DSCR model) and a competing risks model (DCR model). We further consider regression analysis under each of these models.

6.6.2.1 Semi-Competing Risks Model

According to the semi-competing risks model 6.2.2, the likelihood function has been derived in equation 6.2.5 as

$$\mathbf{L} = \prod_{i=1}^{n} f(t_{1i},t_{2i})^{d_{1i}d_{2i}} \cdot S_1^{'}(t_{1i},t_{2i})^{d_{1i}(1-d_{2i})} \cdot S_2^{'}(t_i,t_i)^{(1-d_{1i})d_{2i}} \cdot S(t_i,t_i)^{(1-d_{1i})(1-d_{2i})}$$

and we will use the same competing risks model as we proposed in Chapter 3 as expression (3.5.1) to obtain the log-likelihood function, which is given by

$$\log L = \sum_{i=1}^{n} \left(d_{1i} \log f_1(t_i) + d_{2i} \log f_2(t_i) + (1 - d_{1i} - d_{2i}) \log S(t_i,t_i) \right)$$

The maximum likelihood estimator for semi-competing risks model and competing risks model are given in Tables 6.6 and 6.7 as below.

Table 6.6: MLE Under DSCR Model for Mycobacterium Data

Parameter	μ_1	μ_2	$\beta 1$	β_2	δ	-log(Likelihood)
mean	0.115	0.160	0.388	0.495	0.7	389.727

Table 6.7: MLE Under DCR Model for Mycobacterium Data

Parameter	μ_1	μ_2	$\beta 1$	β_2	δ	-log(Likelihood)
mean	0.224	0.226	0.452	0.456	0.01	356.943

Note that the data used to obtain MLEs under DCR model and DSCR model are different. The former one uses the data as competing risks, whereas we only obtain the minimum. The data

for DSCR model contains additional information which includes the observations for death after relapse; i.e., if relapse occurs first to the patient, we will continue the observation on him/her.

We also conduct a parallel Bayesian analysis on the same data with prior distributions $\mu_1 \sim \text{Gam}(1,1)$, $\mu_2 \sim \text{Gam}(1,1)$, $\beta_1 \sim \text{Gam}(0.01, 2.5)$, $\beta_2 \sim \text{Gam}(0.01, 2.5)$ and $\delta \sim \pi \cdot \text{Unif}(0,1) + (1-\pi) \cdot \text{degenerate}(1)$ with $\pi \sim \text{Unif}(0,1)$. We ran 60000 MCMC iterations respectively for semi-competing risks model and competing risks model and the posterior means of μ_1, μ_2, β_1, $\beta 2$ and δ which were computed from the last 50000 iterations as shown in Tables 6.8 and 6.9.

Table 6.8: Posterior Estimation Under DSCR Model for Mycobacterium Data

Parameter	μ_1	μ_2	$\beta 1$	β_2	δ
mean	0.116	0.161	0.392	0.498	0.709
median	0.115	0.160	0.390	0.496	0.708
95% CI	(0.086, 0.151)	(0.128, 0.198)	(0.301, 0.490)	(0.408, 0.599)	(0.589, 0.828)

Table 6.9: Posterior Estimation Under DCR Model for Mycobacterium Data

Parameter	μ_1	μ_2	$\beta 1$	β_2	δ
mean	0.122	0.159	0.384	0.514	0.718
median	0.112	0.152	0.393	0.510	1
95% CI	(0.074, 0.205)	(0.107, 0.234)	(0.298, 0.495)	(0.417, 0.630)	(0, 1)

The estimates from maximum likelihood and Bayesian approaches are similar for semi-competing risks model and similar for competing risks model except for the estimator of δ. The posterior mean of deviance for DSCR model is 784.5, and we can compute -log(Likelihood) = 392.25. The posterior mean of deviance for DCR model is 723.2 and we can obtain the corresponding

-log(Likelihood) = 361.6. We note here that the semi-competing risks model utilizes an additional level of data (time to terminal event following non-terminal event) and hence the log-likelihoods of the two models are not directly comparable. Figure 6.3 shows the posterior density of δ for DCR model using Mycobacterium data. According to the density plot, the likelihood seems to be flat, which could be the reason why the MLE and Bayesian estimation of δ are different in DCR model.

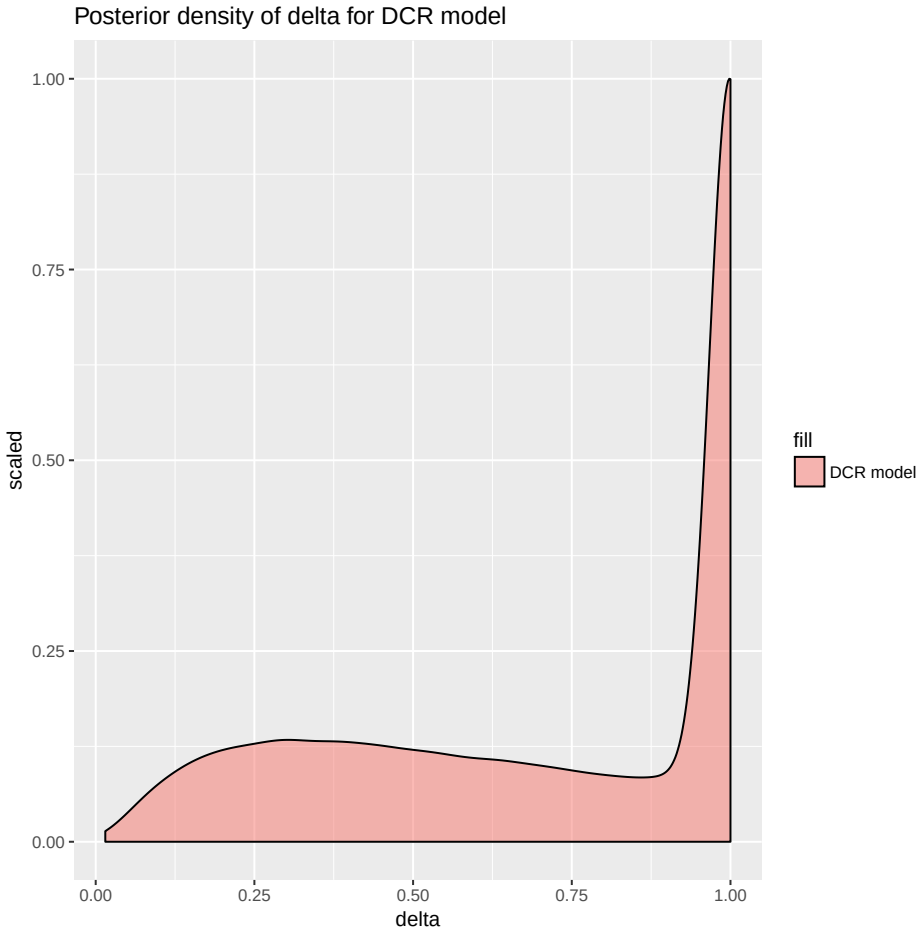


Figure 6.3: The posterior density of δ for DCR model by using Mycobacterium data

The Figure 6.4 gives the plot of cumulative incidence functions for dependent semi-competing risks and dependent competing risks models. Compared with DCR model, the DSCR model presents not only CIF of Constric/Tampo first or death first, but also the CIF of death after Constric/Tampo, which is shown in red solid line.

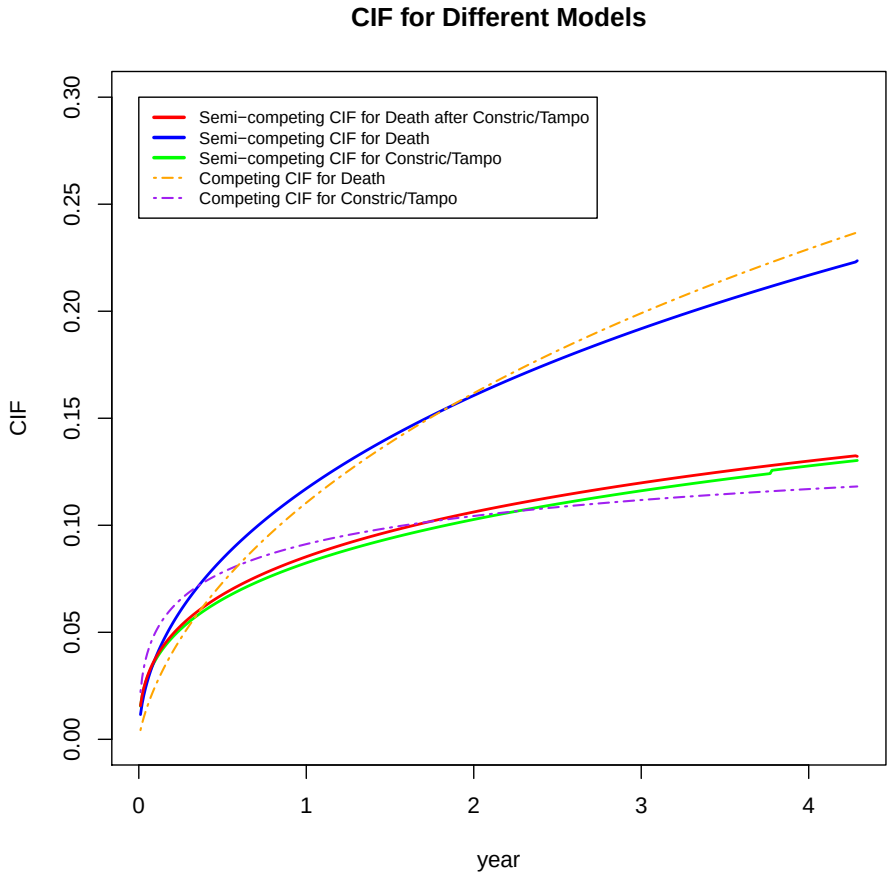


Figure 6.4: CIF Plot for DSCR Model and DCR Model

6.6.2.2 Semi-Competing Risks Regression Model

We next analyzed the same data in regression framework including covariate information. The regression structure is introduced as $\log(\mu_j^{1/\delta}) = X\eta_j$ for $j = 1,2$ in semi-competing risks model (DSCRR) and similarly in the competing risks model (DCRR model). The covariates which have been used for analysis contain age, gender, weight, pred (whether treatment prednisolone is used or not) and myco (whether treatment Mycobacterium is used or not). The MLE of DSCRR is given by Table 6.10.

Table 6.10: MLE Under DSCRR Model for Mycobacterium Data

Parameter	β1	β_2	δ	intercept1	age1	gender1
mean	0.401	0.492	0.69	-0.288	-0.01	-0.781
Parameter	weight1	pred1	myco1	pred1*myco1	intercept2	age2
mean	-0.356	-0.089	-0.114	-0.525	-0.540	-0.006
Parameter	gender2	weight2	pred2	myco2	pred2*myco2	-logL
mean	0.077	-0.302	0.020	-0.243	-0.055	191.15

We again consider a parallel Bayesian analysis with prior distributions that have been assigned as $\eta_1's \sim \text{Normal}(0,5)$, $\eta_2's \sim \text{Normal}(0,5)$, $\beta_1 \sim \text{Gam}(0.01, 2.5)$, $\beta_2 \sim \text{Gam}(0.01, 2.5)$ and $\delta \sim \pi \cdot \text{Unif}(0,1) + (1-\pi) \cdot \text{degenerate}(1)$ with $\pi \sim \text{Unif}(0,1)$. After specifying the prior distributions, we ran 60000 MCMC iterations and looked at posterior means calculated from the last 50000 iterations of the respective parameters involved in the model. The posterior estimates of β_1, β_2, δ and $\eta's$ are shown in Table 6.11; the posterior mean of deviance is 391.

Table 6.11: Posterior Estimation Under DSCRR Model for Mycobacterium Data

Parameter	$\beta 1$	β_2	δ	intercept1	age1
mean	0.392	0.488	0.66	-2.063	0.010
95% CI	(0.30, 0.49)	(0.40, 0.59)	(0.54, 0.79)	(-3.93, 0.34)	(-0.02, 0.04)
Parameter	gender1	weight1	pred1	myco1	pred1*myco1
mean	-0.927	-0.187	-0.577	0.286	-0.402
95% CI	(-1.77, -0.12)	(-0.48, -0.10)	(-1.57, 0.37)	(-0.59, 1.19)	(-1.81, 1.01)
Parameter	intercept2	age2	gender2	weight2	pred2
mean	-1.438	0.005	0.091	-2.62	0.051
95% CI	(-3.11, 0.23)	(-0.02, 0.03)	(-0.54, 0.72)	(-0.54, -0.004)	(-0.72, 0.83)
Parameter	myco2	pred2*myco2			
mean	-0.144	-0.303			
95% CI	(-0.95, 0.64)	(-1.50, 0.83)			

According to Table 6.11, weight of the patient impacts the event Constric/Tampo and death. The credible intervals corresponding to weight does not contain zero and hence can be termed as having "significant" influence on the event Constric/Tampo and death. Heavier weight leads to better survival. And gender is a "significant" difference for Constric/Tampo; females have higher chance to get remission of Constric/Tampo. Figure 6.5 shows the cumulative incidence functions of DSCRR model by gender; there is significant difference between blue solid line (Constric/Tampo in females) and orange dot line (Constric/Tampo in males). The chance to relapse in males is much higher than it in females.

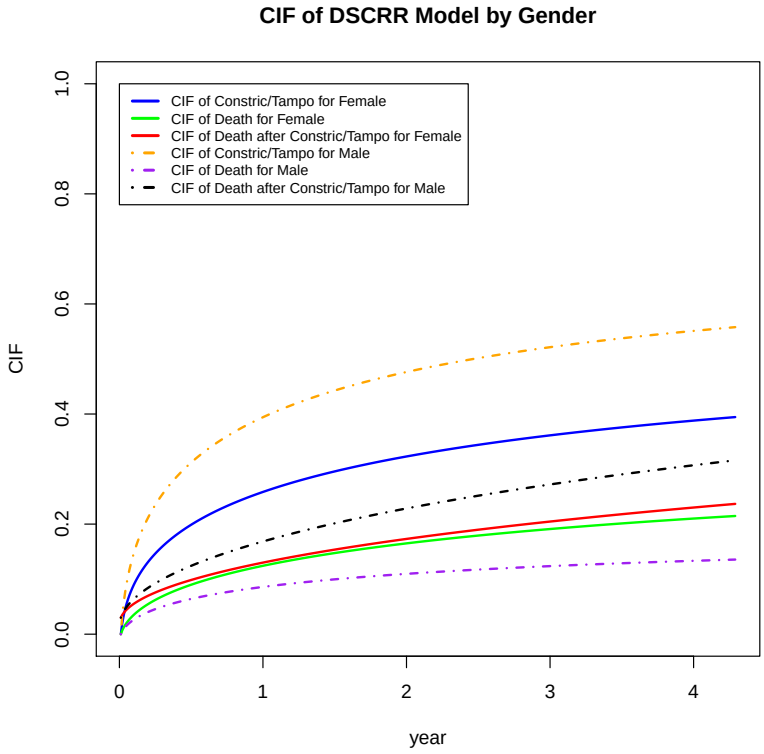


Figure 6.5: CIF of DSCRR model for Mycobacterium data by gender

Tables 6.12 and 6.13 present the estimation from frequentist inference and Bayesian approach of DCRR model; the posterior mean of deviance is 709.

Table 6.12: MLE Under DCRR Model for Mycobacterium Data

Parameter	$\beta 1$	β_2	δ	intercept1	age1	gender1
mean	0.486	0.498	0.02	3.480	-0.772	3.415
Parameter	weight1	pred1	myco1	pred1*myco1	intercept2	age2
mean	-8.165	3.426	-2.073	2.463	4.104	-0.778
Parameter	gender2	weight2	pred2	myco2	pred2*myco2	-log
mean	5.347	-8.068	3.135	1.086	3.837	351.95

Table 6.13: Posterior Estimation Under DCRR Model for Mycobacterium Data

Parameter	β1	β_2	δ	intercept1	age1
mean	0.435	0.446	0.024	-0.968	-1.171
95% CI	(0.36, 0.51)	(0.37, 0.52)	(0.017, 0.035)	(-4.21, 2.17)	(-1.44, -0.91)
Parameter	gender1	weight1	pred1	myco1	pred1*myco1
mean	-1.146	-2.65	-0.858	-0.313	-0.735
95% CI	(-4.24, 2.04)	(-5.34, -0.87)	(-3.94, 2.23)	(-3.53, 2.82)	(-3.75, 2.39)
Parameter	intercept2	age2	gender2	weight2	pred2
mean	-0.543	-1.172	-0.024	-2.581	-0.397
95% CI	(-3.61, 2.68)	(-1.45, -0.91)	(-3.12, 3.13)	(-5.27, -0.78)	(-3.56, 2.64)
Parameter	myco2	pred2*myco2			
mean	-0.878	-0.558			
95% CI	(-4.02, 2.26)	(-3.64, 2.49)			

According to Table 6.13, the credible intervals corresponding to weight and age do not contain zero and hence can be termed as having "significant" influence on the non-terminal event Constric/Tampo and terminal event death. Heavier weight and older age lead the patient to better survival. There is no evidence to show that treatments prednisolone and Mycobacterium indicus pranii are effective to the patients. Note that the data used to obtain MLE under DCRR model and DSCRR model are different. The former one uses the data as competing risks, whereas we only observe the minimum. The data for DSCRR model contains additional information that includes the observations for death after Constric/Tampo; i.e., if Constric/Tampo occurs first, we could observe death after that as well.

CHAPTER 7

FUTURE EXTENSIONS AND CONCLUSIONS

7.1 Introduction

Interval censoring in the context of competing risks has been receiving increasing attention in the literature; see Sun (2007) and Mao et al. (2016). Another important question is variable selection in the context of competing risks. We describe our proposed future research directions in these two important areas in this chapter. In addition, goodness-of-fit of the proposed model could be conducted as well as the model validation.

7.2 Interval-Censored Failure Time in Competing Risks

Interval censoring is a common feature in the study of failure time data. Interval censoring occurs when the failure time of interest is not observed continuously, and as a consequence, the failure time is not observed exactly. Instead, only an interval within which the event of interest occurred is observed. For example, it is common to have pre-specified regular visit schedule for a patient and a changed status may be observed in one such visit. It is obvious that the true event time is greater than the last observation time at which the change has not occurred and less than or equal to the first observation time at which the change has been observed to occur. That is, the interval contains the real but unobserved time of occurrence of the event.

Another special case of interval-censored data is called current status data, see Jewell & van der Laan (1996) and Sun & Kalbfleisch (1993). For this type of censoring, the status of the occurrence

of the event of interest is observed only once for each subject. That is, although the failure time is not observed directly, the information on occurrence of the event of interest at the observation time is known. Thus, the failure time is either left censored or right censored.

Suppose a study that involves n independent subjects and every subject may experience the event of interest whose time is denoted as T and it is not observed. Instead, we know that T lies in an interval if the subject experiences the event of interest during the study period (i.e., if the subject experiences a failure) or that T lies somewhere beyond the largest observation time if the subject does not experience the event during the study period (i.e., if the subject is right censored). Specifically, let J denote the random number of observation times for a subject, and let $U_1 < U_2 < ... < U_J$ denote a random sequence of the observation times. Set $U_0 = 0$ and $U_{J+1} = \infty$. Define a vector of indicators for a subject, $\Delta = (\triangle_1, ..., \triangle_J)$, which contains the information about the specific interval in which the subject experienced the event, where $\triangle_j$ $(j = 1, ..., J)$ is given by

$$\triangle_j = \begin{cases} 1 & \text{if } U_{j-1} < T \leq U_j \\ 0 & \text{o/w} \end{cases}$$

and the indicator Δ_{J+1} gives the information whether the subject is right censored or not, which is defined as

$$\Delta_{J+1} = 1 - \sum_{j=1}^{J} \triangle_j$$

In addition, we assume that there exists K competing risks in the studies and $T_1, ..., T_K$ denote the latent failure times. The unobserved failure time $T = \min(T_1, ..., T_K)$. Let $d_j (j = 1, ..., K)$ indicate, by the values 1 versus 0, whether or not the failure is caused by risk j. Then the observed data for a random sample of n subjects consist of $\{(J_i, \mathbf{U_i}, \Delta_i, d_{ij}), \text{for } i = 1, ..., n;\ j = 1, ..., K\}$, where $\mathbf{U_i} = (U_{i0}, U_{i1}, ..., U_{iJ_i})$ and $\Delta_i = (\triangle_{i1}, ..., \triangle_{iJ_i})$.

Mao et al. (2016) has derived the likelihood of a general class of semiparametric regression models for interval-censored competing risks with potentially time-varying covariates, which is given by

$$
\begin{aligned}
L_n(\beta,\Lambda) = \prod_{i=1}^{n} \Bigg[\prod_{j=1}^{J_i} \prod_{k=1}^{K} \Bigg(& \exp\left[-G_k\left\{ \int_0^{U_{ij-1}} \exp(\beta_{\mathbf{k}}^{\mathbf{T}} Z_i(t)) d\Lambda_k(t) \right\} \right] \\
& - \exp\left[-G_k\left\{ \int_0^{U_{ij}} \exp(\beta_{\mathbf{k}}^{\mathbf{T}} Z_i(t)) d\Lambda_k(t) \right\} \right] \Bigg)^{I(d_{ij}=1,\triangle_{ij}=1)} \\
& \times \left\{ \sum_{k=1}^{K} \left(\exp\left[-G_k\left\{ \int_0^{U_{ij-1}} \exp(\beta_{\mathbf{k}}^{\mathbf{T}} Z_i(t)) d\Lambda_k(t) \right\} \right] \right.\right. \\
& \left.\left. - \exp\left[-G_k\left\{ \int_0^{U_{ij}} \exp(\beta_{\mathbf{k}}^{\mathbf{T}} Z_i(t)) d\Lambda_k(t) \right\} \right] \right) \right\}^{I(d_{ij}=0,\triangle_{ij}=1)} \\
& \times \left(\sum_{k=1}^{K} \exp\left[-G_k\left\{ \int_0^{U_{iJ_i}} \exp(\beta_{\mathbf{k}}^{\mathbf{T}} Z_i(t)) d\Lambda_k(t) \right\} \right] - K + 1 \right)^{I(\Delta_i=\mathbf{0})} \Bigg]
\end{aligned}
$$

where $G_k(\cdot)$ is a known increasing function, $\beta_{\mathbf{k}}$ is a set of regression parameters, $\Lambda_k(\cdot)$ is an arbitrary increasing function with $\Lambda_k(0) = 0$, $I(\cdot)$ is the indicator function and $Z(\cdot)$ denotes a p-vector of possibly time-varying external covariates. The nonparametric maximum likelihood estimation is given in the numerical algorithm section.

Instead of using a general class of semiparametric regression models for competing risks, we propose to consider the proposed parametric dependent competing risks model for interval-censored data and propose to develop likelihood and Bayesian inference for this model in future work.

7.3 Variable Selection on Competing Risks Models

Variable selection remains an important problem in competing risks regression modeling. We will discuss the non-Bayesian way to achieve variable selection in this section.

7.3.1 Penalized Variable Selection in Competing Risks

Penalized variable selection methods for Cox's proportional hazards model and frailty model have been developed by Fan & Li (2002). It is known that penalization method cannot be directly applied to general competing risks model; instead, the proportional hazards model with covariate can be estimated through such methods. Z. Fu et al. (2016) have also proposed a penalization method for variable selection and model estimation of the proportional subdistribution hazards (PSH) model.

Consider the proposed parametric model in Chapter3:

$$S(t_1,t_2) = \exp\left(-\left[(\mu_1 t_1^{\beta_1})^{1/\delta} + (\mu_2 t_2^{\beta_2})^{1/\delta} \right]^{\delta} \right) \tag{7.3.1.1}$$

After setting up the regression structure with using $\mathbf{X}$, the model can be expressed as a mixed proportional hazards competing risk model under a given covariate $\mathbf{X}$ in the representation via a frailty approach 4.4.7,4.4.8, that is,

$$\begin{aligned} h_1(t_1|\beta_1,\mu_1,\delta,z) &= g_1(\mathbf{X}'\eta_{\mathbf{t_1}})\cdot h_{0_1}(t_1|\beta_1,\eta_{0t_1},\delta)\cdot z \\ h_2(t_2|\beta_2,\mu_2,\delta,z) &= g_2(\mathbf{X}'\eta_{\mathbf{t_2}})\cdot h_{0_2}(t_2|\beta_2,\eta_{0t_2},\delta)\cdot z \end{aligned}$$

Then under the Laplace transform $\psi(s) = E[\exp(-sZ)]$, the joint survival function of T_1 and T_2 is given by

$$S(t_1,t_2|\mathbf{X}) = \psi\Big[\psi^{-1}\big(S_1(t_1|\mathbf{X})\big) + \psi^{-1}\big(S_2(t_2|\mathbf{X})\big)\Big] \tag{7.3.1.2}$$

where ψ^{-1} is the inverse function of ψ, and $S_1(t_1|\mathbf{X})$, $S_2(t_2|\mathbf{X})$ are corresponding marginal survival functions, which are given by

$$S_1(t_1|\mathbf{X}) = \psi[g_1(\mathbf{X}'\eta_{\mathbf{t_1}}) \cdot H_{0_1}(t_1|\beta_1, \eta_{0t_1}, \delta)] \tag{7.3.1.3}$$

and

$$S_2(t_2|\mathbf{X}) = \psi[g_2(\mathbf{X}'\eta_{\mathbf{t_2}}) \cdot H_{0_2}(t_2|\beta_2, \eta_{0c_2}, \delta)] \tag{7.3.1.4}$$

The penalized log-likelihood function is

$$\sum_{i=1}^{n} d_{i1} * f_1(t_i|\mathbf{x_i}) + d_{i2} * f_2(t_i|\mathbf{x_i}) + (1 - d_{i1} - d_{i2}) * S(t_i, t_i|\mathbf{x_i}) - n \sum_{k=1}^{2d} p_\lambda(|\eta_k|) \tag{7.3.1.5}$$

where η_k is one of $\eta = (\eta_{0t_1}, \eta_{1t_1}, ..., \eta_{dt_1}, \eta_{0t_2}, \eta_{1t_2}, ..., \eta_{dt_2})$; the marginal density can be obtained from $f_j(t_j|\mathbf{x}) = \frac{\partial S_j(t_j|\mathbf{x})}{\partial t_j}$, and $p_\lambda(|\eta_k|)$ is the penalty function. λ is a tuning parameter that controls the complexity of select models. Maximizing 7.3.1.5 with respect to $(\beta_1, \beta_2, \delta, \eta)$ yields the maximum penalized likelihood estimator. Various penalty functions $p_\lambda(|\eta_k|)$ have been proposed, such as LASSO which refers to Tibshirani (1996), SCAD which refers to Fan & Li (2001) and ALASSO which refers to Zou (2006). The estimators may vary for different penalty functions.

Another way to obtain the penalized likelihood estimator is to maximize the penalized log-likelihood function with components of partial likelihood function, instead of a complete likelihood function in 7.3.1.5, and penalty functions with respect to $(\beta_1, \beta_2, \delta, \eta)$. This approach is considered in Fan & Li (2002).

7.4 Conclusions

In this dissertation, we have proposed a model for dependent competing risks. This model could be motivated via a frailty approach as well as a copula approach. We have provided a detailed review of the literature on statistical methods for competing risks data. We have proposed a dependent competing risks model and its identifiability properties have been discussed. In Chapter 4, we have shown that our proposed dependent competing risks model is identifiable under different shape parameters and a known dependence parameter. In addition, the identifiability property of the model for unknown dependence parameter has also been explored.

This proposed model can be extended to incorporate covariates while maintaining its identifiability properties. We have considered extensive simulation studies under this model. Statistical inference for this model is obtained via maximum likelihood and via Markov chain sampling within a Bayesian framework. We have applied the proposed dependent competing risks model, both with and without a regression structure, to analyze competing risks data of breast cancer patients treated with Tamoxifen.

In addition, we have extended the proposed model to the case of semi-competing risks data. Compared with competing risks data, the semi-competing risks data contains additional information on survival status of the patients, as the terminal event could be observed after the occurrence of the non-terminal event. In Chapter 6, we have shown the identifiability properties under semi-competing risks data. We applied the proposed approach to a semi-competing risks data collected in Africa by the Population Health Research Institute (PHRI).

REFERENCES

Aalen, O. (1976). Nonparametric inference in connection with multiple decrement models. *Scandinavian Journal of Statistics*, 15–27.

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

Altshuler, B. (1970). Theory for the measurement of competing risks in animal experiments. *Mathematical Biosciences*, *6*, 1–11.

Basu, A. P., & Ghosh, J. (1978). Identifiability of the multinormal and other distributions under competing risks model. *Journal of Multivariate Analysis*, *8*(3), 413–429.

Basu, A. P., & Klein, J. P. (1982). Some recent results in competing risks theory. *Lecture Notes-Monograph Series*, *2*, 216–229.

Basu, S., Sen, A., & Banerjee, M. (2003). Bayesian analysis of competing risks with partially masked cause of failure. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *52*(1), 77–93.

Breslow, N. (1974). Covariance analysis of censored survival data. *Biometrics*, 89–99.

Chen, M.-H., Ibrahim, J. G., & Sinha, D. (2002). Bayesian inference for multivariate survival data with a cure fraction. *Journal of Multivariate Analysis*, *80*(1), 101–126.

Chiang, C. L. (1970). Competing risks and conditional probabilities. *Biometrics*, 767–776.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, *65*(1), 141–151.

Cook, R. J., & Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in medicine*, *16*(8), 911–924.

Cornfield, J. (1957). Estimation of the probability of developing a disease in the presence of competing risks. *American Journal of Public Health and the Nations Health*, *47*(5), 601–607.

Cox, D. R. (1959). The analysis of exponentially distributed life-times with two types of failure. *Journal of the Royal Statistical Society. Series B (Methodological)*, 411–421.

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

Cox, D. R. (1975). Partial likelihood. *Biometrika*, *62*(2), 269–276.

Cox, D. R., & Oakes, D. (1984). *Analysis of survival data* (Vol. 21). CRC Press.

Crowder, M. (1991). On the identifiability crisis in competing risks analysis. *Scandinavian Journal of Statistics*, 223–233.

Crowder, M. (2001). *Classical competing risks*. CRC Press.

David, H. A., & Moeschberger, M. L. (1978). *The theory of competing risks* (No. 39). Griffin London.

Day, R., Bryant, J., & Lefkopoulou, M. (1997). Adaptation of bivariate frailty models for prediction, with application to biological markers as prognostic indicators. *Biometrika*, *84*(1), 45–56.

Duchateau, L., & Janssen, P. (2007). *The frailty model*. Springer Science & Business Media.

Ebrahimi, N., Molefe, D., & Ying, Z. (2003). Identifiability and censored data. *Biometrika*, *90*(3), 724–727.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, *96*(456), 1348–1360.

Fan, J., & Li, R. (2002). Variable selection for cox's proportional hazards model and frailty model. *Annals of Statistics*, 74–99.

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

Flinn, C. J., & Heckman, J. J. (1982). *Are unemployment and out of the labor force behaviorally distinct labor force states?* National Bureau of Economic Research Cambridge, Mass., USA.

Fu, H., Wang, Y., Liu, J., Kulkarni, P. M., & Melemed, A. S. (2013). Joint modeling of progression-free survival and overall survival by a bayesian normal induced copula estimation model. *Statistics in medicine*, *32*(2), 240–254.

Fu, Z., Parikh, C. R., & Zhou, B. (2016). Penalized variable selection in competing risks regression. *Lifetime data analysis*, 1–24.

Fyles, A. W., McCready, D. R., Manchul, L. A., Trudeau, M. E., Merante, P., Pintilie, M., ... Olivotto, I. A. (2004). Tamoxifen with or without breast irradiation in women 50 years of age or older with early breast cancer. *New England Journal of Medicine*, *351*(10), 963–970.

Gail, M. (1975). A review and critique of some models used in competing risk analysis. *Biometrics*, 209–222.

Genest, C., Ghoudi, K., & Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, *82*(3), 543–552.

Genest, C., & Nešlehová, J. (2007). A primer on copulas for count data. *Astin Bulletin*, *37*(02), 475–515.

Ghosh, D. (2006). Semiparametric inferences for association with semi-competing risks data. *Statistics in medicine*, *25*(12), 2059–2070.

Ghosh, D., & Lin, D. (2000). Nonparametric analysis of recurrent events and death. *Biometrics*, *56*(2), 554–562.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, *55*(292), 698–707.

Heckman, J. J., & Honoré, B. E. (1989). The identifiability of the competing risks model. *Biometrika*, *76*(2), 325–330.

Holt, J. (1978). Competing risk analyses with special reference to matched pair experiments. *Biometrika*, *65*(1), 159–165.

Hougaard, P. (1986). A class of multivanate failure time distributions. *Biometrika*, *73*(3), 671–678.

Hsieh, J.-J., Wang, W., & Adam Ding, A. (2008). Regression analysis based on semicompeting risks data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *70*(1), 3–20.

Ibrahim, J. G., Chen, M.-H., & Sinha, D. (2005). *Bayesian survival analysis*. Wiley Online Library.

Jewell, N. P., & van der Laan, M. (1996). Generalizations of current status data with applications. In *Lifetime data: models in reliability and survival analysis* (pp. 141–148). Springer.

Jiang, H., Chappell, R., & Fine, J. P. (2003). Estimating the distribution of nonterminal event time in the presence of mortality or informative dropout. *Controlled clinical trials*, *24*(2), 135–146.

Jiang, H., Fine, J. P., & Chappell, R. (2005). Semiparametric analysis of survival data with left truncation and dependent right censoring. *Biometrics*, *61*(2), 567–575.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. CRC Press.

Kalbfleisch, J., & Prentice, R. (1973). Marginal likelihoods based on cox's regression and life model. *Biometrika*, *60*(2), 267–278.

Kalbfleisch, J., & Prentice, R. (1980). The statistical analysis of failure time data. new york: John wiley & sons. *KalbfleischThe Statistical Analysis of Failure Time Data1980*.

Kalbfleisch, J., & Prentice, R. (2011). *The statistical analysis of failure time data* (Vol. 360). John Wiley & Sons.

Klein, J. P., & Moeschberger, M. (1988). Bounds on net survival probabilities for dependent competing risks. *Biometrics*, 529–538.

Klein, J. P., & Moeschberger, M. L. (2005). *Survival analysis: techniques for censored and truncated data*. Springer Science & Business Media.

Lee, K. H., Haneuse, S., Schrag, D., & Dominici, F. (2015). Bayesian semiparametric analysis of semicompeting risks data: investigating hospital readmission after a pancreatic cancer diagnosis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *64*(2), 253–273.

Lee, K. H., Rondeau, V., & Haneuse, S. (2017). Accelerated failure time models for semi-competing risks data in the presence of complex censoring. *Biometrics*.

Lee, L. (1979). Multivariate distributions having weibull properties. *Journal of Multivariate analysis*, *9*(2), 267–277.

Mao, L., Lin, D., & Zeng, D. (2016). Semiparametric regression analysis of interval-censored competing risks data. *arXiv preprint arXiv:1602.01788*.

Marshall, A. W., & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, *62*(317), 30–44.

Mayosi, B. M., Ntsekhe, M., Bosch, J., Pandie, S., Jung, H., Gumedze, F., ... others (2014). Prednisolone and mycobacterium indicus pranii in tuberculous pericarditis. *New England Journal of Medicine*, *371*(12), 1121–1130.

Moeschberger, M., & David, H. (1971). Life tests under competing causes of failure and the theory of competing risks. *Biometrics*, 909–933.

Nelsen, R. B. (1999). *An introduction to copulas, volume 139 of lecture notes in statistics.* Springer-Verlag, New York.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, *84*(406), 487–493.

Peng, L., & Fine, J. P. (2007). Regression modeling of semicompeting risks data. *Biometrics*, *63*(1), 96–108.

Pepe, M. S., & Mori, M. (1993). Kaplan—meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Statistics in medicine*, *12*(8), 737–751.

Pintilie, M. (2006). *Competing risks: a practical perspective* (Vol. 58). John Wiley & Sons.

Prentice, R. L., & Breslow, N. (1978). Retrospective studies and failure time models. *Biometrika*, *65*(1), 153–158.

Prentice, R. L., Kalbfleisch, J. D., Peterson Jr, A. V., Flournoy, N., Farewell, V., & Breslow, N. (1978). The analysis of failure times in the presence of competing risks. *Biometrics*, 541–554.

Rivest, L.-P., & Wells, M. T. (2001). A martingale approach to the copula-graphic estimator for the survival function under dependent censoring. *Journal of Multivariate Analysis*, *79*(1), 138–155.

Schick, G. J., & Wolverton, R. W. (1978). An analysis of competing software reliability models. *IEEE Transactions on Software Engineering*(2), 104–120.

Schweizer, B., & Sklar, A. (1974). Operations on distribution functions not derivable from operations on random variables. *Studia Mathematica*, *52*(1), 43–52.

Schweizer, B., & Sklar, A. (2011). *Probabilistic metric spaces*. Courier Corporation.

Shih, J. H., & Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, 1384–1399.

Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8.

Sklar, M. (1973). *Sklar aphasia scale*. Western Psychological Services.

Slud, E. V., & Rubinstein, L. V. (1983). Dependent competing risks and summary survival curves. *Biometrika*, *70*(3), 643–649.

Sun, J. (2007). *The statistical analysis of interval-censored failure time data*. Springer Science & Business Media.

Sun, J., & Kalbfleisch, J. D. (1993). The analysis of current status data on point processes. *Journal of the American Statistical Association*, *88*(424), 1449–1454.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, *72*(1), 20–22.

Wang, A., Chandra, K., Xu, R., & Sun, J. (2015). The identifiability of dependent competing risks models induced by bivariate frailty models. *Scandinavian Journal of Statistics*, *42*(2), 427–437.

Wang, W. (2003). Estimating the association parameter for copula models under dependent censoring. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(1), 257–273.

Wei, L. (1992). The accelerated failure time model: a useful alternative to the cox regression model in survival analysis. *Statistics in medicine*, *11*(14-15), 1871–1879.

Xu, J., Kalbfleisch, J. D., & Tai, B. (2010). Statistical analysis of illness–death processes and semicompeting risks data. *Biometrics*, *66*(3), 716–725.

Ye, Y., Kalbfleisch, J. D., & Schaubel, D. E. (2007). Semiparametric analysis of correlated recurrent and terminal events. *Biometrics*, *63*(1), 78–87.

Zacks, S. (1971). *The theory of statistical inference* (Vol. 34). Wiley New York.

Zheng, M., & Klein, J. P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, *82*(1), 127–138.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418–1429.