INFERENCE ON CONDITIONAL QUANTILE RESIDUAL LIFE FOR CENSORED SURVIVAL DATA

by

Wen-Chi Wu

B.S. in Statistics, Tamkang University, Taiwan, 2005

M.S. in Biostatistics, University of Pittsburgh, 2008

Submitted to the Graduate Faculty of

the Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2014

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Wen-Chi Wu

It was defended on

July 18, 2014

and approved by

Dissertation Advisor:

Jong-Hyeon Jeong, Ph.D.
Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Committee Members:

Abdus S. Wahed, Ph.D.
Associate Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Ruosha Li, Ph.D.
Assistant Professor
Department of Biostatistics
Graduate School of Public Health
University of Pittsburgh

Yu Cheng, Ph.D.
Associate Professor
Department of Statistics
Dietrich School of Arts and Sciences
University of Pittsburgh

INFERENCE ON CONDITIONAL QUANTILE RESIDUAL LIFE FOR CENSORED SURVIVAL DATA

Wen-Chi Wu, PhD

University of Pittsburgh, 2014

For randomly censored data, the residual life function at a given time determines a life distribution of a subject survived up to that time point. In the situation where the data are censored, or where the underlying distribution is skewed, the quantile residual life function is preferred. A number of studies regarding the quantile residual lifetime have been conducted in the univariate settings by many professionals. However, when a pair of units are observed, i.e. a study of twins, or when patients experience two types of events, i.e. time to morbidity and time to mortality, a bivariate modelling of quantile residual lifetime subject to right censoring might be of utmost interest. In this dissertation, we develop the estimation of conditional quantile residual lifetime on semi-competing risks data. The proposed estimator is conditioning on the occurrence of the nonterminal event beyond time t. The covariate effects on specific pairs of failure times are evaluated based on a log-linear regression on conditional quantile residual lifetime for semi-competing risks data. Numerical studies demonstrate a reasonable performance of the estimator for moderate sample sizes. The proposed method is applied to a study of breast cancer data from a phase III clinical trial.

Public Health Significance: In many survival studies, bivariate correlated failure times can be observed in a pair or in the same individual experiencing multiple failure times. It is of interest to know the additional time to failure of a surviving unit, when another unit is known to have failed at an earlier time. In this dissertation, the proposed estimator of the residual lifetime given the occurrence of a failure demonstrates the importance of lifetime expectancy that patients and their family seek to know before an onset of a new treatment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This dissertation represents a culmination of research work and learning over a period of five years. I was motivated by my advisor Dr. Jong-Hyeon Jeong on his research interest of residual lifetime in survival analysis. I would have never reached the point of finishing my dissertation without his support and guidance. He always shares his insightful ideas and encourages me to go through the problems meticulously.

I sincerely thank my committee members Dr. Abdus S. Wahed, Dr. Rusha Li, and Dr. Yu Cheng for their time and valuable suggestions. I give my gratitude to Dr. Abdus S. Wahed for sharing his experiences and caring my research work. I also greatly appreciate the mentorship from Dr. Ruosha Li. She was very open and willing to discuss problems with me no matter when I came to her. Furthermore, I am grateful to Dr. Yu Cheng for her support to complete my PhD journey. I own my sincere thank to DrPH Evelyn Talbott in the Department of Epidemiology for supporting my graduate study. She served as my GSR supervisor for the past three years. I have gained a lot of collaborative experiences in her research group.

Throughout all these years, it was like a challenging trip, with both ups and downs. Fortunately, I was not alone on this road. I want to give special thanks to all of my friends who have accompanied me to provide countless assistance. Thanks and love my father Tzyy-Arng and my mother Ming-Hwa for their encouragement and support; to my sister Wen-Shih and my brother Yen-Ju. Undoubtedly, my fiance Sean deserves a special word of appreciation for his patience and love. Without them, I could not have reached this level of my life.

August 8, 2014

1.0 INTRODUCTION

In most cases where survival data are studied, traditional approaches focus on modeling the distribution of the time to an event such as Kaplan-Meier estimator [Kaplan and Meier, 1958] and Cox proportional hazard model [Cox, 1972, 1975]. Kaplan-Meier estimator is used to measure the fraction of patients survival after treatment. While Cox proportional hazard model accesses the importance of various covariates in the survival times of individuals through the hazard function. However, the hazard function, interpreted as the "instantaneous rate of failure", does not directly provide numerical measure for long-term lifetime reliability. For instance, patients participated in randomized clinical trials often inquire how much more time they have or whether the new treatment improves their life expectancy. To address these questions, the study of residual lifetime has recently received considerable attention in biomedical research. This desirable information can be provided to medical practitioners in predicting the remaining lifetime.

Quantile residual lifetime function provides a straightforward interpretation relating the potential benefit of a secondary course of treatment for patients seeking for a long-term medical care. In addition, without any strong assumptions, it can be estimated at any specific time point [Jeong et al., 2008]. In this dissertation, such advantages are extended to a bivariate modeling of quantile residual lifetime subject to right censoring.

Multivariate failure times arise with various censoring schemes of interest. There are three main classifications for particular censoring schemes in bivariate failure times. First, parallel failure data describe pairs of units or individuals such as a pair of eyes or twins are followed simultaneously until each unit/individual experiences the event of interest. Second, successive failure times occur when each patient is potentially observed from several related events with a natural chronological order. That suggests the gap time between two successive

events is censored by a dependent variable related to the first duration process if the two duration times are correlated. Third, semi-competing risks data arise when the terminal event censors the nontermial event, but not vice versa. In this dissertation, we study the methodology to quantify conditional residual lifetime for semi-competing risks data and parallel failure data.

Our research interest aims for development of a quantile residual lifetime using bivariate correlated survival data. One can depict conditional quantile residual lifetime through a residual lifetime of an event given the occurrence time of another event. Inference on conditional survival function becomes pertinent to this research. Two types of conditioning on survival function are studied, so the inference of quantile residual lifetime are proposed in two aspects:

1. The main work is constructed on semi-competing risks data. The inference condition focuses on the quantile residual lifetime of the terminal event given that the nonterminal event occurs beyond a specific time point. We treat this proposed method as a semi-parametric inference because a Clayton bivariate model is assumed to assess conditional quantile residual lifetime.
2. A nonparametric inference is adopted to study the quantile residual lifetime of a subject under the condition of another subject who has failed at earlier time t. More specifically, we are interested in the quantile residual lifetime of a subject survived up to time t, when another subject is known to have failed at time t_2, for $t > t_2$. The proposed estimator is obtained relating a smoothing technique. Therefore, the bootstrap estimate of standard error is utilized to measure the accuracy of our proposed estimator.

This dissertation is organized as follows. In Chapter 2 we present background details along with the study motivation and objectives. Chapter 3 is devoted to the proposed method for conditional quantile residual lifetime including the inference procedure, the simulation plan and real data application. Chapter 4 presents the comparison of a nonparametric approach using bivariate failure time data with censoring. The summary of our proposed method and possible future work are discussed in Chapter 5.

2.0 BACKGROUND

In this chapter, we present some important background information for our proposed analysis. First of all, we review techniques for dealing with the estimation of residual life function. Structures of bivariate survival data are then introduced based on the schemes of censoring. The estimation of bivariate survival function specifically gives an idea of how investigators account for the association between two failure times. Moreover, kernel smoothing function is applied to estimate unknown density function. The presence of counting processes provides an efficient method of deriving asymptotic properties for our proposed estimator. Finally, the study motivation and objectives are presented.

2.1 MEAN AND MEDIAN (QUANTILE) RESIDUAL LIFETIME

Historically the mean residual life (MRL) function has been popular to characterize a residual life distribution. Its univariate properties have been studied frequently in actuarial, reliability or survivorship analysis. It is defined as the expected remaining lifetime given survival up to time t. That is,

$$m(t) = E(T - t \mid T > t), \tag{2.1}$$

where $t > 0$. A survival function can be written in terms of the MRL function through an inversion formula [Hall and Wellner, 1981],

$$S(t) = \frac{m(0)}{m(t)} \exp\left\{ - \int_0^t m(u)^{-1} du \right\}, \tag{2.2}$$

where $S(t) = Pr(T > t)$. Nonparametric estimation of the MRL function has been proposed by Yang [1977], Lahiri and Ho Park [1992], Chaubey and Sen [1999], Abdous and Berred

[2005] among others. McLain and Ghosh [2011] identified theoretical limitations of semi-parametric conditional MRL models and compared them with their proposed nonparametric methods in presence of censoring.

However, as Schmittlein and Morrison [1981] first pointed out, the mean residual lifetime has many theoretical and practical shortcomings such as (i) inappropriateness for frequently encountered censored data, (ii) skewness in time-to-event data and (iii) non-existence of the mean residual life function for some distributions. As an alternative, the median residual life function or more generally α-quantile residual life function would be more recommended than the mean. For $0 < \alpha < 1$, the α-quantile residual life function is defined as

$$\begin{aligned} \gamma_\alpha(t) &= \alpha_{\text{quantile}}(T - t \mid T > t) \\ &= \inf\{x : S(t+x) < (1-\alpha)S(t)\} \\ &= S^{-1}\{(1-\alpha)S(t)\} - t \end{aligned} \tag{2.3}$$

at $t > 0$, and interpreted as the α percentile additional time to failure, given no failure by time t. Numerous studies for estimating the quantile residual life function have been conducted in the univariate settings with or without covariates. For example, Csörgő and Csörgő [1987] initiated a nonparametric large sample estimation theory for the percentile residual lifetime and also constructed confidence bands based on non-censored data. Chung [1989] proved that the scaled $(1-p)$ percentile residual lifetime process can be almost surely approximated by a Gaussian process and then constructed confidence bands using bootstraps. Lillo [2005] studied several aspects of the median residual life function and proposed that the median residual life function can determine the distribution uniquely on an interval which fulfills the insufficiency relative to the mean residual life function. They also found that the same pattern of relationship holds for both the mean and median residual life functions. A kernel-type smooth estimator of the quantile residual life function was studied by various authors [Padgett, 1986, Padget and Thombs, 1988, Alam and Kulasekera, 1993]. Later Gelfand and Kottas [2003] proposed a Bayesian approach to fit a median residual life regression that was induced by a semiparametric accelerated failure time (AFT) regression model. Jung et al. [2009] developed a time-specific regression method to model the effect of covariates on quantile residual lifetime without specifying semiparametric model for the underlying failure

time. It is noted that these methods are directly applicable to modeling the residual life function in the univariate settings. Some of the important concepts can be addressed in studying multivariate failure time data.

2.2 STRUCTURES OF BIVARIATE SURVIVAL DATA

Over the past few decades, much attention has been paid to bivariate times to the events and inference. Studies on paired subjects, time to recurrence and time to death of a disease, or the familial dependence for lifetimes of fathers and sons attempt to observe the life length of each subject with the presence of censoring. If the development of those event times is considered as longitudinal, such data can be formulated as multi-state models. A multi-state model is defined as a model for a stochastic process, which at any time points occupied one of a set of discrete state such as healthy, diseased, disease with complications, and dead [Hougaard and Hougaard, 2000]. The state structure varies with a statistical model. Generally, any continuous multivariate distribution can be described by a multi-state model. The limit of censoring pattern, however, requires unique censoring for parallel failure data because of its longitudinal approach.

Unlike the competing risks setting that only allows the observation time ends upon the occurrence of the first failure, semi-competing risks data [Fine et al., 2001] refers to the situation where a subject may experience a nonterminal event, such as recurrence, and/or a terminal event, such as death. The terminal event can censor the nonterminal event but not vice versa. For instance, recurrence is observable before mortality but only mortality is observable otherwise. As illustrated in Figure 2.1, patients may die without recurrence after the initial treatment or experience recurrence but survive beyond time t. Therefore, a pair of event times (T_1, T_2) along with a censoring indicator is recorded for each patient. A censoring indicator is independent of both T_1 and T_2. Since there exists an association between time-to-recurrence and time-to-death from the same patient, the joint survival function of those two event times is assumed to follow a copula model defined as

$$S_{1,2}(t_1, t_2) = Pr(T_1 > t_1, T_2 > t_2) = C_\theta\big(S_1(t_1), S_2(t_2)\big),\ 0 < t_1 < t_2, \tag{2.4}$$

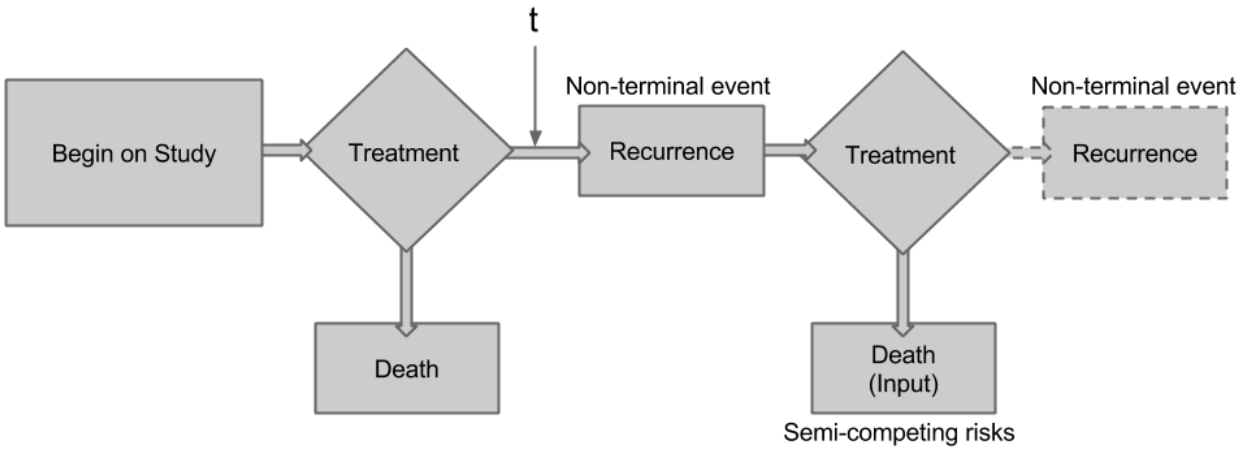


Figure 2.1: Illustration of semi-competing risks data

where T_1, T_2 are times to nonterminal and terminal events, S_1, S_2 are their respective marginal survival functions. Since $S(t_1, t_2)$ is only identifiable on the upper wedge where $T_1 < T_2$, the interpretation of $S_1(t_1)$ as a marginal distribution is controversial and would leave it unspecified [Day et al., 1997, Fine et al., 2001].

For parallel failure data, there are two standard censoring cases, homogeneous and heterogeneous censoring. Homogeneous censoring arises when a simultaneous censoring is observed for both subjects. For example, in a study that the time to a deterioration level is of interest in pairs of eyes, the censoring time for both eyes are observed simultaneously when an individual is withdrawn from the study. In contrast, heterogeneous censoring corresponds to observing a separate censoring time for each subject. Taking a twin pair for example, an individual can be lost to follow-up during the study, even though the other is still followed. The model depicted in Figure 2.2 shows the flow of parallel data and gives some insight of modeling conditional quantile residual lifetime when one of the subjects has failed at time t before another subject.

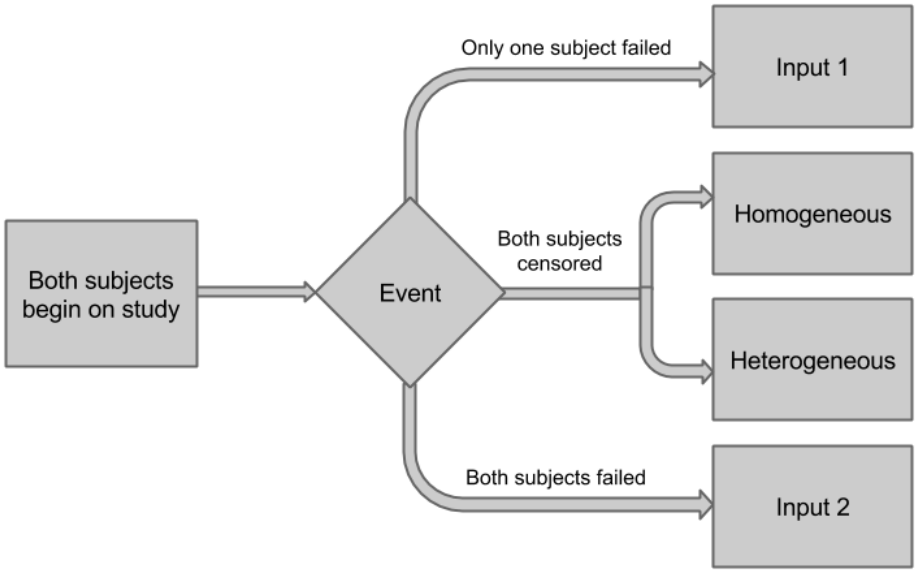


Figure 2.2: Illustration of parallel data

Several possible mechanisms to generate dependence between failure times have been seen in the analysis of bivariate survival data. When common risks appear between the various courses for parallel failure data, conditional independence is assumed that these common risks are responsible for all the dependence seen between failure times [Hougaard and Hougaard, 2000]. The most common model is the frailty model based on a common factor in the hazard. Consider the form of the bivariate survival function

$$S_{1,2}(t_1, t_2) = Pr(T_1 > t_1, T_2 > t_2) \tag{2.5}$$

and its marginals are $S_{1,2}(t_1, 0) = S_1(t_1) = Pr(T_1 > t_1)$ and $S_{1,2}(0, t_2) = S_2(t_2) = Pr(T_2 > t_2)$, respectively. With consideration of dependence between two failure times, Prentice and Cai [1992] developed a covariance function that captures the characterization of the nature of the dependence between two or more correlated failure times and proposed a nonparametric bivariate survival function estimator in terms of marginal survival functions and a conditional covariance function. Dabrowska [1988] developed the product integral representation of

univariate survival functions and generalized to the bivariate case. Both methods have good sample performance but are lack of nonparametrical efficiency. Pruitt [1991] proposed a modification of the self-consistency equation to estimate a bivariate survival function by redistributing singly-censored observations over their associated region and by assigning uncensored observations to give mass 1/n to the observed survival time. His estimator is a distribution function and each uncensored and singly-censored observation is redistributed and smoothed over the region in one direction depending on the other observations. However, instead of smoothing only in one direction, van der Laan [1994] adopted Pruitt′s estimator and generalized it to the two dimension case using the edge corrected bivariate kernel density estimator to develop a nonparametric maximum likelihood approach (NPMLE) in order to deal with nonparametric efficiency.

2.3 KERNEL SMOOTHING FUNCTION

Kernel smoothing is an effective tool for visualising the distribution of data. The performance of this methodology depends on the choice of a smoothing parameter, i.e. bandwidth, which takes an important role to approximate the unknown density [Wand and Jones, 1993, Simonoff, 1996]. There exists numerous methods that have good theoretical properties in selecting the scalar bandwidth in univariate kernel density estimation and most of them can be extended to the multivariate case [Jones et al., 1996]. When bivariate data are considered, let $\mathbf{T} = (T_1, T_2)$ be i.i.d. bivariate random vector drawn from a density f, the kernel density estimator is defined by

$$\widehat{f}(\mathbf{t}; \mathbf{H}) = \frac{1}{n} \sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{t} - \mathbf{T}_i), \tag{2.6}$$

where $\mathbf{t} = (t_1, t_2)^{\mathrm{T}}$ and $K_{\mathbf{H}}(\mathbf{t}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2}\mathbf{t})$. Here the kernel K is a symmetric probability density function; $\mathbf{H}$ is a symmetric and positive-definite bandwidth matrix to determine the performance of $\widehat{f}$. The choice of $\mathbf{H}$ is crucial because of its effect on the shape of the corresponding estimator. The study of data-driven methods for selecting $\mathbf{H}$ provides a general application on multivariate kernel smoothing problems.

It is known that the most common bandwidth matrix seen in multivariate kernel density estimation is a diagonal matrix [Wand and Jones, 1994, Sain et al., 1994]. The selectors for a full (unconstrained) bandwidth matrix are more challenging and rarely discussed in the literature. A full bandwidth is

$$\mathbf{H} = \begin{bmatrix} h_1^2 & h_1 h_2 \\ h_1 h_2 & h_2^2 \end{bmatrix}$$

which provides kernels with an arbitrary orientation whereas a diagonal matrix is only oriented to the co-ordinate axes. Hall et al. [1992] studied a modification of cross-validation method that involved a presmoothing of the pairwise differences of the observations. It was named by smoothed cross-validation and revealed to have an excellent asymptotic performance. Duong and Hazelton [2005] considered cross-validation technique for full bandwidth matrices including unbiased, biased, and smoothed cross-validation approaches and compared their performance. The judge of the performance was according to a global error criteria for $\widehat{f}(\mathbf{t}, \mathbf{H})$ such as mean integrated squared error (MISE) given by

$$\begin{aligned} \mathrm{MISE}(\mathbf{H}) \equiv \mathrm{MISE}\widehat{f}(\cdot; \mathbf{H}) =& \mathrm{E} \int_{R^d} \left(\widehat{f}(\mathbf{t}, \mathbf{H}) - f(\mathbf{t})\right)^2 \mathrm{d}\mathbf{t} \\ =& \int_{R^d} \mathrm{Bias}\{[\widehat{f}(\mathbf{t}, \mathbf{H})]\}^2 \mathrm{d}\mathbf{t} + \int_{R^d} \mathrm{Var}[\widehat{f}(\mathbf{t}, \mathbf{H})] \mathrm{d}\mathbf{t}. \end{aligned} \tag{2.7}$$

The results suggest that smoothed cross-validation for full bandwidth matrices is the most reliable among the selectors that they studied. The smoothed cross-validation function is

$$SCV(\mathbf{H}) = n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} (K_{\mathbf{H}}^2 L_{\mathbf{G}}^2 - 2K_{\mathbf{H}} L_{\mathbf{G}}^2 + L_{\mathbf{G}}^2)(\mathbf{X}_i - \mathbf{X}_j) + n^{-1} R(K) |\mathbf{H}|^{-1/2}, \tag{2.8}$$

where $L_{\mathbf{G}}(\cdot)$ is the pilot kernel with pilot bandwidth matrix $\mathbf{G}$ and $R(K) = \int_{R^d} K(\mathbf{x})^2 d\mathbf{x} < \infty$. Therefore, a smoothed cross-validation bandwidth matrix for estimating bivariate density function is used throughout this work.

2.4 THE COUNTING PROCESS AND MARTINGALE

A counting process is a stochastic process $\{M(t), t > 0\}$ with characteristics of being positive, integers, and an increasing step function. More generally, it is to model the numbers of events in different types that occur over time. A martingale is based on what has happened up to time t, the expected future change is 0. The introduction of martingales in survival analysis was firstly introduced by Aalen [1976] in his Ph.D thesis. Let $M(t), t \geq 0$ be a right-continuous stochastic process with left-hand limits. Filtration at time t, $\mathcal{F}_t$, is a σ-field that increases or enlarges as a function of t. It usually corresponds to history or information collected up to time t. It is a martingale if for any $t \geq s$

$$E[M(t) \mid \mathcal{F}_s] = M(s). \tag{2.9}$$

In order to understand the concepts leading to the discussion of counting process and martingale with censored survival data, the estimation procedure in survival probability is briefly reviewed. Assume T and C are continuous nonnegative and independent random variables. T has a distribution function $F(t)$ and density function $f(t) = dF(t)/dt$. When the Kaplan–Meier estimator is studied, the concept of hazard function is required. The survival function is given by $S(t) = Pr(T > t)$. The hazard function is defined by means of a conditional probability, that is,

$$\begin{aligned} \lambda(t) &= \lim_{\Delta t \to 0} \frac{Pr\{t \leq T < t + \Delta t \mid T \geq t\}}{\Delta t} \\ &= f(t)/S(t) \\ &= -d \log[S(t)]/dt \end{aligned} \tag{2.10}$$

which can be described as the risk at which an event happens, conditional on not having happened previously. The cumulative hazard function $\Lambda(t) = \int_0^t \lambda(u)du$ is defined and the survival probability can be written as $S(t) = \exp\{-\Lambda(t)\}$ for continuous T [Klein and Moeschberger, 2003].

Martingale theory is widely used for right censored data because of its advantages in variance simplification. Thus, we introduce the concepts of how martingale transformation leads to present the properties of two famous estimators in survival analysis. Let the death

process given at time t be defined by $N_i(t) = I(\widetilde{T}_i \leq t, \delta_i = 1)$ where $\widetilde{T}_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$ be the indicator variable. Let $Y_i(t) = I(\widetilde{T}_i \geq t)$ denote as the risk process and let $\overline{N}(t) = \sum_{i=1}^{n} N_i(t)$ and $\overline{Y}(t) = \sum_{i=1}^{n} Y_i(t)$. By theorem 1.3.1 in Fleming and Harrington [2011], it follows that

$$M_i(t) = N_i(t) - \int_0^t Y_i(u) d\Lambda(u) \tag{2.11}$$

is a martingale for each i subject and the counting process $\{N_i : i = 1, \cdots, n\}$ holds for subjects who have failed or have been censored up to and including that time t. The Nelson–Aalen estimator [Nelson, 1972, Aalen, 1978] for cumulative hazard can be expressed as

$$\widehat{\Lambda}(t) = \int_0^t \frac{d\overline{N}(u)}{\overline{Y}(u)}, \tag{2.12}$$

and it follows

$$\widehat{\Lambda}(t) - \Lambda(t) = \int_0^t \frac{\overline{N}(u) - \overline{Y}(u) d\Lambda(u)}{\overline{Y}(u)} = \int_0^t \frac{dM(u)}{\overline{Y}(u)}, \tag{2.13}$$

where $M(u) = \sum_{i=1}^{n} M(u)$ is the sum of martingales for all subjects. The property of zero mean in martingale transformation shows the consistency of the Nelson-Aalen estimator in equation (2.12). Then we can adopt the property of martingale Central Limit Theorem to establish distributional function of cumulative hazard function. By theorem 3.2.1 in Fleming and Harrington [2011], it indicates that $\widehat{\Lambda}(t)$ is an asymptotically unbiased estimator of $\Lambda(t)$, with bias converging to zero at an exponential rate as $n \to \infty$. The variance of $\Lambda(t)$ should approach

$$Var[\widehat{\Lambda}(t)] = \int_0^t \frac{1}{\overline{Y}(t)} \left\{ 1 - \frac{\triangle \overline{N}(u)}{\overline{Y}(u)} \right\} \frac{d\overline{N}(u)}{\overline{Y}(u)}, \tag{2.14}$$

for large n. Since there exists a relationship between Λ and S, the Kaplan–Meier estimator can be written in terms of martingale transformation as well. The variance of $\widehat{S}$ can be obtained as the same calculation for the Nelson-Aalen estimator which is approximately the Greenwood's formula:

$$Var(\widehat{S}(t)) = \widehat{S}^2(t) \int_0^t \frac{d\overline{N}(u)}{\overline{Y}(u)\{\overline{Y}(u) - \triangle \overline{N}(u)}. \tag{2.15}$$

2.5 MOTIVATION AND OBJECTIVES

A life expectancy of the treatment outcome is the most important information that patients and their family are concerned. Although traditional tools such as Kaplan-Meier estimator and Cox proportional hazard models are usually referred by physicians in estimating survival probabilities, the estimates of residual life expectancy cannot be obtained directly for patients who have been followed up for several years after the initial treatment. This desirable estimate becomes critical for patients seeking long-term care on their secondary course of treatment. We consider this idea in the extension of modeling bivariate failure time data with censoring.

The objective of this study is to determine the additional time to failure on the later event given the occurrence of the first event when the event times are correlated. We may encounter a situation that the first event remains event free by a certain time point, saying, under the condition of breast cancer patients staying recurrence free up to 5 years from the initial diagnosis, how much residual life expectancy they will get after survival up to 10 years. We propose the method of conditional quantile residual lifetime by using the concept of semi-competing risks data.

In addition to conditioning on a failure happening beyond a time, another approach is studied when the first event occurs just at time t_1, how much remaining lifetime can be prolonged after the later event stays event-free up to time t_2. One of the challenges that investigators might encounter is to find the underlying density function of the unknown failure time distribution with censoring. The proposed estimator of conditional quantile residual lifetime can be obtained by inverting the estimating equation. However, the issue of the unknown density function still needs to be addressed in the estimation of conditional survival probability. A bivariate kernel density smoothing with an appropriate bandwidth matrix is then used to overcome this challenge.

In the application of bivariate failure time data, we use examples of semi-competing risks data and parallel data to evaluate our proposed estimators. First, the breast cancer data in B-14 clinical trial provide time to recurrence and time to death with the censoring scheme from semi-competing risks. We examine the additional time to the terminal event such as

death given a survivorship from the nontermial event such as recurrence beyond a time point. Second, a study of duration of ventilating tubes in ears provides a good demonstration. The insertion of ventilating tubes is often the treatment of choice for otitis media with effusion in childhood. Since the observed failure times in both ears are recorded from the same child, it takes account of the correlation between failure times. Each failure time is subject to right censoring due to the cession of tube functioning or tube extrusion.

3.0 SEMIPARAMETRIC INFERENCE USING SEMI-COMPETING RISKS DATA

3.1 INTRODUCTION

In many biomedical studies, multivariate failure time data have been commonly encountered by scientific investigators. To consider a series of random variables, the approach of studying residual lifetime analysis was directly toward the multivariate mean residual life (MMRL) function in literature. For instance, Arnold and Zahedi [1988] studied some general characterization properties of MMRL function and also discussed the relationship between the MMRL function and the hazard gradient. Nair and Nair [1989] proved an extended theoretical results for the bivariate case. Shaked and Shanthikumar [1991] introduced a dynamic notion of mean residual life functions in the context of multivariate reliability theory. They studied the properties of mean residual life functions and their relationship to the multivariate conditional hazard rate functions. A natural nonparametric estimator of bivariate mean residual life based on empirical survival function has been studied by Jeong et al. [1996] and Kulkarni and Rattihalli [2002]. They showed an asymptotically unbiased estimator which has the joint weak convergence to a zero-mean Gaussian process. Although certain efficient methods of MMRL function have been proposed, none has been applied to quantile residual lifetime using bivariate failure data under random censoring.

Unlike the competing risks setting where only time-to-first-event is observed, semi-competing risks data concern a situation where a subject may experience the nonterminal event, such as recurrence, and/or the terminal event, such as death. The terminal event can censor the nonterminal event but not vice versa. The semi-competing risks data have been also described as an illness-death model [Keiding, 1991, Xu et al., 2010]. Fine et al. [2001]

proposed a plug-in estimator for the marginal distribution of the nonterminal event using the association parameter from a concordance estimating function. Peng and Fine [2007] modeled the covariate effects on the survival function of the intermediate events via a functional regression model. Li and Peng [2011] applied a quantile regression method to appropriately handle the complexity posed by left-truncated semi-competing risks data. Earlier Ghosh [2006] developed methods to infer dependence of semi-competing risks data across strata of a discrete covariate Z. To the best of our knowledge, however, little attention has been paid to inference on residual lifetime for semi-competing risks.

In this chapter, a method for estimating the conditional α-quantile residual lifetime is proposed for semi-competing risks data. The objective is to infer the conditional residual life distribution of time-to-the terminal event given that a patient has not experienced the nonterminal event by a certain time point, when the event times are correlated. For instance, in breast cancer patients, most of them might have their recurrent breast cancer in the first 3 to 5 years after initial treatment. Before starting the secondary treatment for breast cancer recurrence, it would be informative to know the median of a time-to-death distribution beyond year 10 if they were recurrence-free up to 5 years from the initial diagnosis.

This chapter is organized as follows. Section 3.2 introduces notation and model definition. Section 3.3 shows one-sample inference procedure for the conditional quantile residual life function. In Section 3.4, a time-specific conditional quantile residual life regression is proposed, together with an inference procedure for the regression coefficients. Section 3.5 presents simulation studies to assess performances of the proposed methods. In Section 3.6, we demonstrate the proposed method through an application to NSABP B-14 phrase III breast cancer dataset. Finally, Section 3.7 concludes with a summary and discussion.

3.2 MODEL DEFINITION

For the ith subject, let T_{1i} be the nonterminal event time and T_{2i} be the terminal event time. Assume two event times are correlated and there exists the censoring time C_i independent of both T_{1i} and T_{2i}, such as time to lost of follow-up. Define $Z_i = \min(T_{1i}, T_{2i})$, $\delta_{Z_i} =$

$I(Z_i < C_i)$, $T'_{2i} = \min(T_{2i}, C_i)$, $\delta_{2i} = I(T_{2i} < C_i)$, and $T'_{1i} = \min(T_{1i}, T'_{2i})$, $\delta_{1i} = I(T_{1i} < T'_{2i})$ where $I(\cdot)$ is the indicator function. Thus, the semi-competing risks data are denoted by $\{T'_{1i}, \delta_{1i}, T'_{2i}, \delta_{2i}, i = 1, \cdots, n\}$. Since $0 \leq T'_{1i} \leq T'_{2i}$, this implies that the joint distribution of (T_1, T_2) is only identifiable when observations are restricted to the upper wedge. Let $Y_{Zi}(t) = I(Z_i \geq t)$ and $N_{Zi}(t) = \delta_{Zi} I(Z_i \leq t)$ be the at-risk and death processes for Z. Similarly, let $Y_{2i}(t) = I(T'_{2i} \geq t)$ and $N_{2i}(t) = \delta_{2i} I(T'_{2i} \leq t)$ for T_2. In a univariate setting, the α-quantile residual life function at time t is defined as

$$\gamma_\alpha(t) = \alpha\text{-quantile}(T - t \mid T > t), \ 0 < \alpha < 1, \tag{3.1}$$

which describes the α-quantile residual lifetime among survivors beyond time t. The function (3.1) can be written as $Pr(T > t + \gamma_\alpha) = (1-\alpha) Pr(T > t)$. Note that the conditional survival function of the terminal event time given that the nonterminal event did not occur by t_1 is given by

$$S_{2|1}(t_2 \mid t_1) = Pr(T_2 > t_2 \mid T_1 > t_1). \tag{3.2}$$

Then the conditional quantile residual life function at time t_2 is defined as

$$\gamma^S_{\alpha|t_1}(t_2) = \inf \left\{ x : S_{2|1}(t_2 + x \mid t_1) < (1 - \alpha) S_{2|1}(t_2 \mid t_1) \right\}. \tag{3.3}$$

which is the α-quantile of the residual life distribution of the terminal event evaluated at time t_2 among patients who are recurrence-free up to time t_1.

With semi-competing risks data, the Kaplan-Meier procedure can not be employed to obtain a consistent estimator of $S_1(t) = Pr(T_1 > t)$, the marginal distribution of the non-terminal event. In general, without consideration of the dependent structure, $\widehat{S}_1$ does not converge to S_1 as $n \to \infty$. Fine et al. [2001] proposed a novel plug-in estimator for S_1 using a closed-form estimator for an association parameter θ along with the Kaplan-Meier estimators for Z and T_2.

3.3 INFERENCE

3.3.1 Estimation of θ and S_1

When two event times are assumed to be correlated, the dependence structure between T_1 and T_2 is often formulated via the Clayton copula [Clayton, 1978], that is, for $0 < \theta < \infty$ and $0 \leq t_1 \leq t_2 < \infty$, the joint survival function $S(t_1, t_2)$ is expressed as

$$S_{1,2}(t_1, t_2) = \{S_1(t_1)^{-1/\theta} + S_2(t_2)^{-1/\theta} - 1\}^{-\theta}, \tag{3.4}$$

where $\theta \to 0$ corresponds to the maximal positive dependence and $\theta \to \infty$ to independence. For all $t_1 \leq t_2$, the parameter θ is equivalent to the ratio of the two conditional probabilities (predictive hazard ratio), that is, the conditional probability of $T_2 > t_2$ given $T_1 = t_1$ over the conditional probability of $T_2 > t_2$ given $T_1 > t_1$ being constant [Oakes, 1989]. If the model (3.4) describes the dependency between T_1 and T_2 in the whole plane, the parameter θ has the usual relationship with Kendall's τ, defined by

$$\tau = Pr\{(T_{1i} - T_{1j})(T_{2i} - T_{2j}) > 0\} - Pr\{(T_{1i} - T_{1j})(T_{2i} - T_{2j}) < 0\},\ i \neq j \tag{3.5}$$

where (T_{1i}, T_{2i}) and (T_{1j}, T_{2j}) are independent pairs of (T_1, T_2) and $\tau = 1/(1 + 2\theta)$.

Our proposed estimator will be built upon estimation of the dependent structure and the marginal distribution of the nonterminal event, which will play an important role for inference on the quantile residual lifetime in semi-competing risks data. We briefly review the procedure for estimating θ and S_1 from Fine et al. [2001] and Jiang et al. [2003]. First, let $\widetilde{T}_{1ij} = \min(T_{1i}, T_{1j})$, $\widetilde{T}_{2ij} = \min(T_{2i}, T_{2j})$ and $\widetilde{C}_{ij} = \min(C_i, C_j)$ for $i \neq j$. The concordance indicator $\Delta_{ij} = I\{(T_{1i} - T_{1j})(T_{2i} - T_{2j}) > 0\}$ has the expected value of $(1 + \theta)/(1 + 2\theta)$ under model (3.4). A consistent estimator for θ is obtained from a concordance estimating function $U(\theta)$ which is

$$U(\theta) = \sum_{i<j} W(\widetilde{T}'_{1ij}, \widetilde{T}'_{2ij}) D_{ij} \left\{ \Delta_{ij} - \frac{1 + \theta}{1 + 2\theta} \right\}, \tag{3.6}$$

where $D_{ij} = I(\widetilde{T}_{1ij} < \widetilde{T}_{2ij} < \widetilde{C}_{ij})$, $\widetilde{T}'_{1ij} = \min(T'_{1i}, T'_{1j})$, $\widetilde{T}'_{2ij} = \min(T'_{2i}, T'_{2j})$, and $W(u, v)$ is a random weight function. The equation $U(\theta) = 0$ gives the estimation of θ as

$$\widehat{\theta} = \frac{\sum_{i<j} W(\widetilde{T}'_{1ij}, \widetilde{T}'_{2ij}) D_{ij}(1 - \Delta_{ij})}{\sum_{i<j} W(\widetilde{T}'_{1ij}, \widetilde{T}'_{2ij}) D_{ij}(2\Delta_{ij} - 1)}. \tag{3.7}$$

A useful form of the weight function is

$$W_{a,b}(x, y) = n^{-1} I\{T'_{1i} \geq \min(a, x), T'_{2i} \geq \min(b, y)\}, \tag{3.8}$$

where a and b are chosen to be the pth quantile of the uncensored T_1 and T_2. According to the simulation results from Fine et al. [2001], $\widehat{\theta}$ is efficient when p takes a value between 75 and 95. The survival function of Z denoted as $S_Z(t)$ is equivalent to $Pr(T_1 > t, T_2 > t)$. Therefore, one can manipulate (3.4) to get $S_1(t)$ in terms of $S_Z(t)$, $S_2(t)$ and θ, that is,

$$S_1(t) = \{S_Z(t)^{-1/\theta} - S_2(t)^{-1/\theta} + 1\}^{-\theta}. \tag{3.9}$$

Since Z and T_2 are subject to right censoring by C, $S_Z(t)$ and $S_2(t)$ can be consistently estimated with Kaplan-Meier estimators $\widehat{S}_Z(t)$ and $\widehat{S}_2(t)$, respectively. A closed form estimator for $S_1(t)$ is obtained by replacing $S_Z(t)$, $S_2(t)$ and θ by their estimators $\widehat{S}_Z(t)$, $\widehat{S}_2(t)$ and $\widehat{\theta}$ in equation (3.9).

3.3.2 Estimation of $\gamma^S_{\alpha|t_1}$

One approach to estimate the quantile residual lifetime from (3.3) is to solve the equation $u(\gamma^S_{\alpha|t_1}) = 0$ for $\gamma^S_{\alpha|t_1}$, where

$$u(\gamma^S_{\alpha|t_1}) = S_{2|1}(t_2 + \gamma^S_{\alpha|t_1} \mid t_1) - (1 - \alpha) S_{2|1}(t_2 \mid t_1). \tag{3.10}$$

Based on equations (3.4) and (3.9), the conditional survival probabilities $S_{2|1}(t_2 + \gamma^S_{\alpha|t_1} \mid t_1)$ and $S_{2|1}(t_2 \mid t_1)$ can be written in terms of $S_Z(t_1)$, $S_2(t_1)$, $S_2(t_2)$ and $S_2(t_2 + \gamma^S_{\alpha|t_1})$. To account for the correlation between the marginal survival distributions of T_2 evaluated at t_1, t_2 and $t_2 + \gamma^S_{\alpha|t_1}$, suppose that $k = S_2(t_1)/S_2(t_2 + \gamma^S_{\alpha|t_1})$ and $m = S_2(t_1)/S_2(t_2)$ are the ratios

of S_2 at different time points. Therefore, once can rewrite the equation (3.10) as a function of five elements including $S_Z(t_1)$, $S_2(t_1)$, θ, k, and m shown as below:

$$\begin{aligned}
u(\gamma^S_{\alpha|t_1}) =& S_{2|1}(t_2+\gamma^S_{\alpha|t_1} \mid t_1) - (1-\alpha)S_{2|1}(t_2 \mid t_1) \\
=& \frac{[S_1(t_1)^{-1/\theta} + S_2(t_2+\gamma^S_{\alpha|t_1})^{-1/\theta} - 1]^{-\theta} - (1-\alpha)[S_1(t_1)^{-1/\theta} + S_2(t_2)^{-1/\theta} - 1]^{-\theta}}{S_1(t_1)} \\
=& \frac{[S_Z(t_1)^{-1/\theta} - S_2(t_1)^{-1/\theta} + S_2(t_2+\gamma^S_{\alpha|t_1})^{-1/\theta}]^{-\theta}}{[S_Z(t_1)^{-1/\theta} - S_2(t_1)^{-1/\theta} + 1]^{-\theta}} \\
& - (1-\alpha)\frac{[S_Z(t_1)^{-1/\theta} - S_2(t_1)^{-1/\theta} + S_2(t_2)^{-1/\theta}]^{-\theta}}{[S_Z(t_1)^{-1/\theta} - S_2(t_1)^{-1/\theta} + 1]^{-\theta}} \\
=& \frac{[S_Z(t_1)^{-1/\theta} - (1-k^{1/\theta})S_2(t_1)^{-1/\theta}]^{-\theta} - (1-\alpha)[S_Z(t_1)^{-1/\theta} - (1-m^{1/\theta})S_2(t_1)^{-1/\theta}]^{-\theta}}{[S_Z(t_1)^{-1/\theta} - S_2(t_1)^{-1/\theta} + 1]^{-\theta}} \\
=& \phi\left\{S_Z(t_1), S_2(t_1), \theta, k, m\right\}
\end{aligned} \tag{3.11}$$

After doing some algebra, the estimating equation $u(\gamma^S_{\alpha|t_1})$ simplifies to $\phi\left\{S_Z(t_1), S_2(t_1), \theta, k, m\right\}$, where

$$\phi(a,b,c,d,e) = \frac{\left[a^{-1/c} - \left(1-d^{1/c}\right)b^{-1/c}\right]^{-c} - (1-\alpha)\left[a^{-1/c} - \left(1-e^{1/c}\right)b^{-1/c}\right]^{-c}}{\left[a^{-1/c} - b^{-1/c} + 1\right]^{-c}}$$

and hence the consistent estimator is given by $\widehat{u}(\gamma^S_{\alpha|t_1}) = \phi\left\{\widehat{S}_Z(t_1), \widehat{S}_2(t_1), \widehat{\theta}, \widehat{k}, \widehat{m}\right\}$, where $\widehat{k}$ and $\widehat{m}$ are estimators of the ratios of $\widehat{S}_2$ at different times. Let $\widehat{\gamma}^S_{\alpha|t_1}$ denote the solution. According to Theorem 3.4.2 of Fleming and Harrington [2011], when a continuous failure time random variable T along with a censoring time C are subject to the continuity of S, the uniform consistency of $\widehat{S}$ holds over $0 \leq t \leq \nu$, where $\nu = \sup\{t : Pr(\min(T,C) > t) > 0\}$. Since Z and T_2 are subject to right censoring by C, $\widehat{S}_Z(t_1)$ and $\widehat{S}_2(t_1)$ are strongly consistent for $S_Z(t_1)$ and $S_2(t_1)$ over $0 \leq t_1 \leq \nu$. Therefore, for $t_2 + \gamma^S_{\alpha|t_1} \leq \nu$, $\widehat{u}(\gamma^S_{\alpha|t_1})$ uniformly converges to $u(\gamma^S_{\alpha|t_1})$. Suppose $\gamma^S_{\alpha|t_1,0}$ denotes as the true value of α-quantile residual lifetime such that $u(\gamma^S_{\alpha|t_1,0}) = 0$. One can show that $\widehat{\gamma}^S_{\alpha|t_1}$ is a consistent estimator of $\gamma^S_{\alpha|t_1,0}$ consequently.

It is noted that both the Kaplan-Meier estimator and the Nelson-Aalen estimator can be obtained using the theory of counting process. The derivation of their asymptotic properties

is provided in Chapter 3 of Fleming and Harrington [2011]. The martingale representations for $\widehat{S}_Z$ and $\widehat{S}_2$ indicate the weak convergence of

$$n^{1/2}\left\{\widehat{S}_Z(t) - S_Z(t)\right\} = -S_Z(t)n^{-1/2}\sum_{i=1}^{n}\int_0^t \frac{dM_{Z_i}(s)}{h_Z(s)} + o_p(1)$$

$$n^{1/2}\left\{\widehat{S}_2(t) - S_2(t)\right\} = -S_2(t_1)n^{-1/2}\sum_{i=1}^{n}\int_0^t \frac{dM_{2_i}(s)}{h_2(s)} + o_p(1),$$

where $h_Z(t)$ and $h_2(t)$ are the limits of $\widehat{h}_Z(t) = n^{-1}\sum_{i=1}^n Y_{Zi}(t)$ and $\widehat{h}_2(t) = n^{-1}\sum_{i=1}^n Y_{2i}(t)$, and

$$\begin{aligned} M_{Zi}(t) &= N_{Zi}(t) - \int_0^t Y_{Zi}(s)d\Lambda_Z(s), \\ M_{2i}(t) &= N_{2i}(t) - \int_0^t Y_{2i}(s)d\Lambda_2(s) \end{aligned} \tag{3.12}$$

are martingales, and $\Lambda_Z(s)$ and $\Lambda_2(s)$ are the cumulative hazard functions for Z and T_2', respectively. According to Theorem 3.2.3 in Fleming and Harrington [2011], if $S(t) > 0$,

$$\frac{\widehat{S}(t)}{S(t)} = 1 - \int_0^t \frac{\widehat{S}(s^-)}{S(s)}\left\{\frac{d\overline{N}(s)}{\overline{Y}(s)} - d\Lambda(s)\right\}. \tag{3.13}$$

Since k and m are ratios of the marginal survival probability of T_2 evaluated at different time points, one can approximate $\widehat{k}/k$ using equation (3.13) represented as

$$\begin{aligned} \frac{\widehat{k}}{k} &= \frac{\widehat{S}_2(t_1)S_2(t_2+\gamma^S_{\alpha|t_1})}{\widehat{S}_2(t_2+\gamma^S_{\alpha|t_1})S_2(t_1)} \\ &= \frac{1 - \int_0^{t_1}\frac{\widehat{S}_2(s^-)}{S_2(s)}\left\{\frac{d\overline{N}_2(s)}{\overline{Y}_2(s)} - d\Lambda_2(s)\right\}}{1 - \int_0^{t_2+\gamma^S_{\alpha|t_1}}\frac{\widehat{S}_2(s^-)}{S_2(s)}\left\{\frac{d\overline{N}_2(s)}{\overline{Y}_2(s)} - d\Lambda_2(s)\right\}} \\ &= 1 + \frac{\int_{t_1}^{t_2+\gamma^S_{\alpha|t_1}}\frac{\widehat{S}_2(s^-)}{S_2(s)}\left\{\frac{d\overline{N}_2(s)}{\overline{Y}_2(s)} - d\Lambda_2(s)\right\}}{\left[1 - \int_{t_1}^{t_2+\gamma^S_{\alpha|t_1}}\frac{\widehat{S}_2(s^-)}{S_2(s)}\left\{\frac{d\overline{N}_2(s)}{\overline{Y}_2(s)} - d\Lambda_2(s)\right\}\right]}. \end{aligned} \tag{3.14}$$

Therefore, in terms of martingale representation, weak convergence of $n^{1/2}\big(\widehat{k}-k\big)$ gives

$$n^{1/2}\big(\widehat{k}-k\big)=\sum_{i=1}^{n}\frac{k\cdot n^{1/2}\int_{t_1}^{t_2+\gamma^S_{\alpha|t_1}}\frac{dM_{2i}(s)}{h_2(s)}}{\left[1-\sum_{l=1}^{n}\int_{0}^{t_2+\gamma^S_{\alpha|t_1}}\frac{dM_{2l}(s)}{h_2(s)}\right]}+o_p(1). \tag{3.15}$$

Similarly,

$$n^{1/2}\big(\widehat{m}-m\big)=\sum_{i=1}^{n}\frac{m\cdot n^{1/2}\int_{t_1}^{t_2}\frac{dM_{2i}(s)}{h_2(s)}}{\left[1-\sum_{l=1}^{n}\int_{0}^{t_2}\frac{dM_{2l}(s)}{h_2(s)}\right]}+o_p(1). \tag{3.16}$$

The ratios k and m are shown to converge weakly to a Gaussian process.

Since a concordance estimating equation $U(\theta)$ is a U-statistic, Fine et al. [2001] have shown that with application of the central limit theorem to the U-statistic and Slutsky's law, as $n\to\infty$, $n^{1/2}(\widehat{\theta}-\theta)$ has a limiting distribution with variance $I^{-2}J$ which can be consistently estimated by $\widehat{I}=n^{-2}\sum_{i<j}W(\widetilde{T}'_{1ij},\widetilde{T}'_{2ij})D_{ij}(1+\widehat{\theta})^{-2}$ and $\widehat{J}=2n^{-3}\sum_{k<l<m}(\widehat{Q}_{kl}\widehat{Q}_{km}+\widehat{Q}_{kl}\widehat{Q}_{lm}+\widehat{Q}_{lm}\widehat{Q}_{km})$. An asymptotic normality of $\widehat{\theta}$ gives

$$n^{1/2}(\widehat{\theta}-\theta_0)=I^{-1}\left(n^{-3/2}\sum_{i<j}Q_{ij}\right)+o_p(1), \tag{3.17}$$

where I is the probability limit of $\widehat{I}$, and $Q_{ij}=W(\widetilde{T}'_{1ij},\widetilde{T}'_{2ij})D_{ij}\{\Delta_{ij}-(1+\theta_0)(1+2\theta_0)^{-1}\}$. Now the asymptotic distribution of the estimating equation $u(\gamma^S_{\alpha|t_1})$ is obtained by utilizing the finite-dimensional delta method, which shows that $n^{1/2}\{\widehat{u}(\gamma^S_{\alpha|t_1})-u(\gamma^S_{\alpha|t_1})\}$ is asymptotically equivalent to

$$\begin{aligned}
&\phi_1\big\{S_Z(t_1),S_2(t_1),\theta,k,m\big\}\big[n^{1/2}\{\widehat{S}_Z(t_1)-S_Z(t_1)\}\big]\\
&+\phi_2\big\{S_Z(t_1),S_2(t_1),\theta,k,m\big\}\big[n^{1/2}\{\widehat{S}_2(t_1)-S_2(t_1)\}\big]\\
&+\phi_3\big\{S_Z(t_1),S_2(t_1),\theta,k,m\big\}\big[n^{1/2}\{\widehat{\theta}-\theta\}\big]\\
&+\phi_4\big\{S_Z(t_1),S_2(t_1),\theta,k,m\big\}\big[n^{1/2}\{\widehat{k}-k\}\big]\\
&+\phi_5\big\{S_Z(t_1),S_2(t_1),\theta,k,m\big\}\big[n^{1/2}\{\widehat{m}-m\}\big],
\end{aligned}$$

where the function ϕ is differentiable at $(S_Z(t_1), S_2(t_1), \theta, k, m)^{\mathrm{T}}$, with the first derivative $\phi'_{(S_Z(t_1),S_2(t_1),\theta,k,m)} = \{\phi_1, \cdots, \phi_5\}$ shown in follows:

$$\begin{aligned}
\phi_1 = a^{-(1+1/c)}\Bigg\{&(a^{-1/c} - b^{-1/c} + 1)^c \cdot \big[(a^{-1/c} - (1-d^{1/c})b^{-1/c})^{-(1+c)} \\
&- (1-\alpha)(a^{-1/c} - (1-e^{1/c})b^{-1/c})^{-(1+c)}\big] - \frac{g(a,b,c,d,e)}{(a^{-1/c} - b^{-1/c} + 1)}\Bigg\}, \\
\phi_2 = \ &- b^{-(1+1/c)}\Bigg\{(a^{-1/c} - b^{-1/c} + 1)^c \cdot \big[(a^{-1/c} - (1-d^{1/c})b^{-1/c})^{-(1+c)} \\
&- (1-\alpha)(a^{-1/c} - (1-e^{1/c})b^{-1/c})^{-(1+c)}\big] - \frac{\phi(a,b,c,d,e)}{(a^{-1/c} - b^{-1/c} + 1)}\Bigg\}, \\
\phi_3 =&\phi(a,b,c,d,e)\Bigg\{\log(a^{-1/c} - b^{-1/c} + 1) + \frac{\log(a)a^{-1/c} - \log(b)b^{-1/c}}{c \cdot (a^{-1/c} - b^{-1/c} + 1)}\Bigg\} \\
&- (a^{-1/c} - b^{-1/c} + 1)^c \cdot \Bigg\{(a^{-1/c} - (1-d^{1/c})b^{-1/c})^{-c} \cdot \\
&\Bigg[\log(a^{-1/c} - (1-d^{1/c})b^{-1/c}) + \frac{\log(a)a^{-1/c} - \log(d)d^{1/c}b^{-1/c} - (1-d^{1/c})\log(b)b^{-1/c}}{c \cdot (a^{-1/c} - (1-d^{1/c})b^{-1/c})}\Bigg] \\
&- (1-\alpha)(a^{-1/c} - (1-e^{1/c})b^{-1/c})^{-c} \cdot \Bigg[\log(a^{-1/c} - (1-e^{1/c})b^{-1/c}) \\
&+ \frac{\log(a)a^{-1/c} - \log(e)e^{1/c}b^{-1/c} - (1-e^{1/c})\log(b)b^{-1/c}}{c \cdot (a^{-1/c} - (1-e^{1/c})b^{-1/c})}\Bigg]\Bigg\}, \\
\phi_4 =& -(a^{-1/c} - b^{-1/c} + 1)^c \left\{[a^{-1/c} - (1-d^{1/c})b^{-1/c}]^{-1-c} d^{1/c-1} b^{-1/c}\right\}, \\
\phi_5 =&(a^{-1/c} - b^{-1/c} + 1)^c \left\{[a^{-1/c} - (1-e^{1/c})b^{-1/c}]^{-1-c} e^{1/c-1} b^{-1/c}\right\}.
\end{aligned}$$

Therefore, we obtain

$$n^{1/2}\left\{\widehat{u}(\gamma^S_{\alpha|t_1}) - u(\gamma^S_{\alpha|t_1})\right\} = n^{-3/2}\sum_{i<j} V_{ij} + o_p(1), \tag{3.18}$$

where

$$
\begin{aligned}
V_{ij} = &- \phi_1 \cdot S_Z(t_1) \int_0^{t_1} \frac{dM_{Zi}(s) + dM_{Zj}(s)}{h_Z(s)} - \phi_2 \cdot S_2(t_1) \int_0^{t_1} \frac{dM_{2i}(s) + dM_{2j}(s)}{h_2(s)} \\
&+ \phi_3 \cdot I^{-1} Q_{ij} + \phi_4 \cdot k \int_{t_1}^{t_2+\gamma^S_{\alpha|t_1}} \frac{dM_{2i}(s) + dM_{2j}(s)}{h_2(s)} \left[1 - \sum_{l=1}^{n} \int_0^{t_2+\gamma^S_{\alpha|t_1}} \frac{dM_{2l}(s)}{h_2(s)} \right]^{-1} \\
&+ \phi_5 \cdot m \int_{t_1}^{t_2} \frac{dM_{2i}(s) + dM_{2j}(s)}{h_2(s)} \left[1 - \sum_{l=1}^{n} \int_0^{t_2} \frac{dM_{2l}(s)}{h_2(s)} \right]^{-1}.
\end{aligned}
\tag{3.19}
$$

Finite dimensional convergence for each term of V_{ij} gives $n^{1/2}\{\widehat{u}(\gamma^S_{\alpha|t_1}) - u(\gamma^S_{\alpha|t_1})\} \to N(0, \sigma_u^2)$, where

$$
\begin{aligned}
\sigma_u^2 &= E\left\{ n^{-3/2} \sum_{i<j} V_{ij} \right\}^2 \\
&= n^{-3} \left\{ \sum_{i<j} V_{ij}^2 + 2 \sum_{k<l<m} \left[V_{kl} V_{km} + V_{lm} V_{km} + V_{kl} V_{lm} \right] \right\}.
\end{aligned}
\tag{3.20}
$$

Replacing the unknown parameters by their consistent estimators in equation (3.20) gives

$$
\widehat{\sigma}_u^2 = n^{-3} \left\{ \sum_{i<j} \widehat{V}_{ij}^2 + 2 \sum_{k<l<m} \left[\widehat{V}_{kl} \widehat{V}_{km} + \widehat{V}_{lm} \widehat{V}_{km} + \widehat{V}_{kl} \widehat{V}_{lm} \right] \right\},
$$

where

$$
\begin{aligned}
\widehat{V}_{ij} = &- \phi_1\{\widehat{S}_Z(t_1), \widehat{S}_2(t_1), \widehat{\theta}, \widehat{k}, \widehat{m}\} \cdot \widehat{S}_Z(t_1) \int_0^{t_1} \frac{d\widehat{M}_{Zi}(s) + d\widehat{M}_{Zj}(s)}{\widehat{h}_Z(s)} \\
&- \phi_2\{\widehat{S}_Z(t_1), \widehat{S}_2(t_1), \widehat{\theta}, \widehat{k}, \widehat{m}\} \cdot \widehat{S}_2(t_1) \int_0^{t_1} \frac{d\widehat{M}_{2i}(s) + d\widehat{M}_{2j}(s)}{\widehat{h}_2(s)} \\
&+ \phi_3\{\widehat{S}_Z(t_1), \widehat{S}_2(t_1), \widehat{\theta}, \widehat{k}, \widehat{m}\} \cdot \widehat{I}^{-1} \widehat{Q}_{ij} \\
&+ \phi_4\{\widehat{S}_Z(t_1), \widehat{S}_2(t_1), \widehat{\theta}, \widehat{k}, \widehat{m}\} \cdot \widehat{k} \int_{t_1}^{t_2+\widehat{\gamma}^S_{\alpha|t_1}} \frac{d\widehat{M}_{2i}(s) + d\widehat{M}_{2j}(s)}{\widehat{h}_2(s)} \left[1 - \sum_{l=1}^{n} \int_0^{t_2+\widehat{\gamma}^S_{\alpha|t_1}} \frac{d\widehat{M}_{2l}(s)}{\widehat{h}_2(s)} \right]^{-1} \\
&+ \phi_5\{\widehat{S}_Z(t_1), \widehat{S}_2(t_1), \widehat{\theta}, \widehat{k}, \widehat{m}\} \cdot \widehat{m} \int_{t_1}^{t_2} \frac{d\widehat{M}_{2i}(s) + d\widehat{M}_{2j}(s)}{\widehat{h}_2(s)} \left[1 - \sum_{l=1}^{n} \int_0^{t_2} \frac{d\widehat{M}_{2l}(s)}{\widehat{h}_2(s)} \right]^{-1}.
\end{aligned}
$$

$\widehat{M}_{Zi}(t)$ and $\widehat{M}_{2i}(t)$ are obtained by using Nelson-Aalen estimators of Λ_Z and Λ_2 in equation (3.12). Hence, a $100 \times (1-\alpha)\%$ confidence interval for $\gamma^S_{\alpha|t_1}$ can be obtained by inverting the estimating equation $u(\gamma^S_{\alpha|t_1})$

$$\{\gamma^S_{\alpha|t_1} : \widehat{\sigma}_u^{-2}\widehat{u}(\gamma^S_{\alpha|t_1})^2 < \chi^2_{1,1-\alpha}\}, \tag{3.21}$$

where $\chi^2_{1,1-\alpha}$ is the $100\times(1-\alpha)$th percentile of the χ^2 distribution with 1 degree of freedom.

It is known that using the estimating equation for inference on $\gamma^S_{\alpha|t_1}$ has the merit to avoid estimation of underlying probability density function of bivariate failure times under censoring [Jeong et al., 2008, Jung et al., 2009]. However, having the model-based variance formula would be also worthwhile to compare performances of the estimators from both approaches. Suppose there exists a function, Ψ, such that

$$\Psi\left(u(\gamma^S_{\alpha|t_1})\right) = S^{-1}_{2|1}\left(u(\gamma^S_{\alpha|t_1}) + (1-\alpha)S_{2|1}(t_2 \mid t_1) \mid t_1\right) - t_2, \tag{3.22}$$

where $S^{-1}_{2|1}$ is the inverse function of $S_{2|1}(\cdot \mid t_1)$, and the inverse function $S^{-1}_{2|1}$ has the derivatives of order n, that is $S^{-1(n)}_{2|1}(x|t_1) = \frac{d^n}{dx^n}S^{-1}_{2|1}(x \mid t_1)$ exists, then the Taylor polynomial of order n gives

$$\Psi\left(u(\gamma^S_{\alpha|t_1})\right) = \gamma^{S,0}_{\alpha|t_1} + \sum_{n=1}^{\infty}\left\{\frac{[u(\gamma^S_{\alpha|t_1}) - u(\gamma^{S,0}_{\alpha|t_1})]^n}{n!}S^{-1(n)}_{2|1}\left(u(\gamma^{S,0}_{\alpha|t_1}) + (1-\alpha)S_{2|1}(t_2 \mid t_1) \mid t_1\right)\right\}. \tag{3.23}$$

The approximation of a first-order Taylor series expansion to the variance of $\phi(\widehat{u}(\gamma^S_{\alpha|t_1}))$ (delta method) gives

$$n^{1/2}\left[\Psi(\widehat{u}(\gamma^S_{\alpha|t_1})) - \Psi(u(\gamma^S_{\alpha|t_1}))\right] \equiv n^{1/2}\left[\widehat{\gamma}^S_{\alpha|t_1} - \gamma^S_{\alpha|t_1}\right] \to N\left(0, \left[\Psi'(u(\gamma^S_{\alpha|t_1}))\right]^2\sigma_u^2\right), \tag{3.24}$$

where

$$\begin{aligned}\Psi'(u(\gamma^S_{\alpha|t_1})) &= \left[\frac{\partial S_{2|1}(t_2 + \gamma^{S,0}_{\alpha|t_1} \mid t_1)}{\partial(t_2 + \gamma^{S,0}_{\alpha|t_1})}\right]^{-1} \\ &= \frac{P(T_1 > t_1)}{P(T_1 > t_1, T_2 > t_2 + \gamma^{S,0}_{\alpha|t_1} + \Delta_t) - P(T_1 > t_1, T_2 > t_2 + \gamma^{S,0}_{\alpha|t_1})/\Delta_t}.\end{aligned}$$

When $\Delta_t \to 0$, $\Psi'(u(\gamma^S_{\alpha|t_1}))$ is asymptotically equivalent to

$$\frac{S_1(t_1)}{\int_{t_1}^{\infty} f(x, t_2 + \gamma^{S,0}_{\alpha|t_1})dx}. \tag{3.25}$$

For estimation of the joint density $f(x, y)$ in (3.25), a kernel smoothing is suggested to approximate the unknown density. Performance of this methodology depends on the choice of a smoothing parameter, i.e. bandwidth [Simonoff, 1996, Wand and Jones, 1993]. The kernel density estimator is defined by $\widehat{f}(\mathbf{t}; \mathbf{H}) = n^{-1} \sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{t} - \mathbf{T}_i)$, where $\mathbf{t} = (t_1, t_2)^{\mathrm{T}}$ and $K_{\mathbf{H}}(\mathbf{t}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2}\mathbf{t})$. Here the kernel K is a symmetric probability density function and $\mathbf{H}$ is a symmetric and positive-definite bandwidth matrix to determine performance of $\widehat{f}$. So the variance of $\widehat{\gamma}^S_{\alpha|t_1}$ can be obtained by

$$\frac{n^{-1}\widehat{\sigma}_u^2 \widehat{S}_1(t_1)^2}{\left[\int_{t_1}^{\infty} \widehat{f}(x, t_2 + \widehat{\gamma}^S_{\alpha|t_1})dx\right]^2}, \tag{3.26}$$

and a 95% confidence interval can be constructed based on the variance formula. Although the estimated variance formula involves the unknown density function, it can get close to optimal approximation by means of kernel smoothing with appropriate bandwidth selection and moderate sample sizes to reduce potential bias.

3.4 REGRESSION MODEL

3.4.1 Estimation of Regression Coefficients

Jung et al. [2009] proposed a time-specific log-linear regression method on quantile residual lifetime and evaluated the test statistic without estimating the variance-covariance matrix of the regression estimators. We adopt the similar approach but modify the test statistic accounting for the condition of $T_1 > t_1$. A regression setting for the conditional α-quantile residual lifetime is studied for a single sample. Suppose $\gamma^S_{\alpha|t_1}$ defines the α-quantile residual

lifetimes among patients who have recurrence-free up to time t_1. It can be estimated through an estimating equation expressed as

$$\frac{Pr(T_{2i} > t_2 + \gamma^S_{\alpha|t_1} \mid T_{1i} > t_1)}{Pr(T_{2i} > t_2 \mid T_{1i} > t_1)} = 1 - \alpha. \tag{3.27}$$

Consider a linear regression model for the α-quantile residual lifetimes for patients who survived beyond time t_2 conditioned on $T_1 > t_1$, on a log-scale,

$$\alpha\text{-quantile}\{\log(T_{2i} - t_2) \mid T_{2i} > t_2, T_{1i} > t_1, \mathbf{X}_i\} = \boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i, \tag{3.28}$$

where $\boldsymbol{\beta}_\alpha = (\beta_{\alpha,0}, \beta_{\alpha,1}, \cdots, \beta_{\alpha,p})^{\mathrm{T}}$ denotes a vector of the regression coefficients, and $\mathbf{X}_i = (1, X_{1i}, \cdots, X_{pi})^{\mathrm{T}}$ is a vector of covariates for a subject i. Let $G(t) = Pr(C > t)$ denote the censoring distribution and $E\{I(T_2 > t_2)\} = S_2(t_2)G(t_2)$ assume the terminal event time T_2 is independent of the censoring variable C. Assuming conditional independence between $T_{2i} \mid T_{1i}$ and C_i given $\mathbf{X}_i$ and independence between C_i and $\mathbf{X}_i$, the equation (3.27) can be rewritten as

$$\frac{E\big[I\{T'_{2i} > t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i)\} \mid T_{1i} > t_1, \mathbf{X}_i\big]}{Pr(C_i > t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i) \mid T_{1i} > t_1, \mathbf{X}_i)} \times \frac{Pr(C_i > t_2 \mid T_{1i} > t_1, \mathbf{X}_i)}{E\big[I(T'_{2i} > t_2) \mid T_{1i} > t_1, \mathbf{X}_i\big]} = 1 - \alpha$$

Therefore,

$$\begin{aligned}
&E\big[I\{T'_{2i} > t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i)\} \;\big|\; T_{1i} > t_1, \mathbf{X}_i\big] \\
&= (1-\alpha)\frac{Pr(C_i > t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i) \mid T_{1i} > t_1, \mathbf{X}_i)}{Pr(C_i > t_2 \mid T_{1i} > t_1, \mathbf{X}_i)} E\big[I(T'_{2i} > t_2) \mid T_{1i} > t_1, \mathbf{X}_i\big] \\
&= E\left[(1-\alpha)\frac{G\{t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i)\}}{G(t_2)} I(T'_{2i} > t_2) \,\middle|\, T_{1i} > t_1, \mathbf{X}_i\right]
\end{aligned}$$

Mimicking the least squares principle from the ordinary multiple linear regression model, the estimating equation for the regression parameter $\boldsymbol{\beta}_\alpha$ is given by

$$\mathbb{B}_{\alpha,t_2}(\boldsymbol{\beta}_\alpha) = \sum_{i=1}^{n} I(T_{1i} > t_1)\mathbf{X}_i \left[\frac{I\{T'_{2i} > t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i)\}}{\widehat{G}\{t_2 + \exp(\boldsymbol{\beta}^{\mathrm{T}}_{\alpha}\mathbf{X}_i)\}} - (1-\alpha)\frac{I(T'_{2i} > t_2)}{\widehat{G}(t_2)}\right] \approx 0. \tag{3.29}$$

The invariance property of the log-transformed quantile allows the estimating equation (3.29) to evaluate on the original scale of the observed survival data. A solution $\widehat{\boldsymbol{\beta}}_\alpha$ to the equation (3.29) can be obtained by minimizing the function $\|\mathbb{B}_{\alpha,t_2}(\boldsymbol{\beta}_\alpha)\|$, where $\|\cdot\|$ is defined as the

square root of the sum of squares. Under certain regularity conditions, $\widehat{\boldsymbol{\beta}}_\alpha$ is shown to be a consistent estimator for the true value $\boldsymbol{\beta}_\alpha^0$ according to Jung et al. [2009](Web Appendix A).

3.4.2 Test Statistic and Confidence Interval

To test a statistical hypothesis $H_0 : \boldsymbol{\beta}_\alpha = \boldsymbol{\beta}_\alpha^0$, it is difficult to use a Wald-type statistic based on $\widehat{\boldsymbol{\beta}}_\alpha$ because the corresponding limiting covariance matrix depends on the unknown density functions. In particular, with censored data, the covariance matrix cannot be estimated well nonparametrically. Therefore, the estimating equation $\mathbb{B}_{\alpha,t_2}(\boldsymbol{\beta}_\alpha)$ is used directly to make inference on $\widehat{\boldsymbol{\beta}}_\alpha$. In a similar proof to Web Appendix B of Jung et al. [2009], we show that the distribution of $n^{-1/2}\mathbb{B}_{\alpha,t_2}(\boldsymbol{\beta}_\alpha^0)$ is approximately normal with mean zero and variance-covariance matrix $\Gamma_\alpha = \lim_{n\to\infty} n^{-1}\sum_{i=0}^n \xi_{\alpha,i}\,\xi_{\alpha,i}^{\mathrm{T}}$, where

$$\begin{aligned}\xi_{\alpha,i} =& I(T_{1i} > t_1)\mathbf{X}_i\left[\frac{I\{T'_{2i} > t_2 + \exp(\boldsymbol{\beta}_\alpha^{0,\mathrm{T}}\mathbf{X}_i)\}}{G\{t_2 + \exp(\boldsymbol{\beta}_\alpha^{0,\mathrm{T}}\mathbf{X}_i)\}} - (1-\alpha)\frac{I(T'_{2i} > t_2)}{G(t_2)}\right] \\ &+ \int_0^\infty G^{-1}(s)\int_0^s h_2^{-1}(v)\{dI(T'_{2i} \le v, \delta_{2i} = 0) - I(T'_{2i} \ge v)d\Lambda_G(v)\}d\mathbf{q}_1(s) \qquad (3.30) \\ &- \mathbf{q}_2(t_2)\int_0^{t_2} h_2^{-1}(s)\{dI(T'_{2i} \le s, \delta_{2i} = 0) - I(T'_{2i} \ge s)d\Lambda_G(s)\},\end{aligned}$$

where $\Lambda_G(\cdot)$ is the cumulative hazard function for the censoring distribution. $h_2(t) = \lim_{n\to\infty} n^{-1}\sum_{i=1}^n Y_{2i}(t)$, $\mathbf{q}_1(s) = \lim_{n\to\infty} n^{-1}\sum_{i=1}^n \mathbf{X}_i I(T_{1i} > t_1)I\{\min(s, T'_{2i}) > t_2 + \exp(\boldsymbol{\beta}_\alpha^{0,\mathrm{T}}\mathbf{X}_i)\}$, and

$$\mathbf{q}_2(t_2) = \lim_{n\to\infty} G^{-1}(t_2)\sum_{i=1}^n \mathbf{X}_i I(T_{1i} > t_2)I(T'_{2i} > t_2). \qquad (3.31)$$

A consistent estimator $\widehat{\Gamma}_\alpha$ for the limiting covariance matrix of $n^{-1/2}\mathbb{B}_{\alpha,t_2}(\boldsymbol{\beta}_\alpha^0)$ can be obtained as

$$\widehat{\Gamma}_\alpha = n^{-1}\sum_{i=1}^n \widehat{\xi}_{\alpha,i}\,\widehat{\xi}_{\alpha,i}^{\mathrm{T}},$$

where

$$\begin{aligned}
\widehat{\xi}_{\alpha,i} = & \left[\frac{I\{T'_{2i} > t_2 + \exp(\widehat{\boldsymbol{\beta}}_\alpha^{\mathrm{T}} \mathbf{X}_i)\}}{\widehat{G}\{t_2 + \exp(\widehat{\boldsymbol{\beta}}_\alpha^{\mathrm{T}} \mathbf{X}_i)\}} - (1-\alpha) \frac{I(T'_{2i} > t_2)}{\widehat{G}(t_2)} \right] I(T_{1i} > t_1) \mathbf{X}_i \\
& + \sum_{l=1}^{n} \left[I(T_{1l} > t_1) \mathbf{X}_l \frac{I\{T'_{2l} > t_2 + \exp(\widehat{\boldsymbol{\beta}}_\alpha^{\mathrm{T}} \mathbf{X}_l)\}}{\widehat{G}\{t_2 + \exp(\widehat{\boldsymbol{\beta}}_\alpha^{\mathrm{T}} \mathbf{X}_l)} \right] \\
& \times \left[\frac{(1-\delta_{2i}) I\{T'_{2i} \le t_2 + \exp(\widehat{\boldsymbol{\beta}}_\alpha^{\mathrm{T}} \mathbf{X}_l)\}}{\sum_{m=1}^{n} I(T'_{2m} > T'_{2i})} - \sum_{j=1}^{n} \frac{(1-\delta_{2j}) I\{T'_{2j} \le \min(T'_{2i}, t_2 + \exp(\widehat{\boldsymbol{\beta}}_\alpha^{\mathrm{T}} \mathbf{X}_l))\}}{\left[\sum_{m=1}^{n} I(T'_{2m} > T'_{2j})\right]^2} \right] \\
& - \sum_{l=1}^{n} \left[I(T_{1l} > t_1) \mathbf{X}_l \frac{(1-\alpha) I(T'_{2l} > t_2)}{n \widehat{G}(t_2)} \right] \\
& \times \left[\frac{(1-\delta_{2i}) I(T'_{2i} \le t_2)}{\sum_{m=1}^{n} I(T'_{2m} > T'_{2i})} - \sum_{j=1}^{n} \frac{(1-\delta_{2j}) I\{T'_{2j} \le \min(T'_{2i}, t_2)\}}{\left[\sum_{m=1}^{n} I(T'_{2m} > T'_{2j})\right]^2} \right].
\end{aligned} \tag{3.32}$$

Wei et al. [1990] used the linear rank statistics for censored data to make inference about a subset of the regression coefficients in the linear model without estimating the covariance matrix. Ying et al. [1995] proposed semiparametric procedures to regress the median of the failure time under censoring on potential covariates. By using their similar arguments, a test statistic for testing H_0 would be

$$n^{-1} \mathbb{B}_{\alpha,t_2}^{\mathrm{T}}(\boldsymbol{\beta}_\alpha) \widehat{\Gamma}_\alpha^{-1} \mathbb{B}_{\alpha,t_2}(\boldsymbol{\beta}_\alpha),$$

which approximately follows a χ^2 distribution with $p+1$ degrees of freedom. Consider a partition of the regression coefficients, $\boldsymbol{\beta}_\alpha = (\boldsymbol{\beta}_\alpha^{(1)}, \boldsymbol{\beta}_\alpha^{(2)})^{\mathrm{T}}$, where $\boldsymbol{\beta}_\alpha^{(1)}$ is a $r \times 1$ vector. Suppose that $\widehat{\boldsymbol{\beta}}_\alpha^{(1)}$ and $\widehat{\boldsymbol{\beta}}_\alpha^{(2)}$ are the corresponding estimates, and we are only interested in testing the hypothesis $\widetilde{H}_0 : \boldsymbol{\beta}_\alpha^{(1)} = \boldsymbol{\beta}_\alpha^{0,(1)}$, a specific vector, against a general alternative. For the test statistic, the minimum dispersion statistic [Basawa and Koul, 1988] is considered as

$$V\left(\boldsymbol{\beta}_\alpha^{0,(1)}\right) = \min_{\boldsymbol{\beta}_\alpha^{(2)}} \left\{ n^{-1} \mathbb{B}_{\alpha,t_2}^{\mathrm{T}} \left(\left[\boldsymbol{\beta}_\alpha^{0,(1)}, \boldsymbol{\beta}_\alpha^{(2)} \right]^{\mathrm{T}} \right) \widehat{\Gamma}_\alpha^{-1} \mathbb{B}_{\alpha,t_2} \left(\left[\boldsymbol{\beta}_\alpha^{0,(1)}, \boldsymbol{\beta}_\alpha^{(2)} \right]^{\mathrm{T}} \right) \right\}. \tag{3.33}$$

It can be shown that equation (3.33) is approximately χ^2 distribution with r degrees of freedom. We reject $\widetilde{H}_0$ for a large value of $V\left(\boldsymbol{\beta}_{\alpha}^{0,(1)}\right)$. A $(1-\alpha)$ confidence interval for $\boldsymbol{\beta}_{\alpha}^{(1)}$ can be obtained by

$$\left\{\boldsymbol{\beta}_{\alpha}^{(1)} : V\left(\boldsymbol{\beta}_{\alpha}^{(1)}\right) < \chi^2_{r,1-\alpha}\right\}, \tag{3.34}$$

where $\chi^2_{r,1-\alpha}$ is the upper α-percentile of a chi-squared distribution with r degrees of freedom.

3.5 SIMULATION STUDY

Simulation studies were conducted to evaluate performance of the proposed estimator of conditional quantile residual lifetime. Specifically, n pairs of (T_1, T_2) were generated from the Clayton bivariate exponential distribution. The marginals S_1 and S_2 followed an exponential distribution with the rate parameter equal to 0.5. The independent censoring time C was generated from a uniform$(0, c)$ distribution, where c determined the censoring proportion. Using the transformation from the model (3.4), with $T_1 = -2\log(1-u_1)$, T_2 was generated from the equation of $P(T_2 > t_2 \mid T_1 = t_1) = 1 - u_2$, where u_1 and u_2 were uniform(0,1) random variables. Therefore,

$$T_2 = 2\theta \log\left[(1-a) + a(1-u_2)^{-1/(1+\theta)}\right],$$

where $a = (1-u_1)^{-1/\theta}$. A censoring variable was generated to censor the terminal event time and the observed terminal event time then censored the nonterminal event time, which yielded the observable data restricted to the upper wedge of a plane.

A total of 1000 iterations from a bivariate Clayton distribution were carried out in samples of size 200 and 300. The censoring parameter c was chosen to be 9.9, giving 20% censoring. A combination of association parameter θ=0.25, 0.5 or 1, along with multiple pairs of fixed time points, were used to examine the performance of the proposed estimator. Fixed time points were chosen at values of 0.44, 0.58, 1.03, 1.28, and 1.79, corresponding to marginal survival probabilities of 0.8, 0.75, 0.60, 0.50 and 0.40 from an exponential distribution with rate 0.5. The bandwidth matrix was fixed at diag$(0.3, 0.3)$ for estimating bivariate density

function in the variance formula. Under the assumption of $t_1 \leq t_2$, the proposed method was assessed at six pairs of fixed time points for the approximation of conditional median residual lifetime, i.e., $\alpha = 0.5$ from the Clayton distribution. For a regression setting of the conditional quantile residual lifetime for semi-competing risks data, we have considered the median residual regression

$$\text{med}(T_{2i} - t_2 \mid T_{2i} > t_2, T_{1i} > t_1, x_{1i}) = \exp(\beta_{0.5,0} + \beta_{0.5,1} x_{1i}), \tag{3.35}$$

where x_{1i} is a binary covariate taking values of 0 or 1. The regression coefficients $\beta_{0.5,0}$ and $(\beta_{0.5,0} + \beta_{0.5,1})$ are interpreted as the median residual lifetimes on a log-scale for the control $(x = 0)$ and treatment $(x = 1)$ groups at time t_2, respectively. Under $\widetilde{H}_0 : \beta_{0.5,1} = 0$, the equation (3.35) was written in terms of median residual lifetime

$$\begin{aligned} \gamma^S_{0.5|t_1} \equiv \exp(\beta_{0.5,0}) &= S^{-1}_{T_2|T_1}\big\{(1/2)S(t_2 \mid t_1) \mid t_1\big\} - t_2 \\ &= 2\theta \ \log\big\{2^{1/\theta}(e^{t_1/2\theta} + e^{t_2/2\theta} - 1) - e^{t_1/2\theta} + 1\big\} - t_2. \end{aligned} \tag{3.36}$$

Note that at the origin of time axis, i.e., $t_1 = t_2 = 0$, $\exp(\beta_{0.5,0}) = 2\log(2) = 1.386$. The empirical distribution of the regression parameters $\beta_{0.5,0}$ and $\beta_{0.5,1}$ was evaluated via the mean and the standard deviation of the parameter estimates. The grid search method was used to determine the minimum score of the equation (3.29). The true parameter values of $\beta_{0.5,0}$ for all six pairs of fixed time points were obtained from the equation (3.36). The true value of $\beta_{0.5,1}$ must be 0 because the survival distribution was identical for both control and treatment groups. Similar settings were employed for testing the null hypothesis $\widetilde{H}_0 : \beta_{0.5,1} = 0$ except that we compared Type I error probabilities for 0% and 20% censoring proportions at various time points.

For each pair of fixed time points, the estimator of the median conditional quantile residual lifetime was evaluated by the empirical mean (EST) and standard deviation (SD) of estimates, the mean of estimated standard errors (SE), mean square errors (MSE), the percent of bias (%Bias), and 95% coverage probabilities (Cov95). The weighted function $W_{a,b}$ for θ were equal to $T_{1,0.75}$ and $T_{2,0.75}$ which represent the 75th percentiles of the uncensored T_1 and T_2, respectively. In Table 3.1, the estimated conditional median residual lifetimes $\gamma^S_{\alpha|t_1}$ are virtually unbiased with the percent of bias between -1 and 1 for each combination

of parameter settings. The mean of the estimated standard errors are well approximated to the empirical standard deviation. Regardless of the association level, increasing the value of t_2 also increases the standard deviation of the estimates and the mean square error. It is observed that 95% coverage probabilities from model-based variance formula are slightly less efficient than ones directly from the estimating equation approach when association becomes less dependent, i.e. increasing θ. With a sample size of 300, 95% coverage probabilities from model-based and equation-based variance estimation are improved; there is a great improvement especially when the association becomes less dependent. The result is shown in Table 3.2. Overall, the coverage probabilities tend to approximate the nominal level reasonably well when increasing a sample size.

Table 3.3 presents the results for the estimated regression coefficients using the grid search method. The mean values of the estimates of $\beta_{0.5,0}$ and $\beta_{0.5,1}$ are close to their true values for given fixed time points. As expected, the MSE increased as the value of t_2 increased for both estimates of $\beta_{0.5,0}$ and $\beta_{0.5,1}$. The results in Table 3.4 indicate that the test based on the minimum dispersion statistic tends to be slightly conservative, which is consistent with the observations presented by Jeong et al. [2008] and Jung et al. [2009].

Table 3.1: Simulation summary of $\widehat{\gamma}^{S}_{0.5|t_1}$ reported as mean estimate, mean standard error, empirical standard deviation, MSE, %Bias, and empirical 95% coverage probability:(1)model-based, (2)equation-based for 1000 iterations in the sample size of 200.

θ	n	t_1	t_2	True Value	EST	(SE)	SD	MSE	%Bias	Cov95 (1)	Cov95 (2)
0.25	200	0.44	0.44	1.605	1.607	(0.184)	0.186	0.035	0.101	93.8	94.7
			0.58	1.560	1.557	(0.188)	0.191	0.036	-0.166	94.2	93.9
			1.28	1.435	1.443	(0.216)	0.225	0.051	0.587	92.9	93.4
		1.03	1.03	1.685	1.687	(0.232)	0.229	0.053	0.104	95.0	93.9
			1.28	1.588	1.596	(0.239)	0.247	0.061	0.495	93.7	94.8
			1.79	1.469	1.478	(0.259)	0.279	0.078	0.630	93.0	94.2
0.5	200	0.44	0.44	1.623	1.617	(0.183)	0.196	0.038	-0.344	92.0	92.5
			0.58	1.595	1.593	(0.189)	0.201	0.040	-0.101	93.3	92.4
			1.28	1.495	1.498	(0.224)	0.235	0.055	0.174	92.6	93.7
		1.03	1.03	1.780	1.770	(0.234)	0.251	0.063	-0.536	91.6	93.3
			1.28	1.705	1.704	(0.247)	0.263	0.069	-0.051	92.3	93.7
			1.79	1.590	1.599	(0.275)	0.289	0.084	0.603	92.3	93.3
1.0	200	0.44	0.44	1.575	1.571	(0.177)	0.197	0.039	-0.247	91.3	91.7
			0.58	1.562	1.558	(0.184)	0.202	0.041	-0.269	91.6	92.2
			1.28	1.512	1.513	(0.224)	0.245	0.060	0.033	91.9	93.5
		1.03	1.03	1.753	1.750	(0.228)	0.267	0.072	-0.158	90.8	90.5
			1.28	1.713	1.715	(0.243)	0.277	0.077	0.099	89.2	92.4
			1.79	1.644	1.654	(0.277)	0.307	0.095	0.613	91.7	92.9

Table 3.2: Simulation summary of $\widehat{\gamma}^{S}_{0.5|t_1}$ reported as mean estimate, mean standard error, empirical standard deviation, MSE, %Bias, and empirical 95% coverage probability:(1)model-based, (2)equation-based for 1000 iterations in the sample size of 300.

θ	n	t_1	t_2	True Value	EST	(SE)	SD	MSE	%Bias	Cov95 (1)	Cov95 (2)
0.25	300	0.44	0.44	1.605	1.608	(0.152)	0.154	0.024	0.200	94.4	94.2
			0.58	1.560	1.562	(0.155)	0.159	0.025	0.172	94.1	94.0
			1.28	1.435	1.443	(0.178)	0.185	0.034	0.589	94.0	93.9
		1.03	1.03	1.685	1.686	(0.190)	0.186	0.034	0.066	96.2	95.2
			1.28	1.588	1.593	(0.196)	0.196	0.039	0.304	95.6	95.2
			1.79	1.469	1.470	(0.213)	0.224	0.050	0.087	93.0	94.0
0.5	300	0.44	0.44	1.623	1.623	(0.151)	0.158	0.025	0.017	93.4	93.6
			0.58	1.595	1.595	(0.156)	0.163	0.027	-0.027	94.3	94.3
			1.28	1.495	1.496	(0.184)	0.186	0.035	0.011	94.8	94.1
		1.03	1.03	1.780	1.772	(0.193)	0.194	0.038	-0.435	93.7	92.9
			1.28	1.705	1.700	(0.203)	0.207	0.043	-0.278	93.8	93.4
			1.79	1.590	1.594	(0.227)	0.234	0.055	0.262	93.3	94.3
1.0	300	0.44	0.44	1.575	1.571	(0.146)	0.155	0.024	-0.242	92.8	92.2
			0.58	1.562	1.561	(0.152)	0.159	0.025	-0.109	92.7	92.8
			1.28	1.512	1.514	(0.185)	0.191	0.037	0.123	93.5	95.3
		1.03	1.03	1.753	1.751	(0.188)	0.214	0.046	-0.115	90.8	92.9
			1.28	1.713	1.716	(0.201)	0.218	0.048	0.138	91.9	93.9
			1.79	1.644	1.654	(0.229)	0.243	0.059	0.595	92.6	93.8

Table 3.3: Simulation summary of the empirical estimates of the regression parameters $\beta_{0.5,0}$ and $\beta_{0.5,1}$ reported as mean estimate, empirical standard deviation, and MSE by using the grid search method for 1000 iterations.

θ	t_1	t_2	True $\beta_{0.5,0}$	EST($\beta_{0.5,0}$)	SD	MSE	EST($\beta_{0.5,1}$)	SD	MSE
0.25	0.44	0.44	0.473	0.452	0.182	0.034	0.012	0.269	0.072
		0.58	0.445	0.419	0.198	0.040	0.015	0.289	0.084
		1.28	0.361	0.327	0.245	0.061	0.017	0.364	0.132
	1.03	1.03	0.522	0.500	0.219	0.048	0.002	0.330	0.109
		1.28	0.462	0.435	0.250	0.063	0.009	0.376	0.141
		1.79	0.384	0.331	0.465	0.219	0.028	0.563	0.318
0.5	0.44	0.44	0.484	0.458	0.194	0.038	0.014	0.290	0.084
		0.58	0.467	0.439	0.205	0.043	0.014	0.303	0.092
		1.28	0.402	0.354	0.275	0.078	0.035	0.399	0.160
	1.03	1.03	0.577	0.559	0.283	0.080	-0.008	0.397	0.157
		1.28	0.534	0.506	0.295	0.088	0.003	0.422	0.178
		1.79	0.464	0.416	0.526	0.279	-0.015	0.625	0.391
1.0	0.44	0.44	0.454	0.422	0.215	0.047	0.024	0.320	0.103
		0.58	0.446	0.419	0.222	0.050	0.009	0.330	0.109
		1.28	0.413	0.371	0.293	0.088	0.025	0.418	0.175
	1.03	1.03	0.561	0.549	0.321	0.103	-0.021	0.443	0.196
		1.28	0.538	0.514	0.359	0.130	-0.0001	0.501	0.251
		1.79	0.497	0.445	0.609	0.374	-0.0005	0.756	0.572

Table 3.4: Simulation summary of the Type I error probabilities for testing the null hypothesis $H_0 : \beta_{0.5,1} = 0$ when the true parameter values are $\beta_{0.5,1} = 0$.

			Censoring proportion	
θ	t_1	t_2	0%	20%
0.25	0.44	0.44	0.036	0.026
		0.58	0.033	0.024
		1.28	0.033	0.026
	1.03	1.03	0.037	0.026
		1.28	0.038	0.026
		1.79	0.029	0.021
0.5	0.44	0.44	0.038	0.027
		0.58	0.034	0.025
		1.28	0.031	0.023
	1.03	1.03	0.043	0.027
		1.28	0.039	0.027
		1.79	0.030	0.024
1.0	0.44	0.44	0.038	0.043
		0.58	0.040	0.043
		1.28	0.039	0.020
	1.03	1.03	0.039	0.026
		1.28	0.034	0.023
		1.79	0.039	0.020

3.6 ANALYSIS OF B-14 BREAST CANCER DATA

A total of 2,817 eligible patients with estrogen receptor positive breast cancer and negative axillary lymph nodes were enrolled in B-14 phrase III breast cancer clinical trial from the National Surgical Adjuvant Breast and Bowel Project (NSABP). It was a randomized double-blind multi-center trial comparing the treatment of tamoxifen to placebo following surgery. In this trial, patients were randomly assigned to either the placebo arm (n=1,413) or the tamoxifen arm (n=1,404) following surgery. Patients tested with estrogen receptor positive indicate that their hormone receptor suggests the need of signals from estrogen to promote the growth of cancer cells. Tamoxifen blocks the effects of estrogen on cancer cells to prevent the tumor from growing. Initial results showed that tamoxifen treated women has a significantly better outcome than did those who received placebo [Fisher et al., 1989]. Several follow-up studies relating long-term findings provided substantial support for the initial results. [Fisher et al., 1996] found that the benefit of 5-year tamoxifen treatment persisted through 10 years of follow-up but no additional advantage was found for more than 5 years. The study of Fisher et al. [2004] showed chemotherapy plus tamoxifen was more effective than tamoxifen alone from a conjunctional trial.

The failure times in this example are time-to-recurrence, T_1 and time-to-death, T_2. We apply our proposed method to re-analyze this dataset and evaluate the effect of tamoxifen in breast cancer on 25% percentile residual lifetime. First, we estimate the 25th percentile residual lifetimes for each treatment group. Next, a covariate variable of whether patients receive placebo coded as 0 or tamoxifen coded as 1 is included in a regression model.

Figure 3.1a shows the estimated 25th percentile residual lifetimes in patients treated with tamoxifen and placebo. The 25th percentile residual lifetimes in tamoxifen group tend to be higher than the placebo group throughout all years even though the difference between the two groups becomes closer at the tail. It is interesting to observe that the estimated residual lifetime in tamoxifen group has a dramatic decreasing beyond year 9. This may imply that patients treated with tamoxifen and survived up to year 9 would expect their residual lifetime to drop around 23% at year 15 (9.47 versus 7.34) given no sign of recurrence by year 2. We also evaluate the 25th percentile residual lifetime for patients without recurrence by year

5. In Figure 3.1b, although there still exists a notable difference between tamoxifen and placebo, we can not spot any strong variation by each group throughout the whole examined years. Table 3.5 presents estimates of regression coefficients and p-values from the minimum dispersion statistic. The 25th percentile residual lifetimes are compared between tamoxifen and placebo at each time point. The results indicate that the treatment of tamoxifen has an influence on 25th percentile residual lifetimes from years 8 to year 10 when conditioning on year 2 (borderline significant at years 8 and 9, p-value=0.051, 0.049 respectively). However, we did not find any significant effect of tamoxifen on 25th percentile residual lifetime for patients who remain recurrence-free by year 5, shown in Table 3.6.

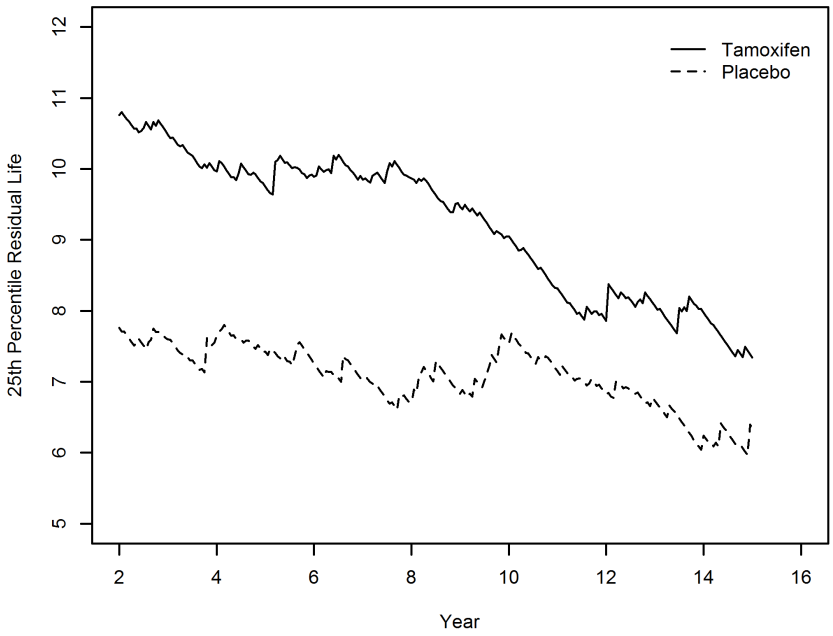
3.7 DISCUSSION

We have proposed a method for semi-competing risks data, by which the residual lifetime can be evaluated and compared at any fixed time point for patients who remain no sign of a morbidity after a prespecified number of years of follow-up for an original disease. Overall, the estimator is nearly unbiased for the conditional median residual lifetime, even though the results indicate that larger values of t_2 are associated with higher mean square error, as expected.

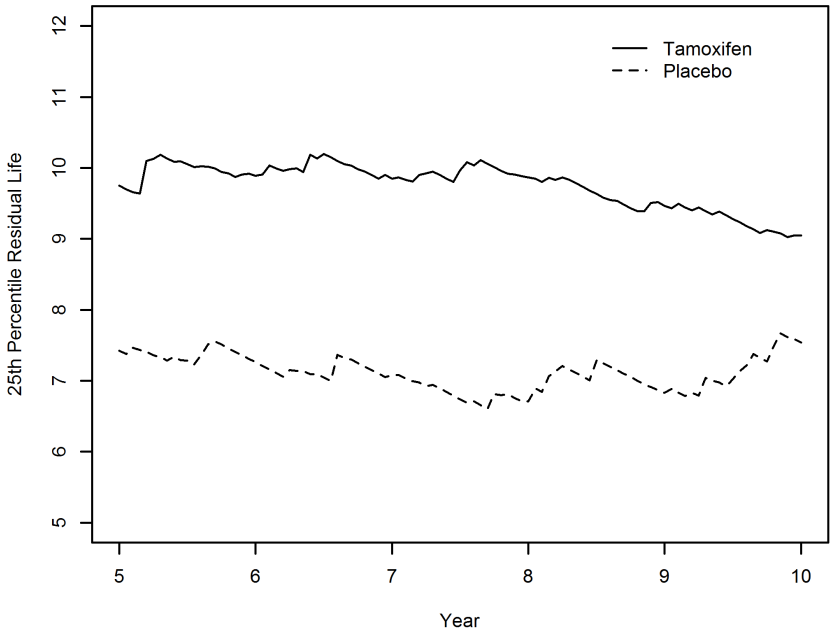
Choosing a value for the bandwidth for density estimation is arbitrary. There exists numerous methods that have good theoretical properties in selecting the bandwidth [Jones et al., 1996]. A modification of cross-validation method named by smoothed cross-validation involves a presmoothing of the pairwise differences of the observations [Hall et al., 1992]. However, the bandwidth selection is time-consuming and the variation of kernel smoothing estimation among iterations needs to be considered using a resampling mechanism. Even though the bandwidth in our simulation is fixed at a diagonal matrix, the estimates of the standard errors result in reasonable coverage probabilities in most cases. Larger sample size might decrease biases resulted from density estimation.

In addition, a regression model was proposed to associate the conditional residual lifetimes with selected covariates under right censoring among survivors without morbidities up to a specific time point. The conditional survival and censoring variables are assumed to be conditionally independent given covariates. However, if a censoring distribution is independent of covariates, then the conditional independence becomes unconditional which would hold in most randomized clinical trials.

The grid search method provided reasonable solutions to the estimating equation in the simulation studies. The proposed estimator was studied under a gamma frailty model. It might be also worthwhile to investigate it in different joint survival models.



(a) Given recurrence free by year 2



(b) Given recurrence free by year 5

Figure 3.1: The comparison of estimated 25th percentile conditional residual lifetimes between the treatment of tamoxifen and placebo in B-14 study

Table 3.5: Regression parameter estimates from the 25% residual life regression model with a single covariate of treatment and associated p-values, given $t_1 > 2$

t_2	min score	$\widehat{\beta}_{0.5,0}$	$\widehat{\beta}_{0.5,1}$	p-value
3	0.287	2.392	0.093	0.198
5	0.261	2.245	0.144	0.155
7	1.157	2.191	0.186	0.083
8	0.265	2.164	0.249	0.051
9	0.132	2.161	0.192	0.049
10	0.533	2.126	0.276	0.083
11	0.230	2.066	0.202	0.421
13	0.499	2.024	0.272	0.603
15	0.214	1.983	0.040	0.919

Table 3.6: Regression parameter estimates from the 25% residual life regression model with a single covariate of treatment and associated p-values, given $t_1 > 5$

t_2	min score	$\widehat{\beta}_{0.5,0}$	$\widehat{\beta}_{0.5,1}$	p-value
6	0.727	2.415	0.055	0.588
7	0.525	2.367	0.044	0.629
8	0.221	2.306	0.118	0.338
9	0.449	2.243	0.116	0.150
10	0.421	2.170	0.189	0.562
11	0.682	2.135	0.209	0.728
12	0.364	2.160	0.097	0.723

4.0 COMPARISON OF A NONPARAMETRIC ESTIMATOR USING BIVARIATE RIGHT CENSORED SURVIVAL DATA

In the previous chapter, we used a semiparametic method to model the conditional quantile residual lifetime. While a great amount of parametric and semiparametric methods have been proposed on the inference of residual lifetime, it is also desirable to have a completely nonparametric estimator that does not require any distributional assumptions. Suppose we are interested in the quantile residual life function of the occurrence time of a unit, say T_2, given the occurrence time of another unit, say T_1. To implement this interesting idea, the first step is to consider estimation of the conditional survival function of T_2 given $T_1 = t_1$ when T_1 is also subject to censoring.

Inference about such a conditional survival function has received attention in several scientific studies. Beran [1981] proposed a generalized Kaplan–Meier estimator to study regression problem with censored data. It is also referred to as conditional Kaplan–Meier estimator. The idea behind this estimator is to use the relationship between the distribution and the cumulative hazard function. In some cases such estimation of the conditional survival function is prerequisite for the estimator of bivariate survival function. The extension of Beran's estimator can be seen in literature. Akritas [1994] considered the problem of estimating the bivariate distribution by averaging estimates of the conditional distribution of T_2 given $T_1 = t_1$ over a range of values of t_1. Later, Akritas and Keilegom [2003] proposed estimators for marginal distributions and also required estimation of a condition survival function when the conditioning variable is subject to censoring. From the mechanism of Pruitt's estimator, those singly-censored observations have to redistribute the mass $1/n$ over their associated lines. It means that in order to estimate the bivariate survival function the conditional probability for each singly-censored observation has to be estimated in an

appropriate way. When the assumption of independence of $\mathbf{T}$ and $\mathbf{C}$ is relaxed to dependent censoring, a nonparametric estimator of the conditional survival function for bivariate failure times has been proposed by Lakhal-Chaieb et al. [2013].

In a bivariate failure time setting, let $\mathbf{T} = (T_1, T_2)$ represent the pair of survival times and $\mathbf{C} = (C_1, C_2)$ the pair of censoring times. Assume that $\mathbf{T}$ is independent of $\mathbf{C}$. The i.i.d. observed random variables are $\widetilde{T}_{ij} = \min(T_{ij}, C_{ij})$, and let $\delta_{ij} = I(T_{ij} \leq C_{ij})$ be the indicator function, for $i = 1, 2,\ j = 1, \cdots, n$. In this case, observed times $\widetilde{\mathbf{T}} = (\widetilde{T}_1, \widetilde{T}_2)$ could be singly or doubly censored, or both uncensored. Each observation in the bivariate censoring model can be illustrated by a region for the bivariate survival time $\mathbf{T}$. The survival time is obtained if both T_1 and T_2 are observed(uncensored), it is known to be on a line if only one of the survival times T_i is right-censored(singly-censored) and in a region of quadrant if both T_1 and T_2 are right-censored(doubly-censored).

The nonparametric estimation of the conditional survival function is considered when the conditioning variable is subject to right censoring. However, the main issue that needs to be addressed is the analyzed region in which the conditioning variable should be placed if it is censored. Since the purpose of considering the conditional survival function is to make inference on the residual lifetime of a unit given another is uncensored at a specific time point, we adopt the approach of using those pairs of observations in the region for which the value of the conditioning variable is uncensored. Therefore, the conditional survival function of T_2, given that T_1 is uncensored at time t_1 is defined as

$$S_{T_2|T_1=t_1}(t_2 \mid t_1) = Pr(T_2 > t_2 \mid T_1 = t_1), \tag{4.1}$$

and such setting can be applied vice versa for $S_{T_1|T_2=t_2}$. However, it is important to note that the problems of dealing with singly-censored observations may cause inconsistency for continuous distributions when using nonparametric maximum likelihood. In the following section, we will compare the sample behavior of two nonparametric estimators of conditional survival probability but restrict attention to the modified nonparametric maximum likelihood estimator (NPMLE) proposed by van der Laan [1994].

This chapter begins with a discussion of nonparametric estimators on conditional survival probability. Next, the inference of conditional quantile residual life function is addressed

and the nonparametric bootstrap method is discussed. Then a small simulation study is conducted to compare the nonparametric estimators with our proposed estimator from semi-competing risks data. Lastly, the potential future work of quantile residual life function is highlighted on the basis of bivariate failure data with censoring.

4.1 ESTIMATION OF CONDITIONAL SURVIVAL FUNCTION

Nonparametric estimation of conditional survival function (4.1) is discussed. The usual empirical distribution function for survival data is used to estimate the underlying distribution when there exists no censoring. Suppose there are two failure time variables and both of them are under independent censoring. The estimated distribution of T_2 can be explained by the available information regarding T_1, and vice versa. If the ith observation of T_2 is censored, the Kaplan–Meier estimator would redistribute its mass to all observations of T_2 which are larger than it. However, the information contained in the ith observation of T_1 is consequential if the correlation of the two variables presents positive.

4.1.1 Beran's estimator and nearest neighbor estimation

In a situation that T_1 is uncensored and T_2 is independent of C_2, kernel estimates of the conditional survival probability based on censored data were firstly introduced by Beran [1981]. It is defined as

$$\widehat{S}_{T_2|T_1=t_1}(t_2 \mid t_1) = \prod_{\widetilde{T}_{2i} \le t_2} \left\{ 1 - \frac{W_i(t_1)}{\sum_{j=1}^{n} W_j(t_1) I(\widetilde{T}_{2j} \ge \widetilde{T}_{2i})} \right\}^{\delta_{2i}}, \tag{4.2}$$

where

$$W_i(t) = \frac{K\left(\frac{t-\widetilde{T}_{1i}}{h}\right)}{\sum_{j=1}^{n} K\left(\frac{t-\widetilde{T}_{1j}}{h}\right)}.$$

Here K is a symmetric kernel function and h is a positive smoothing bandwidth. When T_1 is subject to right censoring, the extension of Beran's estimator [Akritas and Keilegom, 2003]

has been proposed. The analysis was conducted by using an alternative of W_i in equation (4.2) to take advantage of the information from the value of observed T_1 if the value of T_2 is censored. The modified $W_i^m(t)$ is shown as

$$W_i^m(t) = \frac{\delta_{1i} K\left(\frac{t-\widetilde{T}_{1i}}{h}\right)}{\sum_{j=1}^{n} \delta_{1j} K\left(\frac{t-\widetilde{T}_{1j}}{h}\right)}.$$

4.1.2 van der Laan's modified estimator from NPMLE

The idea of van der Laan [1994] is originally taken from Pruitt [1991]. It is used to overcome the problems with the NPMLE calculated from singly-censored observations in bivariate failure data. The strategy suggests that the uncensored component of a singly-censored observation, say $(T_1, T_2, \delta_1 = 1, \delta_2 = 0)$, is censored by a small interval $(T_1 - \lambda, T_1 + \lambda)$ for some $\lambda > 0$. The conditional survival function can be expressed as the equivalent of Kaplan-Meier estimator

$$S_{T_2|T_1=t_1}(t_2 \mid t_1) = \prod_{(0,t_2]} (1 - \Lambda(t_1 \mid ds)) \tag{4.3}$$

where

$$\Lambda(t_1 \mid ds) = \frac{P(T_2 \in ds \mid T_1 = t_1,\ \delta_1 = 1,\ \delta_2 = 1)}{P(T_2 \geq s \mid T_1 = t_1,\ \delta_1 = 1)}$$

Here $\Lambda(t_1 \mid ds)$ is the conditional hazard representing the conditional probability that the first event occurs at the coming moment given that it remains event free right now when the later event has occurred at time t_2. Since Λ can be estimated by the Nelson-Aalen estimator which takes the ratio of the number of deaths to the number at risk, the death and risk processes are denoted as

$$\begin{aligned} N(t_1, t_2) &= \frac{d}{dt_1} P(\widetilde{T}_1 \leq t_1, \widetilde{T}_2 \leq t_2,\ \delta_1 = 1,\ \delta_2 = 1) \\ Y(t_1, t_2) &= \frac{d}{dt_1} P(\widetilde{T}_1 \leq t_1, \widetilde{T}_2 > t_2,\ \delta_1 = 1) \end{aligned} \tag{4.4}$$

In order to estimate conditional densities, a nonparametric data smoothing technique is introduced for visualising the distribution of data. Then the death and risk processes from

the equation (4.4) are estimated using bivariate kernel smoothing function with a full bandwidth matrix to provide different weights for all uncensored and singly-censored observations around the line [van der Laan, 1997]. Therefore, the density estimators are denoted with $\widehat{N}$ and $\widehat{Y}$:

$$\begin{aligned}\widehat{N}(t_1,\ dt_2) &= \frac{1}{nh_1h_2}\sum_{j=1}^{n} K\left(\frac{t_1-\widetilde{T}_{1j}}{h_1}, \frac{dt_2-\widetilde{T}_{2j}}{h_2}\right) I(\widetilde{T}_{2j}\in dt_2,\ \delta_{1j}=1,\ \delta_{2j}=1)\\ \widehat{Y}(t_1,\ t_2) &= \int_{t_2}^{\lambda_2} \frac{1}{nh_1h_2}\sum_{j=1}^{n} K\left(\frac{t_1-\widetilde{T}_{1j}}{h_1}, \frac{s-\widetilde{T}_{2j}}{h_2}\right) I(\delta_{1j}=1)ds\end{aligned} \tag{4.5}$$

where K is the bivariate Gaussian kernel, $K(\mathbf{x}) = (2\pi)^{-1/2}\exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{x}\right)$. The values of bandwidth h_1 and h_2 are chosen from a full bandwidth matrix $\mathbf{H}$ which minimizes the smooth cross-validation function in the equation (2.8). The Kaplan-Meier estimator for $S_{T_2|T_1=t_1}$ under the condition of $t_1 \le t_2$ is then

$$\widehat{S}_{T_2|T_1=t_1}(t_2 \mid t_1) = \prod_{\widetilde{T}_{2j}\le t_2,\ \delta_{1j}=1}\left(1-\frac{\widehat{N}(t_1,\ \triangle\widetilde{T}_{2j})}{\widehat{Y}(t_1,\ \widetilde{T}_{2j})}\right). \tag{4.6}$$

4.2 QUANTILE RESIDUAL LIFETIME ESTIMATOR

Jeong et al. [2008] proposed a method for estimating the median residual lifetimes directly dealing with an estimating equation in the univariate settings. For a bivariate modeling of the residual life function, an estimating equation for quantile residual lifetime is defined by conditioning on the first failure at time t_1. Since the objective is to determine the residual lifetime of a later failed unit, it is reasonable to assume $t_2 \ge t_1$ from a dynamic point of view. The corresponding expression for only right censored T_1 is analogous. For $0 < \alpha < 1$,

the α-quantile residual life function at time t_2 is defined as

$$\begin{aligned}\gamma^B_{\alpha|t_1}(t_2) &= \inf\{x : F_{T_2|T_1=t_1}(x \mid t_1) \geq \alpha\} \\ &= \inf\{x : S_{T_2|T_1=t_1}(t_2 + x \mid t_1) < (1-\alpha) S_{T_2|T_1=t_1}(t_2 \mid t_1)\} \\ &= S^{-1}_{T_2|T_1=t_1}\{(1-\alpha) S_{T_2|T_1=t_1}(t_2) \mid t_1\} - t_2\end{aligned} \tag{4.7}$$

which implies the α-quantile residual lifetime of a unit survived beyond time t_2 when t_1 is fixed. On the basis of this conditional scenario, the proposed method is to estimate conditional survival probability given $T_1 = t_1$ and use it to infer $\gamma^B_{\alpha|t_1}$ through the estimating equation.

The equation (4.7) is equivalent to $P(T_2 - t_2 > \gamma^B_{\alpha|t_1} \mid T_1 = t_1,\ T_2 > t_2) = 1-\alpha$, showing that $P(T_1 = t_1, T_2 > t_2 + \gamma^B_{\alpha|t_1}) = (1-\alpha)P(T_1 = t_1, T_2 > t_2)$. So in terms of conditional survival function defined by equation (4.1), the following is obtained

$$S_{T_2|T_1=t_1}(t_2 + \gamma^B_{\alpha|t_1} \mid t_1) = (1-\alpha) S_{T_2|T_1=t_1}(t_2 \mid t_1). \tag{4.8}$$

The resulting estimator shown in the equation (4.6) can be plugged in the estimating equation $\widehat{u}(\gamma^B_{\alpha|t_1}) = 0$ for $\gamma^B_{\alpha|t_1}$, where

$$\widehat{u}(\gamma^B_{\alpha|t_1}) = \widehat{S}_{T_2|T_1=t_1}(t_2 + \gamma^B_{\alpha|t_1} \mid t_1) - (1-\alpha)\widehat{S}_{T_2|T_1=t_1}(t_2 \mid t_1) \tag{4.9}$$

and $\widehat{\gamma}^B_{\alpha|t_1}$ is the solution of $\widehat{u}(\gamma^B_{\alpha|t_1})$.

4.2.1 Nonparametric bootstrap method

The bootstrap introduced by Efron [1979] is a general methodology to deal with uncertainty of sampling distribution of estimators in almost any nonparametric estimation problem. The uncertainty associated with parameter estimates is usually summarized by approximated biases, standard deviations, and confidence intervals. It can answer questions which are far too complicated for traditional statistical analysis without the assumptions about distributions.

In the case of right censored data, the bootstrap approach answers several questions concerning the Kaplan–Meier survival curve and provides a new justification for Greenwood's

formula using large sample approximation [Efron, 1981]. Basically, the resampling scheme is the same as for the uncensored case, except that the data points become pairs with censoring indicators. The observed bivariate failure times along with their censoring indicators are denoted as $\widetilde{\mathbf{T}} = (\widetilde{T}_1, \widetilde{T}_2, \delta_1, \delta_2)$. Suppose we observe $\widetilde{\mathbf{T}}_i, i = 1, 2, \cdots, n$, where $\widetilde{\mathbf{T}}_i$ are independent and identically distributed according to some unknown probability distribution F. We are interested in calculating a standard error for our proposed estimator, $\widehat{\gamma}^B_{\alpha|t_1}$. Let $\sigma(F)$ denote the standard error of $\widehat{\gamma}^B_{\alpha|t_1}$ as a function of F. The bootstrap estimate of standard error is $\widehat{\sigma} = \sigma(\widehat{F})$, where $\widehat{F}$ is the empirical distribution function putting mass $1/n$ at each observation. The general bootstrap procedure is as follow:

1. A random sample of size n is drawn with replacement from the actual sample $\widetilde{\mathbf{T}} \sim \widehat{F}$ and repeat N times to obtain a large number of bootstrap datasets, denoted as $\widetilde{\mathbf{T}}^*(1), \widetilde{\mathbf{T}}^*(2), \cdots, \widetilde{\mathbf{T}}^*(N)$.
2. For each bootstrap dataset, calculate the statistic of interest, say $\widehat{\gamma}^{B^*}_{\alpha|t_1}(i), j = 1, 2, \cdots, N$.
3. Calculate the sample standard deviation of the $\widehat{\gamma}^{B^*}_{\alpha|t_1}(i)$, which is

$$\widehat{\sigma}_{Boot} = \sqrt{\frac{\sum_{i=1}^{N}\left\{\widehat{\gamma}^{B^*}_{\alpha|t_1}(i)\right\}^2 - \left\{\sum_{i=1}^{N}\widehat{\gamma}^{B^*}_{\alpha|t_1}(i)\right\}^2/N}{N-1}} \tag{4.10}$$

4.2.2 Simulation study

We conducted a simulation study to evaluate the performance of the proposed estimator of conditional α-quantile residual lifetime in a bivariate failure time setting. Bivariate right censored data were generated from Clayton bivariate exponential distribution. A pair of independent censoring variables C_1 and C_2 were uniform$(0, c)$ variates, where c determined the censoring proportion. Values of t_1 and t_2 were generated using the transformation from the model (3.4), where $t_1 = -\log(1 - u_1)$ and $t_2 = 2\theta \log\left[(1-a) + a(1-u_2)^{-1/(1+\theta)}\right], a = (1-u_1)^{-1/\theta}$.

A total of 1000 simulations from a bivariate Clayton distribution were carried out in samples of size 200. The censoring parameter c was chosen to give 20% censoring. The association parameter θ was chosen to be 0.25 or 0.5 to examine the performance of the proposed estimator for various pairs of failure times. We used Gaussian kernel and choose

a suitable bandwidth for each generated sample via smooth cross-validation function. For each simulation, 300 bootstrap samples were generated to obtain the bootstrap estimate of standard error. Under the assumption of $t_1 \leq t_2$, the conditional median residual lifetime, i.e., $\alpha = 0.5$ was evaluated from Clayton distribution. To receive the estimator of nonparametric conditional median residual lifetime in each parameter setting, one dimension root finding from the function uniroot() in R was used to search the interval from 0 to 3 for a root of the estimating equation.

In the following simulation results, the values of t_1 chose to be 0.44 and 1.03 corresponding to marginal survival probabilities of 0.8 and 0.6 and t_2 were at values of 0.44, 0.58, 1.03, 1.28 and 1.79 corresponding to marginal survival probabilities of 0.8, 0.75, 0.60, 0.50 and 0.40 from exponential with rate 0.5. Table 4.2 shows the results of nonparametric estimation of conditional median residual lifetime using the extension of Beran's estimator. The results from Van Der Laan's estimator are shown in Table 4.3. The estimated conditional median residual lifetimes from both estimators have similar performance. There is no obvious changing pattern in terms of 95% coverage. We observed that conditioning on the smaller t_1 would have less bias compared to larger t_1. However, we did not observe any reasonable coverage close to 95 among all parameter settings.

In addition, the mean of estimated standard errors were underestimated. Figure 4.1 illustrates the nonparametric estimation of conditional survival probabilities using van der Laan's approach. From eyeballing, the bias of estimated conditional survival probabilities in Figure 4.1a was small when $\theta = 1$ compared to other association values. In Figure 4.1b, the strength of association does not seem to distinguish the performance of estimated conditional survival overall, but the estimated conditional survival with strong association tended to approach the true distribution when t_2 was increasing. Furthermore, a smaller t_1 resulted in overestimated conditional survival in contrast with a larger t_1 for all θ.

Table 4.1: Simulation results of $\widehat{\gamma}^{B}_{0.5|t_1}$ using nearest neighbor estimation at $\theta = 0.25$. EST is the mean of estimates, SE is the mean of estimated standard errors, SD is standard deviation of the estimates and Cov95 is the 95% coverage.

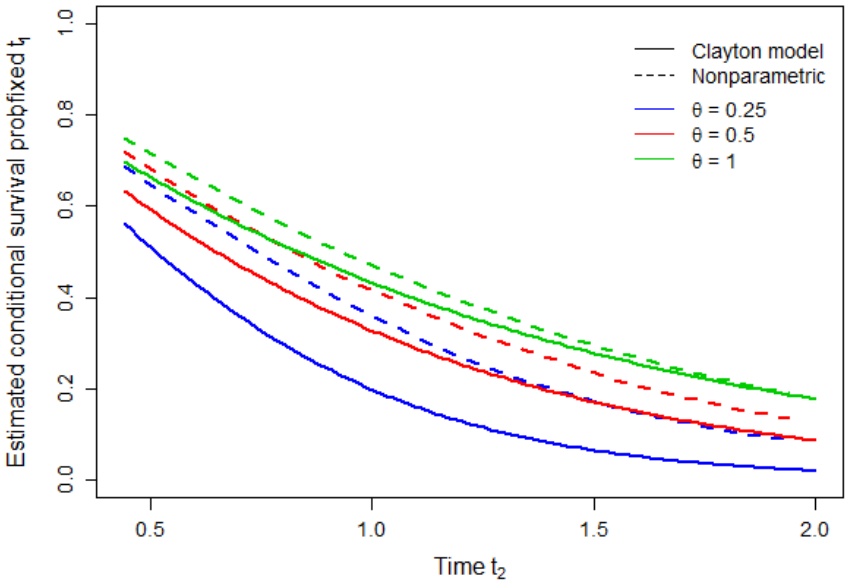
θ	n	t_1	t_2	True Value	EST	(SE)	SD	Cov95
0.25	200	0.44	0.44	0.388	0.440	(0.091)	0.086	91.2
			0.58	0.364	0.416	(0.097)	0.092	91.9
			1.28	0.300	0.362	(0.146)	0.153	88.8
		1.03	1.03	0.435	0.489	(0.122)	0.122	91.0
			1.28	0.379	0.449	(0.129)	0.134	89.3
			1.79	0.316	0.372	(0.149)	0.171	86.7
0.25	400	0.44	0.44	0.388	0.432	(0.064)	0.062	90.0
			0.58	0.364	0.406	(0.069)	0.066	90.8
			1.28	0.300	0.342	(0.115)	0.113	91.7
		1.03	1.03	0.435	0.476	(0.089)	0.087	92.3
			1.28	0.379	0.427	(0.099)	0.094	91.7
			1.79	0.316	0.364	(0.130)	0.129	91.0
0.25	800	0.44	0.44	0.388	0.421	(0.046)	0.045	89.7
			0.58	0.364	0.396	(0.049)	0.047	89.9
			1.28	0.300	0.323	(0.084)	0.079	93.5
		1.03	1.03	0.435	0.462	(0.064)	0.065	91.7
			1.28	0.379	0.410	(0.071)	0.071	91.2
			1.79	0.316	0.346	(0.101)	0.100	92.6

Table 4.2: Simulation results of $\widehat{\gamma}^{B}_{0.5|t_1}$ using nearest neighbor estimation at $\theta = 0.5$. EST is the mean of estimates, SE is the mean of estimated standard errors, SD is standard deviation of the estimates and Cov95 is the 95% coverage.

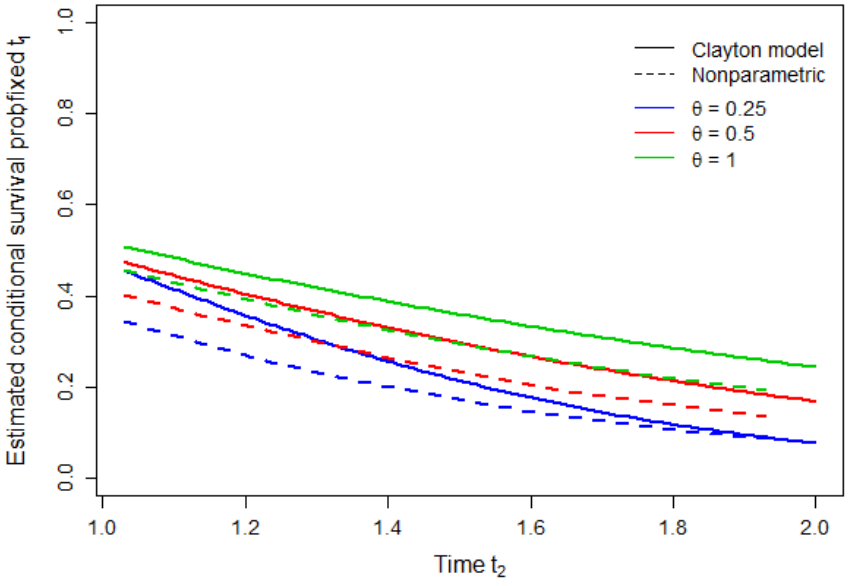
θ	n	t_1	t_2	True Value	EST	(SE)	SD	Cov95
0.5	200	0.44	0.44	0.586	0.612	(0.116)	0.118	91.6
			0.58	0.571	0.600	(0.123)	0.125	90.9
			1.28	0.517	0.546	(0.154)	0.172	86.9
		1.03	1.03	0.676	0.677	(0.136)	0.146	86.1
			1.28	0.632	0.649	(0.143)	0.163	83.1
			1.79	0.568	0.572	(0.161)	0.194	80.8
0.5	400	0.44	0.44	0.586	0.611	(0.088)	0.087	93.0
			0.58	0.571	0.596	(0.094)	0.093	92.9
			1.28	0.517	0.546	(0.129)	0.133	91.9
		1.03	1.03	0.676	0.684	(0.112)	0.116	89.4
			1.28	0.632	0.651	(0.123)	0.131	88.3
			1.79	0.568	0.594	(0.147)	0.163	85.5
0.5	800	0.44	0.44	0.586	0.602	(0.065)	0.063	94.7
			0.58	0.571	0.588	(0.069)	0.068	93.6
			1.28	0.517	0.539	(0.100)	0.103	91.7
		1.03	1.03	0.676	0.684	(0.090)	0.096	92.1
			1.28	0.632	0.642	(0.099)	0.103	91.1
			1.79	0.568	0.580	(0.122)	0.125	90.8

Table 4.3: Simulation results of estimation of $\gamma^B_{0.5|t_1}$ using van der Laan's estimator. EST is the mean of estimates, SE is the mean of estimated standard errors, SD is standard deviation of the estimates and Cov95 is the 95% coverage.

θ	n	t_1	t_2	True Value	EST	(SE)	SD	Cov95
0.25	200	0.44	0.44	0.388	0.449	(0.085)	0.087	89.9
			0.58	0.364	0.421	(0.087)	0.089	89.3
			1.28	0.300	0.365	(0.105)	0.104	90.7
		1.03	1.03	0.435	0.498	(0.113)	0.116	90.0
			1.28	0.379	0.460	(0.118)	0.120	89.4
			1.79	0.316	0.411	(0.130)	0.128	88.8
0.5	200	0.44	0.44	0.586	0.555	(0.106)	0.109	89.5
			0.58	0.571	0.534	(0.110)	0.110	90.2
			1.28	0.517	0.505	(0.131)	0.142	89.3
		1.03	1.03	0.676	0.603	(0.129)	0.136	84.1
			1.28	0.632	0.585	(0.136)	0.150	85.6
			1.79	0.568	0.559	(0.151)	0.159	86.3



(a) Fixed time $t_1 = 0.44$



(b) Fixed time $t_1 = 1.03$

Figure 4.1: Nonparametric estimator of the conditional survival probabilities in comparison of Clayton exponential model for different levels of dependence parameter, $\theta = 0.25,\ 0.50,\ 1$ with $n = 50$ in 500 iterations

4.3 DISCUSSION

The mechanism of nonparametric estimators for the conditional quantile residual life function is built upon redistributing the mass between an interval around the value of uncensored component of $\mathbf{T}$. Two estimators of conditional survival function are studied to estimate the conditional quantile residual lifetime. The bandwidth selection of kernel smoothing function is important relating to how well the estimated conditional survival function can approximate to the underlying distribution. The simulation results indicate that the conditional median residual lifetime inferred by both methods performs biased estimates when using a sample size of 200. The standard errors tend to estimate poorly when t_1 and t_2 get large. The reason might result from the fact of decreasing number of events from the tail. Therefore, we increase a sample size of 400 and 800 to see whether there is an improvement on the estimates. The 95% coverage probability shows a fairly amount of improvement for the estimated conditional median residual lifetime. Further investigation needs to take place in order to find a precise approximation.

5.0 CONCLUSION AND FUTURE WORK

The estimator of conditional quantile residual lifetime for semi-competing risks data has been constructed and studied in this dissertation. Given the occurrence time of the nonterminal event beyond a certain time t, the quantile residual lifetime of the terminal event can be evaluated at any time point. The proposed method considers the nature of the association between event times and provides a practical application for patients who have been followed for years after the initial treatment. In addition, the association between the conditional residual lifetime and selected covariates has been proposed in a regression model. The results indicate that our proposed estimator is nearly unbiased for the conditional median residual lifetime.

A nonparametric approach of conditional quantile residual lifetime has been compared to the proposed estimator. The censoring structure changes to the setting of parallel data, i.e. bivariate independent censoring. By assuming the condition that one of the paired units has failed at time t, we also evaluate the conditional quantile residual lifetime by using the estimating equation. Several existing methods of estimation of conditional survival function are considered. We adopt the methods proposed by Akritas and Keilegom [2003] and van der Laan [1994] to implement our approach. A small simulation study suggests that both estimators do not perform well according to 95% coverage probabilities. The estimator is sensitive to the choice of the bandwidth even though we have used a modification of cross-validation method in a bandwidth selection.

In summary, our proposed estimator performs better compared to a nonparametric approach. The conditioning of one of the paired units just failed at time t before another restricts to the occurrence time of the first failure. One has to consider the information from singly-censored and both uncensored observations to make the resulting estimator consistent.

A bootstrap method is useful because it provides the estimation of the standard deviation of the conditional quantile residual lifetime. Reasons for causing a biased nonparametric estimator might be due to an insufficient sample size or an improper bandwidth. Further investigation is needed.

This work only considered gamma frailty model to derive our proposed estimator in semi-competing risks data. It would be worthwhile to consider other joint survival models. Another issue that can be addressed in the future is to investigate other complex censoring structures in bivariate failure time data such as the analysis of successive event times. It would be reasonable to perform our analysis under various specific censoring schemes.

APPENDIX

ANALYSIS OF VENTILATING TUBES IN EARS

Otitis media (OM), an inflammation of the middle ear, is frequently diagnosed in children. The symptoms of OM include pain in the ear, fever, and temporary hearing loss and general signs such of loss of appetite and irritability. Several subtypes of otitis media have been distinguished but the most common form is otitis media with effusion (OME). It is defined when there is a collection of thick or sticky fluid behind the eardrum in the middle ear without signs of ear infection. In typical, children with OME do not suffer pain or fever but may experience the loss of hearing to affect their behavioural and language development. A number of medical interventions have been suggested for the treatment of OME but failed with limited effects. Therefore, physicians seek with favor of a surgical intervention, that is, ventilating tubes are inserted in the eardrums to let fluid trapped behind the eardrum drain.

Between February 1987 and January 1990, a total of 78 eligible children age 6 months to 8 years with chronic OME were enrolled after receiving therapeutic myringotomy for tympanostomy tube placement [Le and Lindgren, 1996]. Children were randomly assigned to either receive 2-week trials of prednisone and sulfamethoprim treatment after surgery (n=40) or serve as controls without additional treatment (n=38). The duration of ventilating tubes were recorded in order to assess whether the medical treatment could prolong the life of the tubes. The cession of tube functioning or tube extrusion would be treated as the primary endpoint of the study. For each child duration time of vent tube in each ear along with corresponding censoring time are recorded. The data are shown in Figure A1.

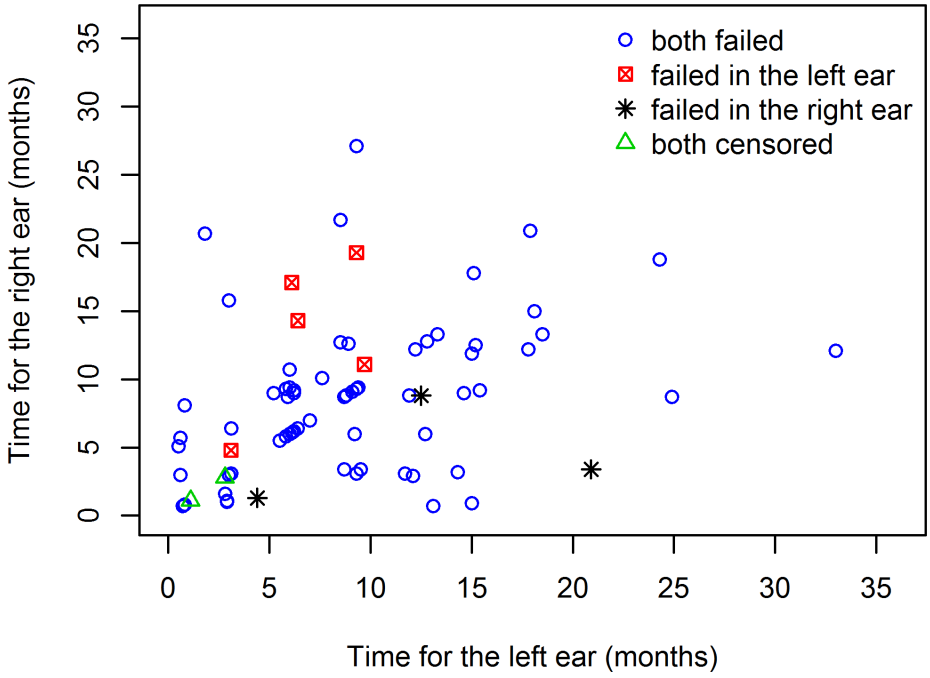


Figure A1: Duration times (months) of ventilating tubes in both ears for 78 children with otitis media effusion

BIBLIOGRAPHY

Odd O. Aalen. *Statistical inference for a family of counting processes*. Institute of Mathematical Statistics, University of Copenhagen, 1976.

Odd O. Aalen. Nonparametric inference for a family of counting processes. *The Annals of Statistics*, pages 701–726, 1978.

Belkacem Abdous and Alexandre Berred. Mean residual life estimation. *Journal of Statistical Planning and Inference*, 132(1):3–19, 2005.

Michael G. Akritas. Nearest neighbor estimation of a bivariate distribution under random censoring. *The Annals of Statistics*, pages 1299–1327, 1994.

Michael G. Akritas and Ingrid Van Keilegom. Estimation of bivariate and marginal distributions with censored data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):457–471, 2003.

Khursheed Alam and KB Kulasekera. Estimation of the quantile function of residual life time distribution. *Journal of statistical planning and inference*, 37(3):327–337, 1993.

Barry C. Arnold and Hassan Zahedi. On multivariate mean remaining life functions. *Journal of multivariate analysis*, 25(1):1–9, 1988.

Ishwar V. Basawa and Hira L. Koul. Large-sample statistics based on quadratic dispersion. *International Statistical Review/Revue Internationale de Statistique*, pages 199–219, 1988.

Rudolf Beran. Nonparametric regression with randomly censored survival data. Technical report, Technical Report, Univ. California, Berkeley, 1981.

Yogendra P. Chaubey and Pranab K. Sen. On smooth estimation of mean residual life. *Journal of Statistical Planning and Inference*, 75(2):223–236, 1999.

Chang-Jo F. Chung. Confidence bands for percentile residual lifetime under random censorship model. *Journal of Multivariate Analysis*, 29(1):94–126, 1989.

David G. Clayton. 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, 1978.

David R. Cox. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society*, 34:187–220, 1972.

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

Miklós Csörgő and Sándor Csörgő. Estimation of percentile residual life. *Operations Research*, 35(4):598–606, 1987.

Dorota M. Dabrowska. Kaplan-meier estimate on the plane. *The Annals of Statistics*, pages 1475–1489, 1988.

Roger Day, John Bryant, and Myrto Lefkopoulou. Adaptation of bivariate frailty models for prediction, with application to biological markers as prognostic indicators. *Biometrika*, 84 (1):45–56, 1997.

Tarn Duong and Martin L. Hazelton. Cross-validation bandwidth matrices for multivariate kernel density estimation. *Scandinavian Journal of Statistics*, 32(3):485–506, 2005.

Bradley Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, pages 1–26, 1979.

Bradley Efron. Censored data and the bootstrap. *Journal of the American Statistical Association*, 76(374):312–319, 1981.

Jason P. Fine, Hongyu Jiang, and Rick Chappell. On semi-competing risks data. *Biometrika*, 88(4):907–919, 2001.

Bernard Fisher, Joseph Costantino, Carol Redmond, Roger Poisson, David Bowman, Jean Couture, Nikolay V. Dimitrov, Norman Wolmark, D. Lawrence Wickerham, Edwin R. Fisher, et al. A randomized clinical trial evaluating tamoxifen in the treatment of patients with node-negative breast cancer who have estrogen-receptor–positive tumors. *New England Journal of Medicine*, 320(8):479–484, 1989.

Bernard Fisher, James Dignam, John Bryant, Arthur DeCillis, D. Lawrence Wickerham, Norman Wolmark, Joseph Costantino, Carol Redmond, Edwin R. Fisher, David M. Bowman, et al. Five versus more than five years of tamoxifen therapy for breast cancer patients with negative lymph nodes and estrogen receptor-positive tumors. *Journal of the National Cancer Institute*, 88(21):1529–1542, 1996.

Bernard Fisher, Jong-Hyeon Jeong, John Bryant, Stewart Anderson, James Dignam, Edwin R. Fisher, and Norman Wolmark. Treatment of lymph-node-negative, oestrogen-receptor-positive breast cancer: long-term findings from national surgical adjuvant breast and bowel project randomised clinical trials. *The Lancet*, 364(9437):858–868, 2004.

Thomas R. Fleming and David P. Harrington. *Counting processes and survival analysis*, volume 169. John Wiley & Sons, 2011.

Alan E. Gelfand and Athanasios Kottas. Bayesian semiparametric regression for median residual life. *Scandinavian Journal of Statistics*, 30(4):651–665, 2003.

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

Peter Hall, JS Marron, and Byeong U. Park. Smoothed cross-validation. *Probability Theory and Related Fields*, 92(1):1–20, 1992.

WJ Hall and Jon A. Wellner. Mean residual life. *Statistics and Related Topics*, 169:184, 1981.

Philip Hougaard and Philip Hougaard. *Analysis of multivariate survival data*, volume 564. Springer New York, 2000.

Dong Myung Jeong, Jae Kee Song, and Joong Kweon Sohn. Nonparametric estimation of bivariate mean residual life function under univariate censoring. *JKSS (Journal of the Korean Statistical Society)*, 25(1):133–144, 1996.

Jong-Hyeon Jeong, Sin-Ho Jung, and Joseph P. Costantino. Nonparametric inference on median residual life function. *Biometrics*, 64(1):157–163, 2008.

Hongyu Jiang, Rick Chappell, and Jason P Fine. Estimating the distribution of nonterminal event time in the presence of mortality or informative dropout. *Controlled clinical trials*, 24(2):135–146, 2003.

Chris Jones, JS Marron, and Simon J. Sheather. Progress in data-based bandwidth selection for kernel density estimation. *Computational Statistics*, (11):337–381, 1996.

Sin-Ho Jung, Jong-Hyeon Jeong, and Hanna Bandos. Regression on quantile residual life. *Biometrics*, 65(4):1203–1212, 2009.

Edward L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

Niels Keiding. Age-specific incidence and prevalence: a statistical perspective. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 371–412, 1991.

John P. Klein and Melvin L. Moeschberger. *Survival analysis: Techniques for censored and truncated data: Statistics for Biology and Health*. Springer, 2003.

Hemangi V. Kulkarni and RN Rattihalli. Nonparametric estimation of a bivariate mean residual life function. *Journal of the American Statistical Association*, 97(459):907–917, 2002.

Parthasarthi Lahiri and Dong Ho Park. Nonparametric bayes and empirical bayes estimators of mean residual life at age t. *Journal of statistical planning and inference*, 29(1):125–136, 1992.

Lajmi Lakhal-Chaieb, Belkacem Abdous, and Thierry Duchesne. Nonparametric estimation of the conditional survival function for bivariate failure times. *Canadian Journal of Statistics*, 41(3):439–452, 2013.

Chap T. Le and Bruce R. Lindgren. Duration of ventilating tubes: a test for comparing two clustered samples of censored data. *Biometrics*, pages 328–334, 1996.

Ruosha Li and Limin Peng. Quantile regression for left-truncated semicompeting risks data. *Biometrics*, 67(3):701–710, 2011.

Rosa E. Lillo. On the median residual lifetime and its aging properties: A characterization theorem and applications. *Naval Research Logistics (NRL)*, 52(4):370–380, 2005.

Alexander C. McLain and Sujit K. Ghosh. Nonparametric estimation of the conditional mean residual life function with censored data. *Lifetime data analysis*, 17(4):514–532, 2011.

KR Muralidharan Nair and N Unnikrishnan Nair. Bivariate mean residual life. *Reliability, IEEE Transactions on*, 38(3):362–364, 1989.

Wayne Nelson. Theory and applications of hazard plotting for censored failure data. *Technometrics*, 14(4):945–966, 1972.

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

William J. Padget and LA Thombs. A smooth nonparametric quantile estimator from right-censored data. *Statistics & probability letters*, 7(2):113–121, 1988.

William J. Padgett. A kernel-type estimator of a quantile function from right-censored data. *Journal of the American Statistical Association*, 81(393):215–222, 1986.

Limin Peng and Jason P. Fine. Regression modeling of semicompeting risks data. *Biometrics*, 63(1):96–108, 2007.

Ross L. Prentice and Jianwen Cai. Covariance and survivor function estimation using censored multivariate failure time data. *Biometrika*, 79(3):495–512, 1992.

RC Pruitt. Strong consistency of self-consistent estimators: general theory and an application to bivariate survival analysis. Technical report, Technical report, 1991.

Stephan R. Sain, Keith A. Baggerly, and David W. Scott. Cross-validation of multivariate densities. *Journal of the American Statistical Association*, 89(427):807–817, 1994.

David C. Schmittlein and Donald G. Morrison. The median residual lifetime: A characterization theorem and an application. *Operations Research*, 29(2):392–399, 1981.

Moshe Shaked and J George Shanthikumar. Dynamic multivariate mean residual life functions. *Journal of applied probability*, pages 613–629, 1991.

Jeffrey S. Simonoff. *Smoothing methods in statistics*. Springer, 1996.

Mark J. van der Laan. Modified em-estimator of the bivariate survival function. *Mathematical Methods of Statistics*, 3:213–43, 1994.

Mark J. van der Laan. Nonparametric estimators of the bivariate survival function under random censoring. *Statistica Neerlandica*, 51(2):178–200, 1997.

Matt P. Wand and M Chris Jones. Comparison of smoothing parameterizations in bivariate kernel density estimation. *Journal of the American Statistical Association*, 88(422):520–528, 1993.

Matt P. Wand and M Chris Jones. Multivariate plug-in bandwidth selection. *Computational Statistics*, 9(2):97–116, 1994.

Lee-Jen Wei, Zhiliang Ying, and Danyu Lin. Linear regression analysis of censored survival data based on rank tests. *Biometrika*, 77(4):845–851, 1990.

Jinfeng Xu, John D. Kalbfleisch, and Beechoo Tai. Statistical analysis of illness–death processes and semicompeting risks data. *Biometrics*, 66(3):716–725, 2010.

Grace Yang. Life expectancy under random censorship. *Stochastic Processes and their Applications*, 6(1):33–39, 1977.

Zhiliang Ying, Sin-Ho Jung, and Lee-Jen Wei. Survival analysis with median regression models. *Journal of the American Statistical Association*, 90(429):178–184, 1995.