ABSTRACT

Title of dissertation: REGRESSION ANALYSIS
OF RECURRENT EVENTS
WITH MEASUREMENT ERRORS

Yixin Ren
Doctor of Philosophy, 2019

Dissertation directed by: Professor Xin He
Department of Epidemiology and Biostatistics
Professor Paul Smith
Department of Mathematics

Recurrent event data and panel count data are often encountered in longitudinal follow-up studies. The main difference between the two types of data is the observation process. Continuous observations will result in recurrent event data; and discrete observations will lead to panel count data. In statistical literature, regression analysis of the two types of data have been well studied; and a typical assumption of those studies is that all covariates are accurately recorded. However, in many applications, it is common to have measurement errors in some of the covariates. For example, in a clinical trial, a medical index might have been measured multiple times. Then dealing with the differences among those measurements is an essential topic for statisticians.

For recurrent event data, we present a class of semiparametric regression models that allow correlations between censoring time and recurrent event process via frailty. An estimating equation based approach is developed to account for the presence of measurement errors in some of the covariates. Both large and finite sample properties of the proposed estimators are established. An example from the study of gamma interferon in chronic granulomatous disease is provided.

For panel count data, we consider two situations in which the observation process is independent or dependent of covariates. Estimating equations are developed for the estimation of the regression parameters for both cases. Simulation studies indicate that the proposed inference procedures perform well for practical situations. An example of bladder cancer study is used to demonstrate the value of the proposed method.

Regression Analysis of Recurrent Events with Measurement Errors

by

Yixin Ren

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2019

Advisory Committee:
Professor Xin He, Co-Chair
Professor Paul Smith, Co-Chair
Professor Mei-Ling Ting Lee
Professor Eric Slud
Professor Yan Li, Dean's Representative

Acknowledgments

I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost I'd like to thank my advisor, Professor Xin He for giving me an invaluable opportunity to work on challenging and extremely interesting projects over the past four years. He has always made himself available for help and advice and without his extraordinary theoretical ideas and computational expertise, this thesis would have been a distant dream. It has been a pleasure to work with and learn from such an extraordinary individual.

I would also like to thank my co-advisor, Professor Paul Smith. There has never been an occasion when I've knocked on his door and he hasn't given me time. Thanks are due to Professor Eric Slud, Professor Mei-Ling Ting Lee and Professor Yan Li for agreeing to serve on my thesis committee and for sparing their invaluable time reviewing the manuscript.

I would also like to acknowledge help and support from some of the staff members. M. Cristina Garcia's help is highly appreciated.

I owe my deepest thanks to my family - my mother and father who have always stood by me and guided me through my career, and have pulled me through against impossible odds at times. Words cannot express the gratitude I owe them. Same to my brother. I would also like to thank my aunt and uncle as well as my cousins, who have supported me over the past ten years in the United States.

I would like to acknowledge financial support from the Department of Math-

ematics and the valuable experience on teaching and conveying mathematical concepts to students.

It is impossible to remember all, and I apologize to those I've inadvertently left out.

Lastly, thank you all!

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

In many study fields, researchers are interested in analyzing the patterns of the occurrences or the event history of certain events. In general, there are two types of event history studies. One is the type of events that can occur only once, and the resulting data of this kind of events are called survival or failure time data. The other type is that the event can occur repeatedly, which are usually referred to as recurrent events. For the first type of events, it might be the case that the event itself can occur only once, for example, death. Or, it might be the case that the event can actually occur multiple times, but the focus of the study is on the first time to the occurrence, for example, first marriage.

With respect to recurrent events, they can also be classified into two types. The first type of recurrent events often arise when the study subjects are being monitored continuously and all occurrence times of the events are observed. This type of data are usually referred to as recurrent event data. For instance, the occurrences of mammary tumors were recorded for 48 female rats (Gail et al., 1980). The other type refers to the situation that the study subjects are only observed at discrete time points, and therefore, only the number of occurrences of events between two consecutive observation times are available. These data have been referred to as panel count data. A typical example is the reliability study of 30 nuclear plants on the loss of feedwater flow (Sun and Kalbfleisch, 1995). Detailed examples of these three types of data are given below.

1.1 Example of Survival Data

One example of the type of events that occur only once is the survival data on remission times of acute leukemia patients discussed by Freireich et al. (1963) and Gehan (1965). Table 1.1 shows the remission times in weeks for 42 patients in two treatment groups. One is the placebo group; and the other is the drug 6-mercaptopurine (6-MP). The patients enrolled into the study at different times and the period of study was one year. The goal of this study was to examine if the patients with drug 6-MP had significantly longer remission times than those in the placebo group.

Table 1.1: Remission times in weeks for acute leukemia patients (Sun and Zhao, 2013)

Treatment	Survival times in weeks
6-MP	6, 6, 6, 6*, 7, 9*, 10, 10*, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25* 32*, 32*, 34*, 35*
Placebo	1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23

The starred numbers indicate censoring times, which means patients were still alive at the end of the study trial, or patients dropped out of the study for some reason. Therefore, the actual remission times for those patients were known only to be greater than the numbers in the table. For the other patients, their remission times were recorded exactly. This common situation in survival data is referred to as right-censored survival data.

Another common format of survival data is interval-censor, which arises, for example, when study subjects are observed at discrete time points. One example of interval-censored survival data is from a retrospective study on early breast cancer patients. In the study, 94 patients, who were given either radiation therapy alone (RT, 46) or radiation therapy plus adjuvant chemotherapy (RCT, 48), were scheduled to visit the clinic every 4 to 6 months, but the actual visit times vary from patient to patient, and times between visits were also different from each other. The goal of this study is to compare the two treatments with respect to their cosmetic effects (Sun, 2006).

1.2 Examples of Recurrent Event Data

The example of recurrent event data given in the previous section, mammary tumors in a carcinogenicity study presented by Gail et al. (1980), contained the times to mammary tumors for 48 female rats (Cook and Lawless, 2007). After being exposed to a carcinogen and further conditioned for 60 days, those rats were randomly assigned to either a treatment or control group. During a follow-up period of 122 days, they were examined every day for the development of new tumors. The observation process was assumed to be continuous.

Figure 1.1 gives the event plots for each rat after randomization, where dots

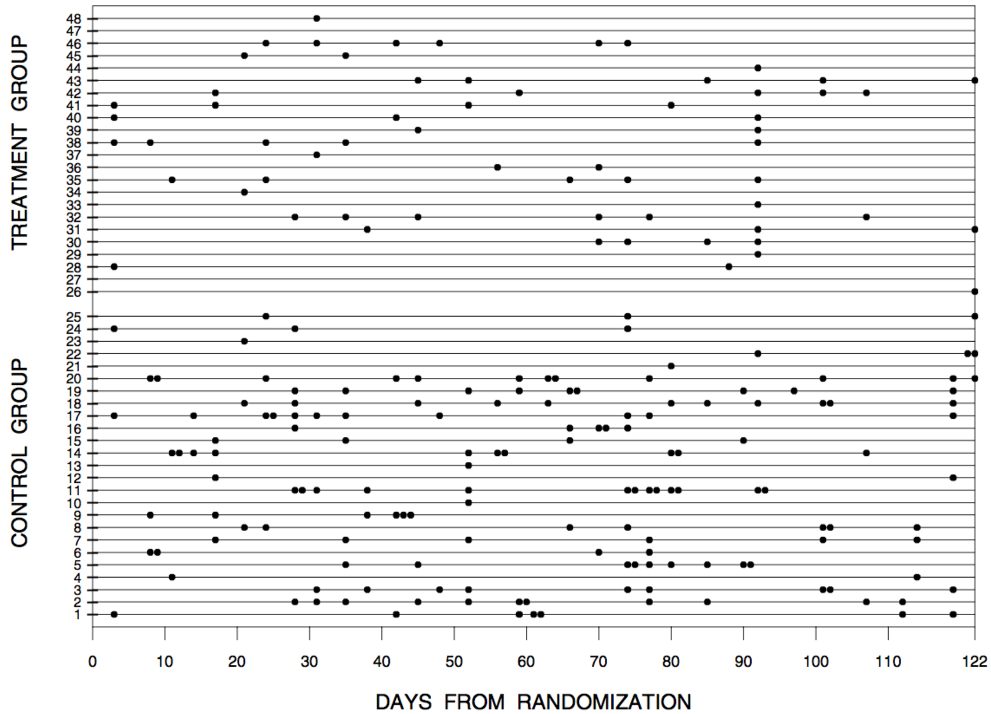

Figure 1.1: Event Plots for Tumor Occurrences in 48 Rats (Cook and Lawless, 2007)

are placed on the day when an event occurred. Dots were slightly separated to indicate that there were multiple events for a given rat on a given day. The event plot also shows the follow-up time for each rat; in this example, all the 48 rats had the same follow-up period, which was 122 days. Table 1.2 gives the times to tumors of the two groups from the beginning of the exposure. The numbers in parentheses indicate the numbers of tumors detected.

Another example of recurrent event data is the one discussed in Chapter 3, which is taken from the study of gamma interferon in chronic granulomatous disease (CGD) presented by Fleming and Harrington (1991). CGD is a group of inherited rare disorders of immune function characterized by recurrent pyogenic infections. The study was a double-blinded clinical trial in which 128 patients were randomly assigned to either placebo group or the gamma interferon treatment group. All patients were followed to record the recurrent infections.

Table 1.2: Times to tumor for 48 female rats (# in parentheses are # of tumors) (Cook and Lawless, 2007)

Treatment group		Control group	
ID	Times to tumor	ID	Times to tumor (in days)
1	182	1	63, 102, 119, 161(2), 172, 179
2		2	88, 91, 95, 105, 112, 119(2), 137, 145, 167, 172
3	63, 68	3	91, 98, 108, 112, 134, 137, 161(2), 179
4	152	4	71, 174
5	130, 134, 145, 152	5	95, 105, 134(2), 137, 140, 145, 150(2)
6	98, 152, 182	6	68(2), 130, 137
7	88, 95, 105, 130, 137, 167	7	77, 95, 112, 137, 161, 174
8	152	8	81, 84, 126, 134, 161(2), 174
9	81	9	68, 77, 98, 102(3)
10	71, 84, 126, 134, 152	10	112
11	116, 130	11	88(2), 91, 98, 112, 134(2), 137(2), 140(2), 152(2)
12	91	12	77, 179
13	63, 68, 84, 95, 152	13	112
14	105, 152	14	71(2), 74, 77, 112, 116(2), 140(2), 167
15	63, 102, 152	15	77, 95, 126, 150
16	63, 77, 112, 140	16	88, 126, 130(2), 134
17	77, 119, 152, 161, 167	17	63, 74, 84(2), 88, 91, 95, 108, 134, 137, 179
18	105, 112, 145, 161, 182	18	81, 88, 105, 116, 123, 140, 145, 152, 161(2), 179
19	152	19	88, 95, 112, 119, 126(2), 150, 157, 179
20	81, 95	20	68(2), 84, 102, 105, 119, 123(2), 137, 161, 179, 182
21	84, 91, 102, 108, 130, 134	21	140
22		22	152, 182(2)
23	91	23	81
		24	63, 88, 134
		25	84, 134, 182

1.3 Examples of Panel Count Data

For panel count data, an example can be taken from a bladder cancer study conducted by the Veterans Administration Cooperative Urological Research Group (Byar et al., 1977; Byar, 1980). The study focused on patients who had superficial bladder tumors when they enrolled in the study. After their initial tumors were removed, they were randomly assigned into three treatment groups, including placebo, thiotepa, and pyridoxine. During the study period, many patients had multiple recurrences of tumors. Additionally, for each patient, the obtained data also included information on two baseline covariates: the size of the largest initial tumor and the number of the initial tumors (Sun and Zhao, 2013).

Table 1.3: Observed number of bladder tumors along with the numbers of initial tumors and the size of the largest initial tumor from a bladder cancer study (Sun and Zhao, 2013).

Patient ID	Size	Months (0 … 10 … 20 … 30)
		Placebo group
1	3	1 0 . . . .
2	1	2 0 . . 0 . . . .
3	1	1 . . . . . . 0 . . . .
4	1	5 . . 0 . . . . . 0 0 . . . .
5	1	4 0 . . 0 . 1 0 . . 0 . . . .
6	1	1 . . 0 . . . . . . 0 . . . 0 . . . .
7	1	1 . 0 . . . . 0 . 2 . . . 3 . 0 . . . .
8	1	1 . . 0 . . . . 0 . . . 0 . . . .
9	3	1 . . . . 2 . . . . 0 . 0 . . . . . 0 . . . .
10	3	1 . . 0 . . . 0 . . 6 . . . . 3 . . . 0 . . . 0 . . . .

The observed data of this study contained a sequence of clinical visit times and the number of reccurrent tumors that occurred between two consecutive visits. Similarly as the initial tumors, the recurrent tumors were also removed at patients' clinical visits. Table 1.3 gives part of the reproduced data from Andrews and Herzberg (1985) and Sun and Wei (2000). The full data included 118 patients, in which 47 were assigned to the placebo group; 38 were assigned to the thiotepa group; and 33 were in the pyridoxine group. Moreover, observation times were recorded in months, with the longest observation time being 53 months. This example will be discussed in Chapter 4.

Table 1.4 presents another example of panel count data. It is a reliability study on the feedwater flow loss of 30 nuclear plants. Gaver and O'Muircheartaigh (1987) and Sun and Kalbfleisch (1995) reproduced the data, which gave both the observation time and the corresponding observed number of losses of feedwater flow for each nuclear plant (Sun and Zhao, 2013).

Table 1.4: Observed numbers of loss of feedwater flow from 30 nuclear plants (Sun and Zhao, 2013)

Observation time t_i (in years) and observed number n_i											
Plant	t_i	n_i	Plant	t_i	n_i	Plant	t_i	n_i	Plant	t_i	n_i
1	15	4	9	4	13	17	2	11	25	1	1
2	12	40	10	3	4	18	2	1	26	3	10
3	8	0	11	4	27	19	2	0	27	2	5
4	8	10	12	4	14	20	1	3	28	4	16
5	6	14	13	4	10	21	1	5	29	3	14
6	5	31	14	2	7	22	1	6	30	11	58
7	5	2	15	3	4	23	5	35			
8	4	4	16	3	3	24	3	12			

As shown in the above examples, there are two key differences between recurrent event data and panel count data. The first main difference is the amount of relevant information available; and the second is the observation process. Due to the incomplete nature of panel count data, the exact event occurrence time is unknown for each study subject; whereas, for recurrent event data, one can observe the event ooccurrence time immediately. The observation process of recurrent event data lasts continuously. However, the observation process of panel count data only includes a sequence of discrete observation times.

1.4 Regression Analysis of Recurrent Events

For parametric regression analysis of recurrent events, likelihood based method is the most common approach. Various parametric counting process regression models have been studied by many researchers, with the establishment of the consistency and asympotic normality of the corresponding maximum likelihood estimators. A detailed discussion can be found in Lawless (1987).

For semiparametric regression analysis of recurrent events, it is common to model the intensity function. A typical example is a counting process formulation of the Cox intensity model proposed by Anderson and Gill (1982), which can be considered as a generalization of the Cox proportional hazards model (Cox, 1972). Another approach is to model the mean or rate function (Lawless and Nadeau, 1995), which is also considered as a generalization of the Cox model (Cox, 1972), but does not require proportional hazards and is more straightforward for interpretation. This approach is more versatile as it allows arbitrary dependence among events and different sizes of jumps. Thus, it can be extended to analyze panel count data as well. Detailed reviews for the above methods of regression analysis of recurrent events are provided in Chapter 2.

1.5 Our Contribution

Regression analysis of recurrent events with completely and accurately observed covariates has been well studied in the statistical literature. However, if there exists some measurement errors in part of the covariates, simply ignoring the error terms is not appropriate. Intensive work has been done for survival and longitudinal data analysis with measurement errors. However, limited research has been conducted for recurrent events.

The goal of our work is to show that by extending the existing estimating equation based methods for both types of recurrent events, the estimators of the regression parameters can still be consistent and asymptotically normal even with presence of relatively large measurement errors in some of the covariates. The simulation results also indicate that our proposed methods are reliable and robust.

Chapter 2: Literature Review

2.1 Counting Process

Modeling recurrent events can be approached in many different ways, which has been discussed in the literature on stochastic processes, and more specifically, point processes. The concepts of intensity functions and counting processes are especially useful for both modeling and statistical analysis (Sun and Zhao, 2013). Andersen and Gill (1982) proposed the Cox intensity model for counting processes by developing a partial likelihood estimation procedure for the regression parameters, and established the large sample theory for the resulting estimators. In the following sections, we will introduce the notations, some commonly used models, and three different approaches for the regression analysis of recurrent events.

2.1.1 Notations

For a single recurrent event process, let $0 \leq T_1 < T_2 < \cdots$ denote the event times, where T_i is the ith event occurrence time. Then the associated counting process $\{N(t), 0 \leq t\}$ gives the cumulative number of events. In other words, $N(t) = \sum_{i=0}^{\infty} I(T_i \leq t)$ is the number of events occurring over the time interval $[0, t]$. In addition, let t^- and t^+ denote the limits from the left and right, respectively. Then the counting processes are right-continuous, which means $N(t) = N(t^+)$. Figure 2.1 shows a right-continuous counting process. Let $\Delta N(t) = N(t + \Delta t) - N(t^-)$ denote the number of events occurring during the time interval $[t, t + \Delta t)$, and let $\mathcal{H}(t) = \{N(s) : 0 \leq s < t\}$ denote the history of the process at time t (a σ-algebra of events). Since events are occurring in continuous time, we assume that no two events can occur simultaneously for mathematical convenience. Then, for the recurrent event data, since the observation process is continuous, we have an intensity function, which gives the instantaneous probability of an event occuring at t, conditioned on the process history. The intensity function is defined formally as

$$\lambda(t|\mathcal{H}(t)) = \lim_{\Delta t \downarrow 0} \frac{Pr\{N(t + \Delta t) - N(t^-) = 1|\mathcal{H}(t)\}}{\Delta t} \tag{2.1}$$

Given $\lambda(t)$, one can obtain the cumulative intensity process $\Lambda(t) = \int_0^t \lambda(s)ds$ and can then directly model $\Lambda(t)$. Let $Z(t)$ denote a vector of covariate processes, $\mathcal{F}_t$ denote the history generated by $\{N(s), Z(s) : 0 \leq s < t\}$ and $\lambda_Z(t)$ the intensity process of $N(t)$ associated with $\mathcal{F}_t$. That is,

$$E[dN(t)|\mathcal{F}_t] = \lambda(t)dt,$$

where $dN(t)$ is the increment $N(t + dt) - N(t-)$ of $N(t)$ over the small interval $[t, t + dt)$.

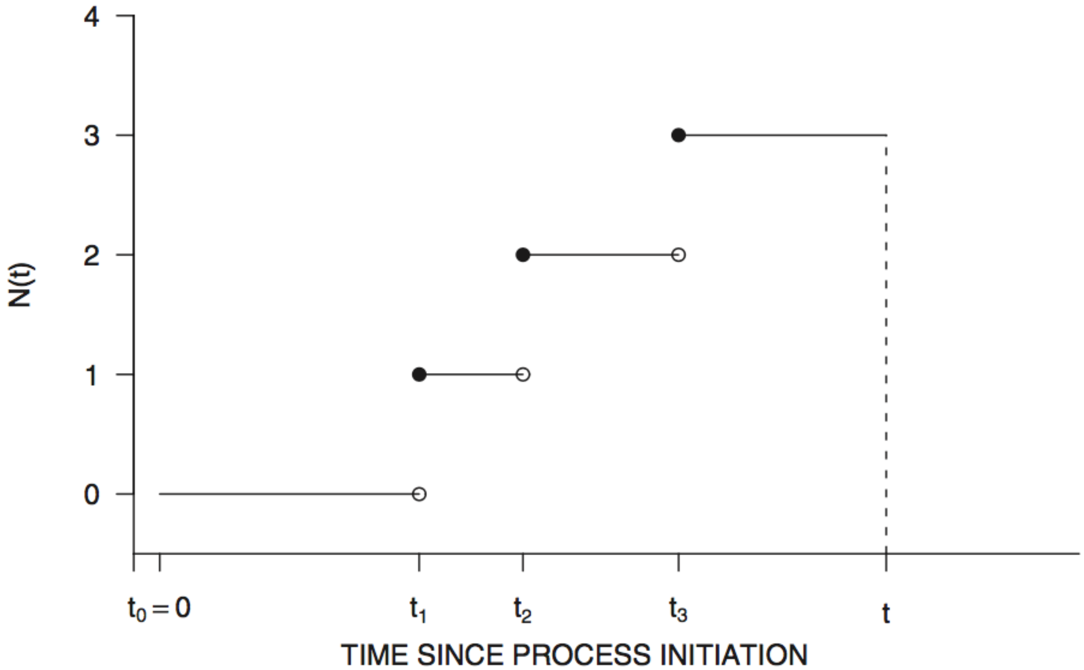

Figure 2.1: Counting process representation on recurrent events.

Due to the incomplete nature of the observed information, researchers usually directly model the mean or rate function of the counting process of panel count data. Commonly used models for analyzing recurrent event data and panel count data are discussed in the next section.

2.1.2 Commonly Used Models

For analyzing recurrent event data, a commonly used model for $\lambda_Z(t)$ given the covariates $Z(t)$, is the Cox intensity model stated by Andersen and Gill (1982):

$$\lambda_Z(t) = \lambda_0(t) \exp\{(\beta' Z(t))\}, \tag{2.2}$$

where $\lambda_0(t)$ is a completely unspecified continous baseline intensity function and β is a vector of regression parameters. However, the Cox intensity model 2.2 may not be very practical because it is too restrictive (Lin et al., 2000). Thus, modeling the mean or rate function of $N(t)$ could be an alternative, since dealing with the mean or rate functions requires fewer assumptions than those needed under the Cox intensity function. The rate function $r(t)$ of $N(t)$ is defined by

$$E\{dN(t)\} = r(t)dt. \tag{2.3}$$

Then, the mean function $\mu(t)$ can be calculated as $\mu(t) = \int_0^\infty r(s)ds$. Given $Z(t)$, the proportional rate model has been a popular choice to model the rate function, which is defined by

$$r_Z(t)dt = E\{dN(t)|Z(t)\} = r_0(t)\exp\{(\beta' Z(t))\}dt, \tag{2.4}$$

where $r_0(t)$ is a completely unknown baseline rate function and β is a vector of regression parameters as defined above. If Z is independent of time, then the mean function would become

$$\mu_Z(t) = E\{N(t)|Z\} = \mu_0(t)\exp\{(\beta' Z)\}, \tag{2.5}$$

where $\mu_0(t) = \int_0^t r_0(s)ds$. This is usually referred to as the proportional mean model (Cook and Lawless, 2007). This model can also be used for analyzing panel count data, since it can model point processes with positive jumps of arbitrary sizes. For the remainder of the dissertation, we only consider covariates that are independent of time.

2.2 Regression Analysis of Recurrent Events

2.2.1 Regression Analysis under the Cox Intensity Model

Let us consider a study of a single type of recurrent events observed on n independent subjects. Let $N_i(t)$ denote the counting process representing the number of occurrences of the event over the time interval $[0, t]$ for subject $i, i = 1, \ldots, n$. Assume that the observation process for each subject i is continuous up to $\min(C_i, \tau)$, where C_i is the follow-up time for subject i and τ is the study ending time. Now define $Y_i(t) = I(t \leq \min(C_i, \tau))$, a left-continuous function, indicating whether or not subject i is being observed at time t. We assume that the follow-up time C_i is independent of the counting process $N_i(t)$ completely or given covariates. Denote by Z_i the vector of covariate for subject i, $i = 1, \ldots, n$. Our goal is to make inferences about the covariate effects. Assume the intensity process of $N_i(t)$ has the form

$$\lambda_i(t) = Y_i(t)\lambda_0(t)\exp\{\beta' Z_i\}, \tag{2.6}$$

where $\lambda_0(t)$ is an unknown continuous baseline intensity function and β is a vector of regression parameters. In order to estimate β, Andersen et al. (1985) suggested to use the solution to the equation

$$\frac{\partial C(\tau;\beta)}{\partial \beta} = 0,$$

where

$$C(t;\beta) = \sum_{i=1}^{n} \int_0^t \beta' Z_i dN_i(s) - \int_0^t \log\left\{\sum_{i=1}^{n} Y_i(s)\exp\{\beta' Z_i\}\right\} d\bar{N}(s)$$

with $\bar{N}(t) = \sum_{i=1}^{n} N_i(t)$.

Let $U(t;\beta) = (\partial C(\tau;\beta))/\partial\beta$ and

$$S^{(j)}(t;\beta) = \frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp\{\beta' Z_i\} Z_i^j(t),$$

$j = 0, 1$. Then we get

$$U(t;\beta) = \sum_{i=1}^{n} \int_0^t Z_i dN_i(s) - \int_0^t \frac{S^{(1)}(s;\beta)}{S^{(0)}(s;\beta)} d\bar{N}(s),$$

then

$$U(t;\beta) = \sum_{i=1}^{n} \int_0^t \{Z_i - \bar{Z}(s;\beta)\} Y_i(s) dN_i(s), \tag{2.7}$$

where $\bar{Z}(t;\beta) = S^{(1)}(t;\beta)/S^{(0)}(t;\beta)$. Let $\hat{\beta}$ be the estimator of β. Given $\hat{\beta}$, $\Lambda_0(t) = \int_0^t \lambda_0(s)ds$ can be estimated by

$$\hat{\Lambda}_0(t;\beta) = \sum_{i=1}^{n} \int_0^t \frac{Y_i(s)dN_i(s)}{nS^{(0)}(s;\hat{\beta})}. \tag{2.8}$$

Note that in the above discussion, we assume that only one type of recurrent event exists. More than one type of events can be referred to as multivariate recurrent events.

The asymptotic properties of this proposed point estimator have been discussed in Cook and Lawless (2007), following Anderson and Gill (1982).

2.2.2 Regression Analysis by a Likelihood-based Approach

Let us first review the basic concepts of likelihood-based approach. Suppose $\theta = (\theta_1, \ldots, \theta_p)$ is a p-dimensional parameter, and $L(\theta, X_1, \ldots, X_n)$ is a likelihood for θ based on data $X_1, \ldots, X_n$, which is the joint density or joint probability mass function with random-variable values substituted for the arguments, informally, $L(\theta, X_1, \ldots, X_n)$ is the probability that we observe the sample in hand. Thus, if $x_1, \ldots, x_n$ are the realizations of $X_1, \ldots, X_n$, then the likelihood is given by

$$L(\theta, X_1, \ldots, X_n) = P_\theta(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n).$$

The maximum likelihood estimator (MLE) $\hat{\theta}$ of θ satisfies the equation

$$L(\hat{\theta}, X_1, \ldots, X_n) = \max_{\theta \in \Theta} L(\theta, X_1, \ldots, X_n),$$

if a solution exists. Also, $\hat{\theta}$ maximizes the log-likelihood

$$\mathcal{L}(\theta, X_1, \ldots, X_n) = \log L(\theta, X_1, \ldots, X_n),$$

because applying any strictly increasing transformation to the likelihood does not change its maximizer. Finding MLE is always proceeded by solving the following system of equations

$$\frac{\partial}{\partial\theta}\mathcal{L}(\theta, X_1, \ldots, X_n) = 0. \tag{2.9}$$

The solutions are not necessarily the global maximizers, but under the situations where MLE exists and is unique, solving the above equations gives the MLE. This system of equations is a vector of score equations; and the term

$$U(\theta) = \frac{\partial}{\partial\theta}\mathcal{L}(\theta, X_1, \ldots, X_n)$$

is called the score statistic for θ based on L (Fleming and Harrington, Chapter 4, pg. 145, 1991).

Now, let us consider recurrent events observed on n independent subjects. Let $N_i(t)$ and Z_i denote the counting process and a covariate vector for subject i, $i = 1, \ldots, n$, respectively. Let $0 < T_{i,1} < \cdots < T_{i,m_i}$ be the observation time points for subject i, where m_i is the number of observations for subject i, and $N_i(T_{i,j})$ be the observed value of $N_i(t)$ at $t = T_{i,j}$ for $j = 1, \ldots, m_i; i = 1, \ldots, n$. For estimating the regression parameter in model (2.5), assume that the counting processes $N_i(t)$'s are non-homogeneous Poisson processes. Then following Sun and Zhao (2013), by ignoring the dependence of $\{N_i(T_{i,j}), j = 1, \ldots, m_i\}$ for each i, we can derive the following log likelihood function

$$\mathcal{L}(\mu_0, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} \{N_{i,j} \log[\mu_0(T_{i,j})] + N_{i,j}\beta' Z_i - \mu_0(T_{i,j}) \exp(\beta' Z_i)\} . \qquad (2.10)$$

Hence, a natural estimator of β is some $\hat{\beta}$ such that for some function $\hat{\mu}(t)$, the pair $(\hat{\mu}(t), \hat{\beta})$ maximizes $\mathcal{L}(\mu_0, \beta)$.

The formulation of the asymptotic normality of this type of estimator is derived by Fleming and Harrington (1991).

2.2.3 Regression Analysis by the Estimating Equation Approach

Let $N_i(t)$, Z_i, $T_{i,j}$, and $N_i(T_{i,j})$ be defined as above. Let $\tilde{H}_i(t)$ be the underlying observation process, and let C_i be the follow-up time for subject $i, i = 1, \ldots, n$. If the observation process is continuous, then we have recurrent event data, so $\tilde{H}_i(t)$ is the identity function that constantly equals to 1; and therefore, by Lawless and Nadeau (1995), we have the following estimating equation for β

$$U_0(\beta) = \sum_{i=1}^{n} \int \left\{ Z_i - \frac{\sum_{l=1}^{n} I(C_l \geq t) \exp(\beta' Z_l) Z_l}{\sum_{l=1}^{n} I(C_l \geq t) \exp(\beta' Z_l)} \right\} dN_i(t). \qquad (2.11)$$

For panel count data, define $H_i(t) = \tilde{H}_i\{\min(t, C_i)\} = \sum_{j=1}^{m_i} I(T_{i,j} \leq t)$ as the true observation process on subject i, presenting the number of observations up to time t for subject i; and assume that the mean function of $N_i(t)$ follows model (2.5). $N_i(t)$ is observed only at time points where $H_i(t)$ jumps. To model the observation process on covariates, we assume that the conditional mean function of $\tilde{H}_i(t)$ has the form

$$\lambda_{Z_i}(t) = E\{\tilde{H}_i(t)|Z_i\} = \lambda_0(t)\exp(\gamma' Z_i). \tag{2.12}$$

Let us first consider a simple case where $\gamma = 0$. That is, the observation process is completely independent of the covariates. Then, under model (2.5), following Wilcoxon-type statistic, we have a natural estimating equation to estimate β, which is given by

$$\begin{aligned} U_1(\beta) = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}(Z_i - Z_j)\Big[\quad & \exp(-\beta' Z_i)\int N_i(t)dH_i(t) \\ & - \exp(\beta' Z_j)\int N_j(t)dH_j(t)\Big] = 0, \end{aligned}$$

where $\int N_i(t)dH_i(t) = \sum_{j=1}^{m_i} N_i(T_{i,j})$. $U_1(\beta)$ is an estimating function for model (2.5), and the resulting $\hat{\beta}$ is a consistent estimator of β.

For the general case where $\gamma \neq 0$, to estimate γ in model (2.12), since the observation process is another recurrent event process, we get a similar estimating equation as (2.11) for γ

$$U_\gamma(\gamma) = \sum_{i=1}^{n}\int\left\{Z_i - \frac{\sum_{l=1}^{n} I(C_l \geq t)\exp(\gamma' Z_l)Z_l}{\sum_{l=1}^{n} I(C_l \geq t)\exp(\gamma' Z_l)}\right\} dH_i(t). \tag{2.13}$$

Now, define

$$d\tilde{M}_i(t) = dH_i(t) - I(C_i \geq t)\exp(\gamma' Z_i)d\lambda_0(t).$$

Then, we have

$$\bar{N}_i = \int N_i(t)dH_i(t) = \int N_i(t)d\tilde{M}_i(t) + \int N_i(t)\exp(\gamma' Z_i)I(C_i \geq t)d\lambda_0(t).$$

Under model (2.11), we have the conditional mean

$$E\left\{\bar{N}_i|Z_i\right\} = \exp\{(\gamma+\beta)'Z_i\}\int \lambda_0(t)P(C_i \geq t)\mu_0(t)dt, \tag{2.14}$$

which can be rewritten as

$$E\left\{\bar{N}_i|Z_i\right\} = \exp[(\gamma+\beta)'Z_i) + \xi_0],$$

where $\xi_0 = \log \int \lambda_0(t)P(C_i \geq t)\mu_0(t)dt$; and the resulting estimating equation is given by

$$U_2(\beta|\gamma, Z) = \frac{1}{n}\sum_{i=1}^{n} Z_i\{\bar{N}_i - \exp[(\gamma+\beta)'Z_i) + \xi_0]\} = 0. \tag{2.15}$$

The performance of this estimator's asymptotic normality has been discussed in Sun and Zhao (2013).

2.3 Measurement Error in Survival Analysis

In survival analysis, Prentice (1982) proposed likelihood approaches for Cox regression (induced proportional hazards model), assuming normal measurement errors and rare disease. Regression calibration (RC) was applied by Wang et al (1997) to obtain a partial score function. They also investigated the performance of the estimator through simulation studies. When repeated surrogate measurements are available, Xie, Wang, and Prentice (2001) discussed the estimation of the calibration function for the proportional hazards model. Under an additive error model, Carroll et al (2006) developed methods on estimating the calibration function. Based on the concept of corrected scores, Nakamura (1992) constructed unbiased estimating equations to estimate the regression parameters. Lin and Ying (1993) proposed a non-parametric approach under the missing data context. A comprehensive discussion on survival analysis with measurement errors can be found in Yi (2017). We will apply the classical additive measurement error model and borrow ideas from survival analysis on the estimation of measurement error variance in both Chapters 3 and 4.

2.4 Measurement Error in Recurrent Event Analysis

In the field of analysis of recurrent event data, limited research has been conducted on measurement errors in covariates. For parametric models, Turnbull et al. (1997) proposed adjustments for the usual maximum likelihood estimators under Poisson regression to analyze recurrent event data with measurement errors, including the standard errors and the associated significance tests in order to account for the presence of measurement errors in some of the covariates. Under a semiparametric model, Jiang et al. (1999) considered a moment corrected method to adjust the bias produced by ignoring the measurement error with a normal measurement error assumption. However, their approach assumes non-informative and completely independent censoring. To incorporate informative censoring, Yu et al. (2016) developed two corrected approaches through a random effect intensity model with an unspecified frailty distribution, and studied the large sample properties of their proposed estimators. In Chapter 3, we will present estimating equation based approaches to accommodate the presence of measurement errors.

Even less has been studied for panel count data with measurement errors on some of the covariates. Kim (2007) studied a measurement error model in which the continuous covariates may be measured with errors. He used a pseudo-likelihood and iterative procedure for the estimation procedure. In Chapter 4, we will propose estimating equation based approaches to correct the bias produced by ignoring measurement errors.

Chapter 3: Regression Analysis of Recurrent Event Data with Measurement Errors

3.1 Introduction

In many study fields, researchers are interested in analyzing the patterns of the occurrences or the event history of certain events. In general, there are two types of event history studies. One is the type of events that can occur only once, and the resulting data of this kind of events are called survival or failure time data. The other type is that the event can occur repeatedly, which are usually referred to as recurrent events. For the first type of events, it might be the case that the event itself can occur only once, for example, death. Or, it might be the case that the event can actually occur multiple times, but the focus of the study is on the first time to the occurrence, for example, first marriage.

With respect to recurrent events, they can also be classified into two types. The first type of recurrent events often arise when the study subjects are being monitored continuously and all occurrence times of the events are observed. This type of data is usually referred to as recurrent event data. For instance, the occurrences of mammary tumors were recorded for 48 female rats (Gail et al., 1980). The other type refers to the situation in which the study subjects are only observed at discrete time points, and therefore, only the number of occurrences of the events between two consecutive observation times are available. These data have been referred to as panel count data. In Chapter 3, we will focus on recurrent event data.

Recurrent event data are frequently encountered in longitudinal follow-up studies. In such studies, the observations of recurrent events could be terminated at or before the end of the study. It could be the end of the study, subject dropout, death, loss to follow-up, and so on. Regression analysis of recurrent event data with accurate recording of all the covariates have been discussed by many researchers in the past decades. For instance, Lin et al. (2000) proposed an estimating equation based method for the estimation of the regression parameters, and the asymptotic properties were derived using martingale approaches. However, in many applications, part of the covariates may be recorded with possible measurement errors, and it is inappropriate to simply assume that those measurement errors are ignorable.

Therefore, Turnbull et al. (1997) proposed adjustments for the usual maximum likelihood estimators to analyze recurrent event data with measurement errors, including the standard errors and the associated significance tests in order to account for the presence of measurement errors in some of covariates. They illustrated the technique with a mixed effects Poisson regression model. To incorporate informative censoring, Yu et al (2016) proposed two corrected approaches through a random effect intensity model with an unspecified frailty distribution, and studied the large sample properties of their proposed estimators. In this chapter, we consider two approaches that are based on estimating equations and provide adjusted estimating equations to obtain consistent and asymptotic estimators of the regression parameters. Our new methods are straightforward and easy to implement.

The remainder of this chapter is organized as follows. First, we will introduce the notations and two semiparametric regression models in Section 3.2. Estimation procedures for both models are presented in Sections 3.3 and 3.4, respectively. In Section 3.5, we will present simulation results for the performance evaluation of the proposed methods. We will apply both methods to the study of gamma interferon in chronic granulomatous disease in Section 3.6. Conclusions and discussions are given in Section 3.7.

3.2 Notation and Models

Consider a study involving n subjects who may experience recurrent events. Denote by X_i and Z_i the p_1 and p_2 dimensional vectors of covariates, and by C_i the censoring time for $i = 1, \ldots, n$. Let $N_i(t)$ be the cumulative number of events that have occurred before time t for $i = 1, \ldots, n$ and $0 \leq t \leq \tau$, where τ denotes the end of the study. Suppose that X_i's are completely observed, and that Z_i's are not available but we can observe the corresponding surrogate W_i's such that

$$W_i = Z_i + \epsilon_i, \tag{3.1}$$

where $\epsilon_i \sim N(0, \Sigma)$ with a known covariance matrix Σ. The observed data consist of $\{O_i, W_i, i = 1, \ldots, n\}$, where $O_i = (T_{i,1}, \ldots, T_{i,m_i}, m_i, C_i, X_i)$, $0 < T_{i,1} < \cdots < T_{i,m_i}$ are the observation time points for subject i, and m_i is the total number of observations for subject i.

Next we will consider two models with different assumptions.

Model A. Suppose that the censoring time C_i is conditionally independent of $N_i(\cdot)$ given X_i and Z_i. Without making any assumptions on the distribution of the counting process, the conditional mean of $N_i(t)$ given (X_i, Z_i) takes the form

$$E(N_i(t)|X_i, Z_i) = \mu_0(t) \exp(\beta'_{x0} X_i + \beta'_{z0} Z_i) , \tag{3.2}$$

where $\beta_0 = (\beta'_{x0}, \beta'_{z0})'$ is the vector of regression parameters to be estimated, and $\mu_0(t)$ is a completely unspecified baseline mean function.

Model A is versatile since it allows arbitrary dependence structure among recurrent events and is applicable to any counting processes. This model has been widely studied in the current literature if there is no measurement error in the X_i's and Z_i's. Among others, Lin et al. (2000) proposed a semiparametric regression analysis method for the mean function when the Z_i's are fully observed.

Model B. In many applications, however, as Lin et al. (2000) stated, some subjects are more prone to recurrent events than others. In addition, the observation of recurrent events could be terminated by informative dropouts or failure events and it is unrealistic to assume that the censoring times are conditionally independent given covariates. This heterogeneity and informative censoring can be characterized through the random-effect intensity model:

$$\lambda_i(t|X_i, Z_i, \varphi_i) = \varphi_i \lambda_0(t) \exp(\beta'_{x0} X_i + \beta'_{z0} Z_i), \tag{3.3}$$

where $\lambda_i(t)$ is the intensity function of $N_i(t)$ associated with $\mathcal{F}_t = \{N_i(s), X_i, Z_i, \varphi_i : 0 \le s < t, i = 1, \ldots, n\}$, $\lambda_0(t)$ is a completely unspecified baseline function, and φ_i is a non-negative valued latent variable.

Given (X_i, Z_i, φ_i), $N_i(\cdot)$ is conditionally independent of C_i, and we assume that $N_i(t)$ is a non-stationary Poisson process with the intensity function given by (3.3). The latent variable φ_i satisfies $E[\varphi_i | X_i, Z_i] = 1$.

Let 0_d denote a d dimensional zero vector and $0_{d\times d}$ a $d \times d$ zero matrix. For a vector a, let $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$, $a^{\otimes 2} = aa'$.

3.3 Estimation Procedure for Model A

In this section, we discuss the estimation of $\beta_0 = (\beta'_{x0}, \beta'_{z0})'$ under model A. When Z_i's are completely observed, Lin et al. (2000) proposed the following estimating equation:

$$U^*_{nA}(\beta) = \sum_{i=1}^{n} \int \left[\xi_i - \frac{\sum_{j=1}^{n} Y_j(t) \exp\big(\xi'_j \beta\big) \xi_j}{\sum_{i=1}^{n} Y_j(t) \exp\big(\xi'_j \beta\big)} \right] dN_i(t), \tag{3.4}$$

where $\xi_i = (X'_i, Z'_i)'$, $Y_i(t) = I(C_i \ge t)$ and $N_i(t) = N^*_i(t \wedge C_i)$.

When the covariates Z_i's are completely observed, the estimating equation (3.4) is used for the estimation of β and its asymptotic properties can be derived by martingale approaches (Lin et al., 2000). However, replacing the unobserved Z_i's with the observed surrogate W_i's would result in biased estimators.

Following Lemma A1 in Section 3.9.2, we propose the following estimating equation

$$U_{nA}(\beta)=\sum_{i=1}^{n}\int\left[\xi_{i1}-\frac{\sum_{j=1}^{n}Y_j(t)\exp\big(\xi_{j1}'\beta\big)\xi_{j2}}{\sum_{j=1}^{n}Y_j(t)\exp\big(\xi_{j1}'\beta\big)}\right]dN_i(t), \tag{3.5}$$

where $\xi_{i1}=(X_i',W_i')'$, $\xi_{i2}=(X_i',(W_i-\Sigma\beta_z)')'$.

Denote $\hat{\beta}_{nA}=(\hat{\beta}_{nAx}',\hat{\beta}_{nAz}')'$ as the solution to $U_{nA}(\beta)=0$. In Section 3.9.2, we show that the proposed estimator $\hat{\beta}_{nA}$ is consistent and asymptotically normal, which is stated as the following theorem.

Theorem 3.1. Suppose that the conditions (C1a)-(C4a) shown in Section 3.9.2 hold and Model A holds. Then $\sqrt{n}(\hat{\beta}_{nA}-\beta_0)$ weakly converges to a zero-mean normal distribution with covariance matrix $D_A^{-1}(\beta_0)\Sigma_A D_A^{-1}(\beta_0)$, where D_A and Σ_A can be consistently estimated by

$$\hat{D}_{nA}=\frac{1}{n}\sum_{i=1}^{n}\int\Bigg[\quad\frac{\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)\hat{\xi}_{j2}\hat{\xi}_{j1}'}{\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)}$$

$$+\frac{\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)\hat{\xi}_{j2}\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)\hat{\xi}_{j1}'}{\Big\{\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)\Big\}^2}\Bigg]dN_i(t),$$

and

$$\hat{\Sigma}_{nA}=\frac{1}{n}\sum_{i=1}^{n}\left\{\int\left[\hat{\xi}_{i1}-\frac{\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)\hat{\xi}_{j2}}{\sum_{j=1}^{n}Y_j(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)}\right]d\hat{M}_i(t)\right\}^{\otimes 2}$$

where $d\hat{M}_i(t)=dN_i(t)-Y_i(t)\exp\Big(\xi_{j1}'\hat{\beta}_{nA}\Big)d\hat{\mu}_0(t)$.

3.4 Estimation Procedure for Model B

In this section, we discuss the estimation of $\beta_0 = (\beta'_{x0}, \beta'_{z0})'$ under Model B. Wang et al. (2001) discussed the estimation of β_0 when the covariate Z_i's are completely observed. Yu et al. (2016) considered the same model with measurement errors. They proposed two different corrections to adjust for the bias produced by ignoring measurement errors. Our model B is similar as those in Wang et al. (2001) and Yu et al. (2016), but will employ a different approach on correcting the bias.

Define $F_0(t) = \lambda_0(t)/\Lambda_0(\tau)$. Assume that $\lambda_0(t)$ has a first derivative function with respect to t. Following Wang et al. (2001), although the covariate Z_i's are not completely observed, the distribution function $F_0(t)$ can be estimated by a simple product-limit estimator $\hat{F}_n(t)$, where

$$\hat{F}_n(t) = \prod_{s_{(l)} \geq t} (1 - d_{(l)}/N_{(l)}).$$

In the above, $\{s_{(l)}\}$ are the ordered and distinct values of the event times $\{T_{i,j}\}$, $d_{(l)}$ is the number of events occurring at time $s_{(l)}$, and $N_{(l)}$ is the total number of events with the event time and the censoring time satisfying $T_{i,j} \leq s_{(l)} \leq C_i$.

Let $\gamma_0 = (\beta_{I0}, \beta'_{x0}, \beta'_{z0})'$ with $\beta_{I0} = \log(\mu_0(\tau))$. By observing

$$E[m_i F_0^{-1}(C_i)|X_i, Z_i, C_i] = \exp(\xi'_i \gamma_0)\mu_0(\tau),$$

when Z_i's are completely observed, Wang et al. (2001) proposed the following estimating equation:

$$U^*_{nB}(\gamma) = \sum_{i=1}^{n} \bar{\xi}_i \Big[m_i \hat{F}_n^{-1}(C_i) - \exp\big(\bar{\xi}_i' \gamma\big) \Big]$$

where $\bar{\xi}_i = (1, X_i', Z_i')'$, $\gamma' = (\beta_I, \beta_x', \beta_z')'$.

Following similar ideas as those leading to the estimating equation (3.5) and Lemma A1 in Section 3.9.2, when the covariates Z_i's are not observed, we propose the following estimating equation:

$$U_{nB}(\gamma) = \sum_{i=1}^{n} \Big[\bar{\xi}_{i2} m_i \hat{F}_n^{-1}(C_i) - \bar{\xi}_{i2} \exp\Big(-\frac{1}{2} \beta_z' \Sigma \beta_z \Big) \exp\big(\bar{\xi}_{i1}' \gamma\big) \Big] \tag{3.6}$$

where $\bar{\xi}_{i1} = (1, X_i', W_i')'$, $\bar{\xi}_{i2} = (1, X_i', (W_i - \Sigma\beta_z)')'$.

Denote $\hat{\gamma}_{nB} = (\widehat{\log(\mu_0(\tau))}, \hat{\beta}_{nBx}', \hat{\beta}_{nBz}')'$ as the solution to $U_{nB}(\gamma) = 0$. To derive the asymptotic covariance matrix of $\hat{\gamma}_{nB}$, we introduce some notations. Let $Q_n(t) = (1/n) \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(T_{i,j} \le t)$, $R_n(t) = (1/n) \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(T_{i,j} \le t \le C_i)$, and

$$b_{in}(t) = \sum_{j=1}^{m_i} \Big\{ \int_t^\tau \frac{I(T_{i,j} \le u \le C_i)}{R_n^2(u)} dQ_n(u) - \frac{I(t \le T_{i,j} \le \tau)}{R_n(T_{i,j})} \Big\}.$$

In Section 3.9.2, we show that the proposed estimator $\hat{\gamma}_{nB}$ is consistent and asymptotically normal, which is stated as the following theorem.

Theorem 3.2. When the conditions (C1b)-(C4b) shown in Section 3.9.2 hold and Model B holds, $\sqrt{n}(\hat{\gamma}_{nB} - \gamma_0)$ weakly converges to a zero-mean normal distribution with covariance matrix $D_B^{-1}(\gamma_0) \Sigma_B D_B^{-1}(\gamma_0)$, where D_B and Σ_B can be consistently estimated by

$$\hat{D}_{nB} = \frac{1}{n}\sum_{i=1}^{n} \exp\left(-\frac{1}{2}\hat{\beta}'_{nBz}\Sigma\hat{\beta}_{nBz}\right)\exp\left(\bar{\xi}'_{i1}\hat{\gamma}_{nB}\right)[\hat{\bar{\xi}}_{i2}^{\otimes 2} - \mathrm{diag}(0, 0_{p_1\times p_1}, \Sigma)],$$

and

$$\hat{\Sigma}_{nB} = \frac{1}{n}\sum_{i=1}^{n}\Bigg[\bar{\xi}_{i1}\frac{m_i}{\hat{F}_n(C_i)} - \hat{\bar{\xi}}_{i2}\exp\left(-\frac{1}{2}\hat{\beta}'_{nBz}\Sigma\hat{\beta}_{nBz}\right)\exp\left(\bar{\xi}'_{i1}\hat{\gamma}_{nB}\right) \\ -\frac{1}{n}\sum_{j=1}^{n}\bar{\xi}_{j1}\frac{m_j}{\hat{F}_n(C_j)}b_i(C_j)\Bigg]^{\otimes 2},$$

respectively.

3.5 Estimation of Σ when Σ is unknown

In some cases, Σ can be obtained from some prior knowledge. The estimating equations (3.5) and (3.6) rely on a known Σ. However, in general, Σ is unknown, and thus it needs to be estimated first.

Borrowing ideas from survival analysis on measurement error covariance estimation, we can estimate it in two possible ways. One is that we may use validation data. This means that for some subjects, we can observe (Z_i, W_i) simultaneously. Given model (3.1), then Σ can be consistently estimated as

$$\hat{\Sigma}_{nV} = \frac{1}{n_v}\sum_{validation\ data}(W_i - X_i)^{\otimes 2},$$

where n_v is the total number of subjects in the validation data. If $\liminf_n n_v/n >$ a positive constant, the estimating equations (3.5) and (3.6) still work.

The second way to estimate Σ is to use replication data. This means that all the Z_i's can not be completely observed, but for each $i = 1, \ldots, n$ we can make r_i independent measurements of Z_i ($r_i \geq 2$) for $i = 1, \ldots, n$. Assuming

$$W_{i,j} = Z_i + \epsilon_{i,j}, \quad i = 1, \ldots, n; \quad j = 1, \ldots, r_i; \quad (\epsilon_{i,1}, \ldots, \epsilon_{i,r_i})' \;\; i.i.d \sim N(0, \Sigma),$$

a maximum likelihood estimator of Σ can be obtained.

3.6 Simulation

To evaluate the performance of both proposed methods under various situations, simulation studies have been conducted with a focus on the estimation of β's in the presence of measurement errors on some of the covariates. We assume that X_i, the covariate without measurement errors, follows a Bernoulli distribution with success probability 0.5; and the covariate with measurement errors $W_i = Z_i + e_i$ for $i = 1, \ldots, n$, where Z_i's follow a normal distribution with mean 0 and variance 1; and e_i's follow a normal distribution with mean 0 and variance $\Sigma = \sigma^2$. The censoring times C_i's are generated from a uniform distribution over $(\tau/3, \tau)$ for Model A, where $\tau = 1$.

For Model B, conditioning on X_i, Z_i, the latent variable φ_i is defined as $\varphi_i = \exp(-X_i \ln(2.75))\varphi_i^*$, where φ_i^* follows the density

$$f(\varphi^*|x_i, z_i) = I(0.5 \leq \varphi^* \leq 1.5)(1 - x_i)^{I(z_i \leq 0)} + I(1.5 \leq \varphi^* \leq 4)x_i^{I(z_i \geq 0)}/2.5.$$

It can be verified that $E[\varphi_i|X_i, Z_i] = 1$. C_i's are generated from a truncated exponential distribution with mean $1/\varphi_i$ on $(\tau/3, \tau)$.

Following He et al. (2008), the observation times $(T_{i,1}, \ldots, T_{i,m_i})$ are generated as

$$T_{i,j+1} = T_{i,j} - \log(U_{i,j+1})\{\exp(\xi'_{i1}\beta)d\mu_0\}^{-1},$$

censored by C_i, where $T_{i,0} = 0, U_{i,j} \sim U(0,1)$, and m_i is the number of observation times for subject i, $i = 1, \ldots, n$.

The results given below are based on $\mu_0(t) = \theta t$, where $\theta = 1$, $n = 200$ and 500, $\sigma = 0.1, 0.3, 0.5$, and 500 replications. For detailed tables with more values for σ, please refer to Section 3.9.1.

Table 3.1: Simulation Results for Method A, $n = 200$

σ	Methods	β	Relative Bias	SEE	SSE	CP
0.1	Proposed	(1,1)	(-0.40%, **-0.53%**)	(0.129, 0.058)	(0.138, 0.064)	(0.93, 0.91)
		(1,-1)	(-1.24%, **-0.14%**)	(0.130, 0.058)	(0.132, 0.063)	(0.95, 0.91)
		(-1,1)	(-0.08%, **-1.02%**)	(0.212, 0.092)	(0.219, 0.092)	(0.94, 0.91)
	Lin et al. (2000)	(1,1)	(-1.66%, **-2.59%**)	(0.130, 0.058)	(0.133, 0.061)	(0.94, 0.91)
		(1,-1)	(-0.87%, **-2.64%**)	(0.129, 0.057)	(0.128, 0.065)	(0.94, 0.89)
		(-1,1)	(0.523%, **-2.267%**)	(0.213, 0.095)	(0.215, 0.098)	(0.94, 0.92)
0.3	Proposed	(1,1)	(-0.77%, **-0.36%**)	(0.141, 0.062)	(0.152, 0.073)	(0.92, 0.90)
		(1,-1)	(0.04%, **-0.79%**)	(0.143, 0.061)	(0.149, 0.077)	(0.94, 0.90)
		(-1,1)	(-1.39%, **-1.60%**)	(0.224, 0.094)	(0.237, 0.124)	(0.93, 0.88)
	Lin et al. (2000)	(1,1)	(-0.43%, **-9.63%**)	(0.142, 0.059)	(0.136, 0.069)	(0.96, 0.57)
		(1,-1)	(-1.56%, **-9.58%**)	(0.140, 0.059)	(0.146, 0.079)	(0.94, 0.57)
		(-1,1)	(-1.55%, **-9.77%**)	(0.220, 0.088)	(0.220, 0.102)	(0.95, 0.75)
0.5	Proposed	(1,1)	(1.29%, **2.14%**)	(0.169, 0.072)	(0.190, 0.131)	(0.92, 0.74)
		(1,-1)	(-0.71%, **1.14%**)	(0.168, 0.072)	(0.182, 0.133)	(0.94, 0.73)
		(-1,1)	(0.21%, **1.56%**)	(0.240, 0.093)	(0.261, 0.160)	(0.92, 0.76)
	Lin et al. (2000)	(1,1)	(-1.03%, **-20.90%**)	(0.156, 0.056)	(0.159, 0.075)	(0.95, 0.14)
		(1,-1)	(-1.24%, **-20.57%**)	(0.156, 0.057)	(0.162, 0.074)	(0.93, 0.14)
		(-1,1)	(2.01%, **-20.67%**)	(0.230, 0.077)	(0.244, 0.108)	(0.95, 0.29)

Tables 3.1 and 3.2 present the simulation results for Method A with different setups for $n = 200$ and $n = 500$, repectively. Both tables include the relative bias (Relative Bias) given by the average of the point estimates $\hat{\beta}$ minus the true value

Table 3.2: Simulation Results for Method A, $n = 500$

σ	Methods	β	Relative Bias	SEE	SSE	CP
0.1	Proposed	(1,1)	(-0.36%, **-0.15%**)	(0.041, 0.019)	(0.040, 0.012)	(0.94, 0.93)
		(1,-1)	(-0.27%, **-0.38%**)	(0.042, 0.019)	(0.043, 0.019)	(0.94, 0.94)
		(-1,1)	(-0.67%, **-0.24%**)	(0.068, 0.030)	(0.070, 0.032)	(0.94, 0.94)
	Lin et al. (2000)	(1,1)	(-0.28%, **-1.37%**)	(0.042, 0.019)	(0.043, 0.018)	(0.94, 0.88)
		(1,-1)	(0.20%, **-1.31%**)	(0.042, 0.019)	(0.040, 0.020)	(0.96, 0.86)
		(-1,1)	(0.21%, **-1.43%**)	(0.068, 0.030)	(0.070, 0.031)	(0.95, 0.91)
0.3	Proposed	(1,1)	(-0.19%, **-0.11%**)	(0.051, 0.025)	(0.050, 0.026)	(0.96, 0.92)
		(1,-1)	(-0.20%, **-0.37%**)	(0.051, 0.025)	(0.054, 0.027)	(0.94, 0.91)
		(-1,1)	(-0.26%, **-0.31%**)	(0.074, 0.033)	(0.070, 0.041)	(0.96, 0.88)
	Lin et al. (2000)	(1,1)	(-0.15%, **-8.98%**)	(0.070, 0.032)	(0.075, 0.034)	(0.93, 0.22)
		(1,-1)	(-0.96%, **-8.88%**)	(0.069, 0.031)	(0.071, 0.038)	(0.94, 0.23)
		(-1,1)	(-0.31%, **-9.30%**)	(0.102, 0.043)	(0.102, 0.048)	(0.95, 0.41)
0.5	Proposed	(1,1)	(0.43%, **0.52%**)	(0.067, 0.035)	(0.070, 0.041)	(0.93, 0.92)
		(1,-1)	(-0.56%, **0.62%**)	(0.067, 0.035)	(0.071, 0.035)	(0.93, 0.96)
		(-1,1)	(-0.04%, **0.24%**)	(0.086, 0.039)	(0.090, 0.042)	(0.94, 0.91)
	Lin et al. (2000)	(1,1)	(0.03%, **-20.44%**)	(0.060, 0.0250)	(0.059, 0.031)	(0.95, 0.00)
		(1,-1)	(0.37%, **-20.33%**)	(0.060, 0.025)	(0.064, 0.032)	(0.93, 0.01)
		(-1,1)	(-0.14%, **-20.59%**)	(0.080, 0.030)	(0.082, 0.039)	(0.95, 0.00)

of β over β, the average of the standard deviation estimates (SEE), the sample standard deviation of $\hat{\beta}$ (SSE), and the empirical 95% coverage probability (CP) for β. All numbers are given in parentheses with a comma separating the results of β_x and β_z. It can be seen that the point estimates are unbiased. Moreover, the SEE's and SSE's are quite close to each other; and the 95% coverage probabilities are close to 0.95, which suggests that the proposed variance estimate is reasonable and reliable. The tables also show that even when the measurement errors have a large variance, the proposed method, as expected, can relatively accurately estimate the regression parameters, especially when sample size is large.

For comparison, we applied the method proposed by Lin et al. (2000) by ignoring the measurement errors. It is apparent that when σ gets larger, Lin's estimators are more biased than our estimators. It indicates that inappropriately ignoring the measurement errors leads to biased estimates.

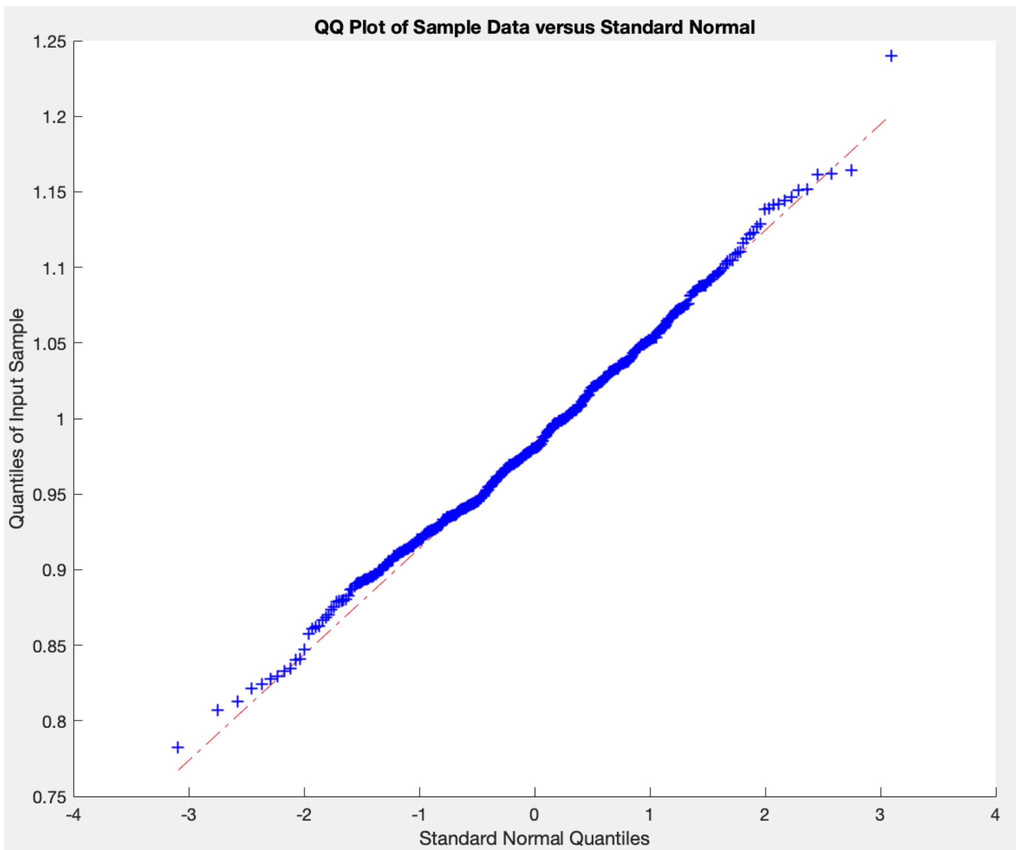

Figure 3.1: Method A QQ Plot for $\hat{\beta}_z$ with $n = 200, \sigma = 0.3$

Figure 3.1 displays a Normal Q-Q plot of $\hat{\beta}_z$ with $n = 200, \sigma = 0.3$. From the plot, we can see that most of the points lie very close to the line, indicating that our normal approximation performs well.

Table 3.3: Simulation Results for Method B, $n = 200$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.1	Proposed	(1,1)	(0.03%, **0.15%**)	(0.123, 0.058)	(0.119, 0.059)	(0.96, 0.95)
		(1,-1)	(0.02%,**0.60%**)	(0.122, 0.059)	(0.124, 0.062)	(0.94, 0.93)
		(-1,1)	(0.65%,**0.50%**)	(0.206, 0.097)	(0.196, 0.092)	(0.95, 0.96)
	Wang et al. (2001)	(1,1)	(-0.16%, **-1.24%**)	(0.124, 0.058)	(0.116, 0.059)	(0.97, 0.90)
		(1,-1)	(0.05%, **-1.10%**)	(0.113, 0.058)	(0.130, 0.059)	(0.95, 0.90)
		(-1,1)	(-0.05%, **-1.31%**)	(0.203, 0.095)	(0.199, 0.092)	(0.95, 0.94)
0.3	Proposed	(1,1)	(0.05%, **1.09%**)	(0.139, 0.078)	(0.139, 0.078)	(0.94, 0.94)
		(1,-1)	(1.35%, **1.75%**)	(0.138, 0.079)	(0.144, 0.082)	(0.94, 0.94)
		(-1,1)	(0.64%, **1.42%**)	(0.216, 0.116)	(0.214, 0.109)	(0.95, 0.96)
	Wang et al. (2001)	(1,1)	(-0.37%, **-7.64%**)	(0.134, 0.065)	(0.140, 0.065)	(0.93, 0.75)
		(1,-1)	(1.06%, **-7.80%**)	(0.137, 0.064)	(0.134, 0.069)	(0.95, 0.73)
		(-1,1)	(1.52%, **-8.07%**)	(0.212, 0.095)	(0.220, 0.098)	(0.94, 0.84)
0.5	Proposed	(1,1)	(1.35%, **3.69%**)	(0.178, 0.141)	(0.179, 0.130)	(0.95, 0.96)
		(1,-1)	(2.44%, **4.17%**)	(0.180, 0.145)	(0.180, 0.143)	(0.95, 0.96)
		(-1,1)	(3.22%, **5.07%**)	(0.255, 0.198)	(0.240, 0.182)	(0.97, 0.97)
	Wang et al. (2001)	(1,1)	(0.38%, **-19.88%**)	(0.151, 0.069)	(0.154, 0.074)	(0.94, 0.21)
		(1,-1)	(0.27%, **-20.07%**)	(0.151, 0.070)	(0.152, 0.075)	(0.94, 0.22)
		(-1,1)	(1.57%, **-20.49%**)	(0.221, 0.095)	(0.228, 0.098)	(0.95, 0.40)

The simulation results for Model B are given in Tables 3.3 and 3.4. Since the variance estimation procedure for Model B is too complicated to compute, we apply a bootstrap approach to estimate the standard deviations for the point estimates. Thus, the BSE's in Tables 3.3 and 3.4 represent the means of bootstrap standard deviation estimates. The results are similar to those for Model A and suggest that the estimation procedure proposed in Section 3.4 also works well for the situations considered here.

The comparison of our proposed estimators and estimators using the approach by Wang et al. (2001) is similar to those in Model A, manifesting that measurement errors are not ignorable.

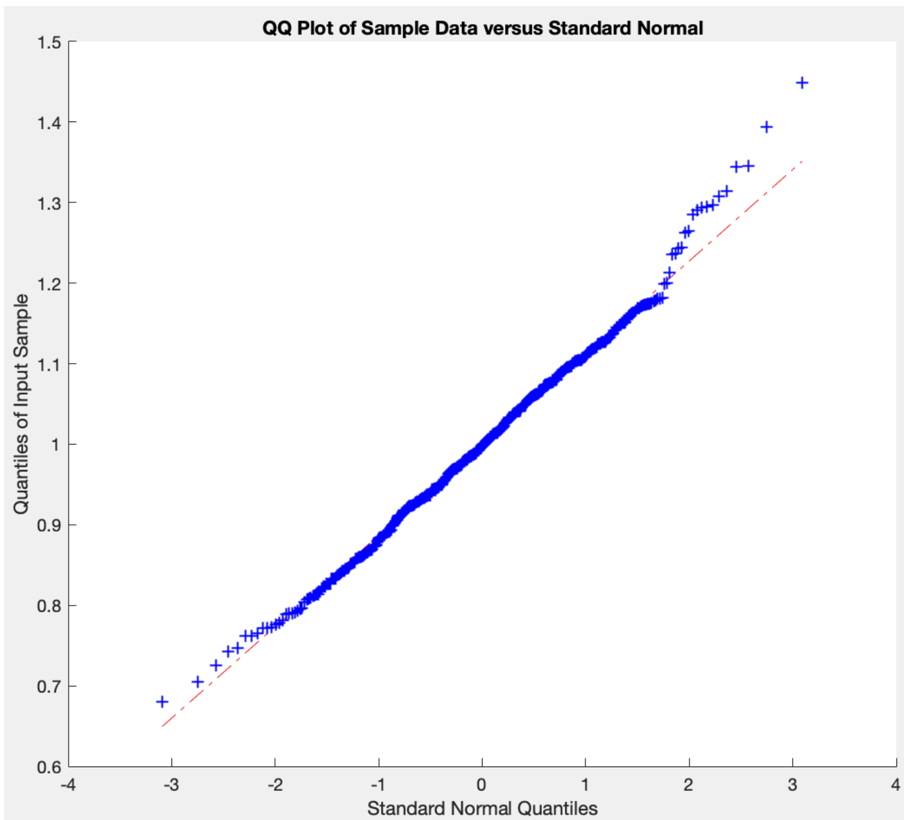

Figure 3.2: Method B QQ Plot for $\hat{\beta}_z$ with $n = 200, \sigma = 0.3$

Table 3.4: Simulation Results for Method B, $n = 500$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.1	Proposed	(1,1)	(0.33%, **-0.11%**)	(0.038, 0.018)	(0.040, 0.019)	(0.93, 0.92)
		(1,-1)	(-0.15%, **0.01%**)	(0.038, 0.018)	(0.038, 0.019)	(0.96, 0.93)
		(-1,1)	(-0.36%, **-0.20%**)	(0.062, 0.029)	(0.062, 0.029)	(0.95, 0.93)
	Wang et al. (2001)	(1,1)	(0.13%, **-1.01%**)	(0.039, 0.018)	(0.039, 0.018)	(0.95, 0.92)
		(1,-1)	(-0.12%, **-1.12%**)	(0.039, 0.018)	(0.038, 0.019)	(0.96, 0.88)
		(-1,1)	(0.07%, **-1.14%**)	(0.062, 0.028)	(0.063, 0.028)	(0.95, 0.92)
0.3	Proposed	(1,1)	(0.08%, **0.64%**)	(0.073, 0.041)	(0.078, 0.047)	(0.91, 0.92)
		(1,-1)	(0.42%, **0.49%**)	(0.073, 0.042)	(0.071, 0.046)	(0.96, 0.93)
		(-1,1)	(-0.02%, **0.69%**)	(0.107, 0.057)	(0.110, 0.060)	(0.95, 0.94)
	Wang et al. (2001)	(1,1)	(0.11%, **-8.51%**)	(0.071, 0.035)	(0.074, 0.036)	(0.95, 0.34)
		(1,-1)	(0.09%, **-8.35%**)	(0.070, 0.035)	(0.070, 0.037)	(0.95, 0.32)
		(-1,1)	(-0.24%, **-8.29%**)	(0.105, 0.049)	(0.104, 0.053)	(0.95, 0.60)
0.5	Proposed	(1,1)	(0.67%, **0.88%**)	(0.065, 0.048)	(0.070, 0.050)	(0.93, 0.93)
		(1,-1)	(0.45%, **-0.67%**)	(0.065, 0.048)	(0.071, 0.051)	(0.93, 0.95)
		(-1,1)	(-0.39%, **0.86%**)	(0.081, 0.056)	(0.080, 0.052)	(0.95, 0.92)
	Wang et al. (2001)	(1,1)	(0.13%, **-19.89%**)	(0.056, 0.027)	(0.057, 0.029)	(0.94, 0.00)
		(1,-1)	(0.31%, **-19.93%**)	(0.056, 0.029)	(0.059, 0.033)	(0.94, 0.00)
		(-1,1)	(0.46%, **-19.92%**)	(0.075, 0.035)	(0.072, 0.038)	(0.95, 0.01)

The normal approximation can also be assessed by Figure 3.2. The plot shows that the quantiles of our estimators match the quantiles for a standard normal distribution, indicating an appropriate normal approximation.

3.7 Application

We now apply the proposed methods to analyze the multiple-infection data taken from the study of gamma interferon in chronic granulomatous disease (CGD) in Fleming and Harrington (1991). CGD is a group of inherited rare disorders of immune function characterized by recurrent pyogenic infections. In order to study the ability of gamma interferon in reducing the rate of serious infections, a double-blinded clinical trial was conducted in which 128 patients were randomized to placebo vs. gamma interferon between October 1988 and March 1989; and through July 15, 1989, all patients were followed to record recurrent infections. By the end of the trial, 30 of the 65 placebo patients and 14 of the 63 gamma interferon patients experienced at least one infection. Among the 30 untreated patients with at least one infection, 5 experienced two infections, 4 had three, and 3 had four or more. Among the 14 treated patients with at least one infection, 4 had two and others had three infections.

For the analysis, we consider $X_i, i = 1, \ldots, 128$ to be the treatment indicator as 0 for the placebo and 1 for the treatment. W_i's are the ages of the patients. Since the ages were recorded only at the time of enrollment, and the study lasted for more than a year, we consider adding a normally distributed error term with mean 0 and variance 0.25. The results for Models A and B are shown in Tables 3.5 and Table 3.6, respectively.

The parameter estimate for treatment in Model A is negative, indicating that the gamma interferon treatment has a significant effect on curing CGD. For the other covariate age, the parameter estimate is negative as well. This suggests that the younger the patients, the more likely they get recurrent infections. This finding is reasonable because CGD usually presents early in life and may lead to death in childhood.

Table 3.5: Estimation of the effects of treatment and age for Model A

Covariate	Parameter Estimate	SEE	95% CI
Treatment	-1.1223	0.3092	(-1.6309, -0.6137)
Age	-0.0303	0.0144	(-0.0540, -0.0066)

Similar results from Model B are shown in Table 3.6. Both our models A and B give very comparable results with the results in Lin et al. (2001), in which the

Table 3.6: Estimation of the effects of treatment and age for Model B

Covariate	Parameter Estimate	BSE	95% CI
Treatment	-1.0755	0.3226	(-1.6061, -0.5449)
Age	-0.0401	0.0198	(-0.0729, -0.0077)

estimated effects of treatment and age were -1.12 and -0.03 with robust SEE 0.309 and 0.014, respectively.

3.8 Conclusion and Discussion

The presence of measurement errors in some of the covariates is often seen in clinical studies, especially for multiple clinical measurements. In the preceding sections, two models and their estimating procedures were presented for regression analysis. Comparing to Yu et al. (2016), we proposed a different adjustment approach (model B) to correct the bias caused by measurement errors. One main advantage of our proposed method is that it is more flexible and can be relatively easy to implement.

The data example presented in Section 3.7 may not be typical, since age is usually not considered as a clinical measurement. In order to demonstrate that our approaches can indeed solve the problems with the presence of measurement errors on some of the covariates, we added some normally distributed errors to age and compared our results to those in Lin et al. (2000). This comparison indicates that our proposed method produces reasonable estimates assuming the presence of measurement errors.

Future research may lie on studies when there are more than one type of event of interest. The resulting data are often referred to as multivariate recurrent event data. It would be interesting to generalize our proposed method to the situation in which the multiple types of recurrent events are dependent. A joint modeling approach will be utilized and similar estimating equation based methods can be developed.

3.9 Appendix

3.9.1 Detailed Tables in Simulation

Table 3.7: Detailed Simulation Results for Model A, $n = 200$

σ	Methods	β	Relative Bias	SEE	SSE	CP
0.1	Proposed	(1,1)	(-0.3926%,**-0.5278%**)	(0.1288,0.0581)	(0.1379,0.0640)	(0.926,0.908)
		(1,-1)	(-1.2386%,**-0.1387%**)	(0.1301,0.0581)	(0.1317,0.0631)	(0.948,0.912)
		(-1,1)	(-0.0801%,**-1.0191%**)	(0.2117,0.0921)	(0.2194,0.0921)	(0.944,0.914)
	Lin et al. (2000)	(1,1)	(-1.6590%,**-2.5860%**)	(0.1295,0.0579)	(0.1328,0.0613)	(0.940,0.908)
		(1,-1)	(-0.8729%,**-2.6355%**)	(0.1291,0.0573)	(0.1278,0.0645)	(0.938,0.886)
		(-1,1)	(0.5229%,**-2.2665%**)	(0.2126,0.0949)	(0.2153,0.0980)	(0.940,0.922)
0.2	Proposed	(1,1)	(-0.7696%,**-1.2048%**)	(0.1344,0.0595)	(0.1382,0.0684)	(0.942,0.894)
		(1,-1)	(-1.6608%,**-1.5191%**)	(0.1347,0.0587)	(0.1438,0.0702)	(0.936,0.874)
		(-1,1)	(1.1355%,**-0.5517%**)	(0.2191,0.0945)	(0.2323,0.1078)	(0.946,0.912)
	Lin et al. (2000)	(1,1)	(-1.5226%,**-5.4127%**)	(0.1347,0.0589)	(0.1332,0.0656)	(0.950,0.770)
		(1,-1)	(-0.5951%,**-5.1530%**)	(0.1340,0.0583)	(0.1345,0.0638)	(0.960,0.824)
		(-1,1)	(0.3937%,**-4.3252%**)	(0.2180,0.0931)	(0.2216,0.1045)	(0.938,0.788)
0.3	Proposed	(1,1)	(-0.7696%,**-0.3553%**)	(0.1413,0.0617)	(0.1520,0.0731)	(0.922,0.898)
		(1,-1)	(0.0440%,**-0.7885%**)	(0.1429,0.0614)	(0.1488,0.0763)	(0.938,0.892)
		(-1,1)	(-1.3852%,**-1.5962%**)	(0.2237,0.0933)	(0.2365,0.1237)	(0.928,0.878)
	Lin et al. (2000)	(1,1)	(-0.4298%,**-9.6349%**)	(0.1419,0.0589)	(0.1356,0.0690)	(0.964,0.572)
		(1,-1)	(-1.5640%,**-9.5773%**)	(0.1398,0.0585)	(0.1463,0.0725)	(0.938,0.570)
		(-1,1)	(-1.5546%,**-9.7719%**)	(0.2204,0.0875)	(0.2204,0.1019)	(0.950,0.746)
0.4	Proposed	(1,1)	(-0.2700%,**0.6238%**)	(0.1536,0.0657)	(0.1600,0.0966)	(0.934,0.876)
		(1,-1)	(0.4424%,**0.9290%**)	(0.1537,0.0656)	(0.1621,0.1071)	(0.932,0.808)
		(-1,1)	(-0.9662%,**0.0669%**)	(0.2303,0.0931)	(0.2338,0.1441)	(0.958,0.824)
	Lin et al. (2000)	(1,1)	(-0.9547%,**-15.2571%**)	(0.1484,0.0575)	(0.1520,0.0727)	(0.942,0.304)
		(1,-1)	(-0.9091%,**-15.1182%**)	(0.1491,0.0582)	(0.1525,0.0744)	(0.936,0.308)
		(-1,1)	(-0.0577%,**-14.8999%**)	(0.2258,0.0830)	(0.2226,0.1017)	(0.946,0.524)
0.5	Proposed	(1,1)	(1.2867%,**2.1383%**)	(0.1689,0.0717)	(0.1899,0.1306)	(0.916,0.736)
		(1,-1)	(-0.7052%,**1.1447%**)	(0.1679,0.0724)	(0.1820,0.1331)	(0.936,0.734)
		(-1,1)	(0.2141%,**1.5966%**)	(0.2391,0.0928)	(0.2609,0.1593)	(0.920,0.762)
	Lin et al. (2000)	(1,1)	(-1.0307%,**-20.9044%**)	(0.1564,0.0563)	(0.1588,0.0749)	(0.948,0.144)
		(1,-1)	(-1.2355%,**-20.5700%**)	(0.1562,0.0567)	(0.1623,0.0743)	(0.932,0.136)
		(-1,1)	(2.0072%,**-20.6685%**)	(0.2299,0.0771)	(0.2436,0.1080)	(0.952,0.292)

Table 3.8: Detailed Simulation Results for Model A, $n = 500$

σ	Methods	β	Relative Bias	SEE	SSE	CP
0.1	Proposed	(1,1)	(-0.3572%,**-0.1466%**)	(0.0413,0.0188)	(0.0403,0.0198)	(0.940,0.930)
		(1,-1)	(-0.2685%,**-0.3822%**)	(0.0420,0.0190)	(0.0432,0.0192)	(0.944,0.944)
		(-1,1)	(-0.6664%,**-0.2437%**)	(0.0679,0.0303)	(0.0699,0.0317)	(0.938,0.940)
	Lin et al. (2000)	(1,1)	(-0.2836%,**-1.3678%**)	(0.0418,0.0190)	(0.0425,0.0182)	(0.940,0.884)
		(1,-1)	(0.2017%,**-1.3130%**)	(0.0420,0.0189)	(0.0402,0.0200)	(0.964,0.862)
		(-1,1)	(0.2080%,**-1.4290%**)	(0.0680,0.0302)	(0.0702,0.0305)	(0.948,0.910)
0.2	Proposed	(1,1)	(-0.2727%,**-0.4751%**)	(0.0638,0.0293)	(0.0655,0.0343)	(0.944,0.898)
		(1,-1)	(-0.3935%,**-0.6526%**)	(0.0634,0.0293)	(0.0646,0.0334)	(0.950,0.902)
		(-1,1)	(-0.1545%,**-0.6364%**)	(0.0988,0.0436)	(0.1002,0.0468)	(0.950,0.926)
	Lin et al. (2000)	(1,1)	(-0.3997%,**-4.7056%**)	(0.0638,0.0291)	(0.0642,0.0303)	(0.954,0.608)
		(1,-1)	(-0.6513%,**-4.7504%**)	(0.0637,0.0293)	(0.0642,0.0322)	(0.948,0.580)
		(-1,1)	(0.2368%,**-4.5199%**)	(0.0983,0.0428)	(0.1005,0.0466)	(0.954,0.788)
0.3	Proposed	(1,1)	(-0.1924%,**-0.1058%**)	(0.0512,0.0254)	(0.0495,0.0262)	(0.964,0.924)
		(1,-1)	(-0.2009%,**-0.3720%**)	(0.0510,0.0254)	(0.0536,0.0271)	(0.938,0.910)
		(-1,1)	(-0.2555%,**-0.3139%**)	(0.0738,0.0329)	(0.0698,0.0406)	(0.956,0.878)
	Lin et al. (2000)	(1,1)	(-0.1502%,**-8.9808%**)	(0.0700,0.0316)	(0.0746,0.0335)	(0.926,0.218)
		(1,-1)	(-0.9569%,**-8.8831%**)	(0.0694,0.0308)	(0.0713,0.0375)	(0.936,0.226)
		(-1,1)	(-0.3077%,**-9.2952%**)	(0.1024,0.0425)	(0.1024,0.0484)	(0.952,0.414)
0.4	Proposed	(1,1)	(-0.7212%,**-0.0463%**)	(0.0807,0.0396)	(0.0849,0.0475)	(0.954,0.912)
		(1,-1)	(0.0499%,**0.4032%**)	(0.0581,0.0299)	(0.0604,0.0330)	(0.944,0.926)
		(-1,1)	(-0.0280%,**0.1691%**)	(0.0787,0.0655)	(0.0769,0.0376)	(0.952,0.886)
	Lin et al. (2000)	(1,1)	(-0.5690%,**-14.2862%**)	(0.0762,0.0326)	(0.0784,0.0425)	(0.934,0.068)
		(1,-1)	(0.0648%,**-14.2544%**)	(0.0759,0.0328)	(0.0704,0.0395)	(0.954,0.070)
		(-1,1)	(-0.3418%,**-14.1591%**)	(0.1078,0.0430)	(0.1092,0.0514)	(0.962,0.148)
0.5	Proposed	(1,1)	(0.4349%,**0.5189%**)	(0.0664,0.0345)	(0.0702,0.0412)	(0.930,0.920)
		(1,-1)	(-0.5580%,**0.6245%**)	(0.0665,0.0345)	(0.0708,0.0352)	(0.930,0.958)
		(-1,1)	(-0.0400%,**0.2410%**)	(0.0857,0.0388)	(0.0897,0.0419)	(0.936,0.908)
	Lin et al. (2000)	(1,1)	(0.0335%,**-20.4390%**)	(0.0596,0.0245)	(0.0591,0.0312)	(0.954,0.002)
		(1,-1)	(0.3744%,**-20.3304%**)	(0.0594,0.0249)	(0.0638,0.0324)	(0.930,0.006)
		(-1,1)	(-0.1361%,**-20.5887%**)	(0.0799,0.0303)	(0.0816,0.0384)	(0.948,0.004)

Table 3.9: Detailed Simulation Results for Model B, $n = 200$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.1	Proposed	(1,1)	(0.0331%,**0.1515%**)	(0.1229, 0.0580)	(0.1187, 0.0591)	(0.964,0.950)
		(1,-1)	(0.0182%,**0.6048%**)	(0.1217, 0.0588)	(0.1236, 0.0615)	(0.944, 0.932)
		(-1,1)	(0.6464%,**0.4985%**)	(0.2060, 0.0969)	(0.1964, 0.0924)	(0.952, 0.960)
	Wang et al. (2001)	(1,1)	(-0.1594%,**-1.2392%**)	(0.1239, 0.0576)	(0.1159, 0.0589)	(0.966, 0.898)
		(1,-1)	(0.0493%,**-1.0949%**)	(0.1129, 0.0578)	(0.1302, 0.0585)	(0.946, 0.902)
		(-1,1)	(-0.0461%,**-1.3081%**)	(0.2032, 0.0945)	(0.1991, 0.0920)	(0.950, 0.942)
0.2	Proposed	(1,1)	(0.8904%,**0.4873%**)	(0.1295,0.0655)	(0.1216, 0.0680)	(0.962, 0.928)
		(1,-1)	(1.6762%,**0.1863%**)	(0.1275,0.0658)	(0.1345,0.0667)	(0.930, 0.942)
		(-1,1)	(0.3578%,**1.4186%**)	(0.2085,0.1038)	(0.2037,0.1015)	(0.960, 0.950)
	Wang et al. (2001)	(1,1)	(0.2575%,**-4.2608%**)	(0.1275,0.0604)	(0.1321,0.0627)	(0.940,0.874)
		(1,-1)	(0.1936%,**-4.0011%**)	(0.1276,0.0606)	(0.1283,0.0600)	(0.952,0.890)
		(-1,1)	(2.9702%,**-3.3931%**)	(0.2115,0.0964)	(0.2070,0.0926)	(0.950,0.898)
0.3	Proposed	(1,1)	(0.0535%,**1.0940%**)	(0.1386,0.0784)	(0.1392,0.0782)	(0.940,0.936)
		(1,-1)	(1.3460%,**1.7489%**)	(0.1383,0.0793)	(0.1444,0.0818)	(0.936,0.938)
		(-1,1)	(0.6367%,**1.4214%**)	(0.2162,0.1158)	(0.2141,0.1090)	(0.954,0.962)
	Wang et al. (2001)	(1,1)	(-0.3705%,**-7.6406%**)	(0.1344,0.0652)	(0.1397,0.0652)	(0.934,0.746)
		(1,-1)	(1.0558%,**-7.8034%**)	(0.1366,0.0637)	(0.1341,0.0693)	(0.952,0.728)
		(-1,1)	(1.5197%,**-8.0734%**)	(0.2123.0.0952)	(0.2203,0.0981)	(0.938,0.844)
0.4	Proposed	(1,1)	(1.0480%,**2.3218%**)	(0.1525,0.0990)	(0.1553,0.1025)	(0.948,0.936)
		(1,-1)	(0.8511%,**2.1337%**)	(0.1528,0.0989)	(0.1538,0.1003)	(0.948,0.944)
		(-1,1)	(2.5495%,**2.3986%**)	(0.2292,0.1437)	(0.2139,0.1307)	(0.974,0.964)
	Wang et al. (2001)	(1,1)	(0.5308%,**-14.0642%**)	(0.1415,0.0664)	(0.1471,0.0702)	(0.930,0.432)
		(1,-1)	(-1.0407%,**-14.0902%**)	(0.1434,0.0673)	(0.1362,0.0698)	(0.956,0.448)
		(-1,1)	(-0.2114%,**-14.3016%**)	(0.2156,0.0963)	(0.2178,0.0977)	(0.940,0.642)
0.5	Proposed	(1,1)	(1.3538%,**3.6922%**)	(0.1779,0.1414)	(0.1788,0.1295)	(0.946,0.964)
		(1,-1)	(2.4405%,**4.1745%**)	(0.1800,0.1449)	(0.1801,0.1432)	(0.954,0.964)
		(-1,1)	(3.2163%,**5.0670%**)	(0.2547,0.1982)	(0.2401,0.1820)	(0.968,0.974)
	Wang et al. (2001)	(1,1)	(0.3803%,**-19.8761%**)	(0.1513,0.0692)	(0.1535,0.0744)	(0.940,0.206)
		(1,-1)	(0.2743%,**-20.0670%**)	(0.1506,0.0697)	(0.1517,0.0749)	(0.942,0.220)
		(-1,1)	(1.5682%,**-20.4892%**)	(0.2213,0.0948)	(0.2280,0.0978)	(0.948,0.402)

Table 3.10: Detailed Simulation Results for Model B, $n = 500$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.1	Proposed	(1,1)	(0.3338%,**-0.1062**%)	(0.0383,0.0180)	(0.0398,0.0187)	(0.932,0.918)
		(1,-1)	(-0.1486%,**0.0100**%)	(0.0383,0.0181)	(0.0375,0.0192)	(0.960,0.930)
		(-1,1)	(-0.3573%,**-0.1978**%)	(0.0621,0.0288)	(0.0623,0.0286)	(0.950,0.934)
	Wang et al. (2001)	(1,1)	(0.1301%,**-1.0092**%)	(0.0385,0.0179)	(0.0388,0.0181)	(0.950,0.916)
		(1,-1)	(-0.1223%,**-1.1150**%)	(0.0387,0.0180)	(0.0384,0.0186)	(0.964,0.884)
		(-1,1)	(0.0745%,**-1.1415**%)	(0.0618,0.0281)	(0.0632,0.0278)	(0.950,0.922)
0.2	Proposed	(1,1)	(-0.1465%,**0.2024**%)	(0.0654,0.0339)	(0.0694,0.0355)	(0.932,0.916)
		(1,-1)	(-0.0024%,**0.0508**%)	(0.0660,0.0343)	(0.0685,0.0363)	(0.942,0.918)
		(-1,1)	(0.2255%,**-0.2281**%)	(0.1017,0.0494)	(0.1059,0.0510)	(0.932,0.936)
	Wang et al. (2001)	(1,1)	(0.3064%,**-3.9134**%)	(0.0646,0.0309)	(0.0639,0.0333)	(0.944,0.690)
		(1,-1)	(-0.2155%,**-4.0571**%)	(0.0647,0.0315)	(0.0620,0.0332)	(0.968,0.726)
		(-1,1)	(-0.6536%,**-3.7357**%)	(0.1013,0.0471)	(0.1013,0.0462)	(0.934,0.866)
0.3	Proposed	(1,1)	(0.0813%,**0.6382**%)	(0.0726,0.0414)	(0.0784,0.0467)	(0.912,0.918)
		(1,-1)	(0.4187%,**0.4892**%)	(0.0729,0.0424)	(0.0712,0.0464)	(0.956,0.930)
		(-1,1)	(-0.0195%,**0.6873**%)	(0.1069,0.0573)	(0.1104,0.0598)	(0.948,0.942)
	Wang et al. (2001)	(1,1)	(0.1083%,**-8.5051**%)	(0.0709,0.0350)	(0.0740,0.0360)	(0.948,0.336)
		(1,-1)	(0.0874%,**-8.3486**%)	(0.0704,0.0345)	(0.0697,0.0369)	(0.948,0.318)
		(-1,1)	(-0.2418%,**-8.2903**%)	(0.1054,0.0486)	(0.1042,0.0529)	(0.954,0.594)
0.4	Proposed	(1,1)	(0.1280%,**1.0826**%)	(0.0861,0.0580)	(0.0861,0.0580)	(0.936,0.920)
		(1,-1)	(0.4385%,**0.8091**%)	(0.0837,0.0544)	(0.0869,0.0551)	(0.942,0.930)
		(-1,1)	(0.0064%,**0.32261**%)	(0.1153,0.0684)	(0.1127,0.0672)	(0.956,0.960)
	Wang et al. (2001)	(1,1)	(-0.0567%,**-13.9094**%)	(0.0774,0.0386)	(0.0792,0.0410)	(0.938,0.108)
		(1,-1)	(0.1901%,**-13.9777**%)	(0.0772,0.0385)	(0.0750,0.0423)	(0.954,0.098)
		(-1,1)	(0.7772%,**-13.9399**%)	(0.1096,0.0507)	(0.1098,0.0533)	(0.936,0.230)
0.5	Proposed	(1,1)	(0.6683%,**0.8845**%)	(0.0646,0.0475)	(0.0697,0.0503)	(0.934,0.934)
		(1,-1)	(0.4549%,**-0.6661**%)	(0.0653,0.0480)	(0.0707,0.0510)	(0.928,0.950)
		(-1,1)	(-0.3856%,**0.8618**%)	(0.0808,0.0557)	(0.0800,0.0524)	(0.948,0.922)
	Wang et al. (2001)	(1,1)	(0.1251%,**-19.8881**%)	(0.0564,0.0286)	(0.0570,0.0293)	(0.944,0.000)
		(1,-1)	(0.3091%,**-19.9309**%)	(0.0562,0.0287)	(0.0587,0.0326)	(0.940,0.000)
		(-1,1)	(0.4553%,**-19.9152**%)	(0.0747,0.0352)	(0.0718,0.0383)	(0.950,0.014)

3.9.2 Asymptotic Properties

To derive the asymptotic properties of the proposed estimator $\hat{\beta}_{1n}$, we need the following regularity conditions for Model A.

(C1a). The true value β_0 belongs to a known open, convex, and bounded set $\mathcal{B}$ in $R^{p_1+p_2}$.

(C2a). The measurement error e_i is independent of $N_i^*(\cdot)$, X_i, Z_i and C_i.

(C3a). $P(C \geq \tau) > 0$, where τ is the largest follow-up time.

(C4a). The matrix

$$D_A(\beta) = E \int E\{Y(t)\exp(\xi'\beta)\}^{-1}\{EY(t)\exp(\xi'\beta)(\xi - \bar{\xi}(\beta;t))^{\otimes 2}\}dN_i(t),$$

where $\bar{\xi}(\beta;t) = \{EY(t)\exp(\xi'\beta)\}^{-1}\{EY(t)\exp(\xi'\beta)\xi\}$, is positive definite for $\beta \in \mathcal{B}$, and $\xi = (X', Z')'$.

For Model B, we need the additional following regularity conditions.

(C1b). The true value γ_0 belongs to a known open, convex, and bounded set Γ in $R^{p_1+p_2+1}$.

(C2b). The measurement error e_i is independent of $N_i^*(\cdot)$, X_i, Z_i and C_i.

(C3b). $\inf_{x,z} P(C \geq \tau | X = x, Z = z) > 0$, and $G(t) = E\varphi I(C \geq t)$ is a continuous function for $t \in [0, \tau]$.

(C4b). The matrix $D_B(\gamma) = E[\exp(\bar{\xi}'\gamma)\bar{\xi}^{\otimes 2}]$ is positive definite for $\gamma \in \Gamma$ where $\bar{\xi} = (1, X', Z')'$.

Lemma A1 Under the assumptions of Theorem 1, one obtains that

$$E\left[\exp(W_i'\beta_z)|O_i, Z_i\right] = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z), \qquad (i)$$

$$E\left[(W_i - \Sigma\beta_z)\exp(W_i'\beta_z)|O_i, Z_i\right] = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)Z_i\exp(Z_i'\beta_z), \qquad (ii)$$

$$E\left[(W_i - \Sigma\beta_z)^{\otimes 2}\exp(W_i'\beta_z)|O_i, Z_i\right] = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)[Z_iZ_i' + \Sigma]\exp(Z_i'\beta_z). \quad (iii)$$

Proof. Note that

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\exp(W_i'\beta_z)\exp\left(-\frac{1}{2}(W_i - Z_i)'\Sigma^{-1}(W_i - Z_i)\right) \\
&= \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&\quad \times\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(W_i - Z_i - \Sigma\beta_z)'\Sigma^{-1}(W_i - Z_i - \Sigma\beta_z)\right),
\end{aligned}$$

which is the normal density function with mean $Z_i + \Sigma\beta_z$ and covariance matrix Σ, muliplied by a term $\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z)$. Detailed derivations of the three equalities in Lemma A1 appear at the end of this section.

Proof of Theorem 3.1 Now we show the consistency and asymptotic normality of the proposed estimator $\hat{\beta}_{nA}$.

Firstly, we show that

$$n^{-1}U_{nA}(\beta_0) \to 0 \ \text{ a.s.} \qquad (3.7)$$

as $n \to \infty$.

Note that the following two classes of functions $\mathcal{F}_1$ and $\mathcal{F}_2$, indexed by $\beta \in \mathcal{B}$, where $\mathcal{B}$ is a open, convex, and bounded neighborhood of β_0,

$$\mathcal{F}_1 = \{Y(t)\exp(X'\beta_x + W'\beta_z) : \quad \beta \in \mathcal{B}, t \in [0,\tau]\}$$

and

$$\mathcal{F}_2 = \{Y(t)\exp(X'\beta_x + W'\beta_z)(X', (W - \Sigma\beta_z)')' : \quad \beta \in \mathcal{B}, t \in [0,\tau]\}$$

are Glivenko-Cantelli classes, because $Y(t)$ is Donsker as a process in $\ell^\infty([0,\tau])$; $\{(X,Z)\}$ are Donsker; e follows a zero-mean normal distribution with known variance, so $\{(X, Z+e)\} = \{(X,W)\}$ are Donsker; and therefore so is $\{X'\beta_x + W'\beta_z\}$. The class $\{\exp(X'\beta_x + W'\beta_z)\}$ is Donsker since exponential is Lipschitz continuous on compacts. Hence, $\mathcal{F}_1$ and $\mathcal{F}_2$ are all Donsker classes, and therefore also Glivenko-Cantelli classes (Kosorok, 2008).

Therefore, it follows from Lemma A1 that uniformly for $\beta \in \mathcal{B}$ and $t \in [0,\tau]$,

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp(X_i'\beta_x + W_i'\beta_z) \to \quad & EY(t)\exp(X'\beta_x + W'\beta_z) \\ = \quad & E\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)Y(t)\exp(X'\beta_x + Z'\beta_z) \end{aligned}$$

and

$$
\begin{aligned}
& \frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp(X_i'\beta_x + W_i'\beta_z)\begin{pmatrix} X_i \\ \\ W_i - \Sigma\beta_z \end{pmatrix} \\
\to \quad & EY(t)\exp(X'\beta_x + W'\beta_z)\begin{pmatrix} X \\ \\ W - \Sigma\beta_z \end{pmatrix} \\
= \quad & E\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)Y(t)\exp(\xi'\beta)\begin{pmatrix} X \\ \\ Z \end{pmatrix}.
\end{aligned}
$$

Consequently one obtains that

$$
\begin{aligned}
& \frac{1}{n}U_{nA}(\beta) \\
\to \quad & E\int\left[(X_i', W_i')' - \frac{EY(t)\exp(X'\beta_x + W'\beta_z)\Big(X', (W-\Sigma\beta_z)'\Big)'}{EY(t)\exp(X'\beta_x + W'\beta_z)}\right]dN_i(t) \\
= \quad & E\int\left[\xi_i - \frac{EY(t)\exp(X'\beta_x + Z'\beta_z)\xi}{EY(t)\exp(X'\beta_x + Z'\beta_z)}\right]dN_i(t)
\end{aligned}
$$

and thus $n^{-1}U_{nA}(\beta_0) \to 0$ a.s.

Secondly, we show that as $n \to \infty$ uniformly for $\beta \in \mathcal{B}$,

$$
\frac{1}{n}\frac{\partial U_{nA}(\beta)}{\partial \beta} \to -D_A(\beta) \quad \text{a.s.} \tag{3.8}
$$

By direct calculation, one obtains that

$$
\begin{aligned}
&\frac{1}{n}\frac{\partial U_{nA}(\beta)}{\partial\beta} \\
&= \frac{1}{n}\sum_{i=1}^{n}\int\Bigg\{\frac{\sum_{i=1}^{n}Y_i(t)\exp(X_i'\beta_x+W_i'\beta_z)\Big(X_i',(W_i-\Sigma\beta_z)'\Big)'^{\otimes 2}}{\sum_{i=1}^{n}Y_i(t)\exp(X_i'\beta_x+W_i'\beta_z)} \\
&\quad -\left[\frac{\sum_{i=1}^{n}Y_i(t)\exp(X_i'\beta_x+W_i'\beta_z)\Big(X_i',(W_i-\Sigma\beta_z)'\Big)'}{\sum_{i=1}^{n}Y_i(t)\exp(X_i'\beta_x+W_i'\beta_z)}\right]^{\otimes 2} \\
&\quad -\mathrm{diag}(0_{p_1\times p_1},\Sigma)\Bigg\}dN_i(t).
\end{aligned}
$$

It follows from Lemma A1 and similar arguments to (3.7) that uniformly for $\beta\in\mathcal{B}$, as $n\to\infty$

$$
\begin{aligned}
\frac{1}{n}\frac{\partial U_{nA}(\beta)}{\partial\beta} &\to -E\int\left\{\frac{EY(t)\exp(\xi'\beta)\xi^{\otimes 2}}{EY(t)\exp(\xi'\beta)}-\left[\frac{EY(t)\exp(\xi'\beta)\xi}{EY(t)\exp(\xi'\beta)}\right]^{\otimes 2}\right\}dN_i(t) \\
&= -E\int\frac{EY(t)\exp(\xi'\beta)(\xi-\bar{\xi}(\beta;t))^{\otimes 2}}{EY(t)\exp(\xi'\beta)}dN_i(t) \\
&= -D_A(\beta),
\end{aligned}
$$

which is strictly negative definite from (C4a). Combining the above two facts, the equation $U_{nA}(\beta)=0$ has a unique solution and thus $\hat{\beta}_{nA}$ is asymptotically consistent.

Finally, we prove the asymptotic normality. By the Central Limit Theorem, one can obtain that

$$
\sqrt{n}(\hat{\beta}_{nA}-\beta_0) = D_A(\beta_0)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int\left[\xi_{i1}-\frac{EY(t)\exp(\xi_1'\beta_0)\xi_{20}}{EY(t)\exp(\xi_1'\beta_0)}\right]dN_i(t)+o_p(1),
$$

$\xi_{20} = (X', (W - \Sigma\beta_{z0})')'$ and thus

$$n^{1/2}(\hat{\beta}_{nA} - \beta_0) \to N(0, D_A^{-1}(\beta_0)\Sigma_A D_A^{-1}(\beta_0)) \text{ in distribution,}$$

where

$$\Sigma_A = \mathrm{Cov}\biggl(\int \biggl[\xi_{i1} - \frac{EY(t)\exp(\xi_1'\beta_0)\xi_{20}}{EY(t)\exp(\xi_1'\beta_0)}\biggr] dN_i(t)\biggr).$$

$\square$

Proof of Theorem 3.2

Now we show the consistency and asymptotic normality of the proposed estimator $\hat{\gamma}_{nB}$. We first need to cite the i.i.d representation of $\sqrt{n}(\hat{F}_n(t) - F_0(t))$ in Wang et al. (2001). Using the same notations as Wang et al. (2001), we define $G(u) = E[\varphi I(C \geq u)]$, $Q(u) = \int_0^u g(v)d\mu_0(v)$, $R(u) = G(u)\mu_0(u)$ and for $i = 1, \ldots, n$

$$b_i(t) = \sum_{j=1}^{m_i} \biggl[\int_t^\tau \frac{I(T_{ij} \leq u \leq C_i)}{R^2(u)} dQ(u) - \frac{I(t \leq T_{ij} \leq \tau)}{R(T_{ij})}\biggr].$$

Then we have

$$\sqrt{n}(\hat{F}_n(t) - F_0(t)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n F_0(t)b_i(t) + o_p(1). \tag{3.9}$$

Firstly, we show that

$$n^{-1}U_{nB}(\gamma_0) \to 0 \quad \text{a.s.} \tag{3.10}$$

as $n \to \infty$. Note that the following class of functions

$$\bigg\{\bar{\xi}_1 \frac{m}{F(C)} - \bar{\xi}_2 \exp\Big(-\frac{1}{2}\beta_z' \Sigma \beta_z\Big) \exp\big(\bar{\xi}_1' \gamma\big) : \gamma \in \Gamma$$

$$F \text{ is a nondecreasing distribution function in } [0, \tau]\bigg\}$$

is a Glivenko-Cantelli class by similar reasoning in the proof of Theorem 3.1. Therefore,

$$\begin{aligned}\sup_{\gamma \in \Gamma}\bigg| \quad & \frac{1}{n}\sum_{i=1}^{n}\bigg[\bar{\xi}_{i1}\frac{m_i}{\hat{F}_n(C_i)} - \bar{\xi}_{i2}\exp\Big(-\frac{1}{2}\beta_z'\Sigma\beta_z\Big)\exp\big(\bar{\xi}_{i1}'\gamma\big)\bigg] \\ & - E\bigg[\bar{\xi}_1\frac{m}{\hat{F}_n(C_i)} - \bar{\xi}_2\exp\Big(-\frac{1}{2}\beta_z'\Sigma\beta_z\Big)\exp\big(\bar{\xi}_1'\gamma\big)\bigg]\bigg| \to 0, \quad \text{a.s.}\end{aligned}$$

Note that by law of large number, $\hat{F}_n(t) \to F_0(t)$ almost surely in $t \in [0, \tau]$, strenghtened to uniform convergence by Glivenko and Cantelli (1933), we conclude that

$$\sup_{\gamma \in \Gamma}\bigg|\frac{1}{n}U_{nB}(\gamma) - E[\bar{\xi}_1\frac{m}{F_0(C_i)} - \bar{\xi}_2\exp\Big(-\frac{1}{2}\beta_z'\Sigma\beta_z\Big)\exp\big(\bar{\xi}_1'\gamma\big)]\bigg| \to 0, \quad \text{a.s.}.$$

It follows from Lemma A1 that

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{n}U_{nB}(\gamma) - E\left[\bar{\xi}\big(\exp\big(\bar{\xi}'\gamma_0\big) - \exp\big(\bar{\xi}'\gamma\big)\big)\right]\right| \to 0 \quad \text{a.s.,}$$

which immediately yields that $n^{-1}U_{nB}(\gamma_0) \to 0$ a.s..

Secondly, we show that as $n \to \infty$,

$$\frac{1}{n}\frac{\partial U_{nB}(\gamma)}{\partial\gamma} \to -D_B(\gamma) \quad \text{a.s. for } \gamma \in \Gamma. \tag{3.11}$$

By direct calculation, one obtains that

$$\frac{1}{n}\frac{\partial U_{nB}(\gamma)}{\partial\gamma} = -\frac{1}{n}\sum_{i=1}^{n}\exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\big(\bar{\xi}_{i2}'\gamma\big)[\bar{\xi}_{i2}^{\otimes 2} - \operatorname{diag}(0, 0_{p_1\times p_1}, \Sigma)].$$

Note that the following class of functions

$$\left\{\exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\big(\bar{\xi}_2'\gamma\big)[\bar{\xi}_2^{\otimes 2} - \operatorname{diag}(0, 0_{p_1\times p_1}, \Sigma)] : \ \gamma \in \Gamma\right\}$$

is a Glivenko-Cantelli class. Therefore,

$$\sup_{\gamma\in\Gamma}\left|-\frac{1}{n}\frac{\partial U_{nB}(\gamma)}{\partial\gamma} - E\exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\big(\bar{\xi}_2'\gamma\big)[\bar{\xi}_2^{\otimes 2} - \operatorname{diag}(0, 0_{p_1\times p_1}, \Sigma)]\right| \to 0, \quad \text{a.s.}$$

It follows from Lemma A1 that

$$D_B(\gamma) = E\exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\big(\bar{\xi}_2'\gamma\big)[\bar{\xi}_2^{\otimes 2} - \operatorname{diag}(0, 0_{p_1\times p_1}, \Sigma)]$$

and thus (3.11) holds.

Finally, we prove the asymptotic normality by showing that $n^{-1}U_{nB}(\gamma_0)$ can be rewritten as the mean of i.i.d. random variables, plus a term $o_p(n^{-1/2})$, which is

$$\frac{1}{n}U_{nB}(\gamma_0) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\bar{\xi}_{i1}m_i}{F_0(C_i)} - \bar{\xi}_{i2}\exp\left(-\frac{1}{2}\beta_{z0}'\Sigma\beta_{z0}\right)\exp\left(\bar{\xi}_{i1}'\gamma_0\right) - E\frac{\bar{\xi}_1 m b_i(C)}{F_0(C)}\right] + o_p(n^{-1/2}). \tag{3.12}$$

Note that the following class of functions

$$\left\{\bar{\xi}_1\frac{m}{F(C)} - \bar{\xi}_2\exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\left(\bar{\xi}_1'\gamma\right): \quad \gamma \in \Gamma, \right.$$
$$\left. \text{and } F \text{ is a nondecreasing distribution function in } [0,\tau]\right\}$$

is a P-Donsker class (Class notes by Han, 2018). Therefore,

$$\begin{aligned}\sup_{\gamma\in\Gamma}\Bigg| & \frac{1}{n}U_{nB}(\gamma) - E\left[\bar{\xi}_1\frac{m}{\hat{F}_n(C)} - \bar{\xi}_2\exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\left(\bar{\xi}_1'\gamma\right)\right] \\ & - \left\{\frac{1}{n}U_{nB}(\gamma_0) - E\left[\bar{\xi}_1\frac{m}{F_0(C)} - \bar{\xi}_2\exp\left(-\frac{1}{2}\beta_{z0}'\Sigma\beta_{z0}\right)\exp\left(\bar{\xi}_1'\gamma_0\right)\right]\right\}\Bigg| \\ & = o_p(n^{-1/2}).\end{aligned}$$

It follows from Lemma A1 that $E[\bar{\xi}_1\frac{m}{F_0(C)} - \bar{\xi}_2\exp\left(-\frac{1}{2}\beta_{z0}'\Sigma\beta_{z0}\right)\exp\left(\bar{\xi}_1'\gamma_0\right)] = 0$ and from (3.9) that

$$\begin{aligned}E\frac{\bar{\xi}_1 m}{\hat{F}_n(C)} &= E\frac{\bar{\xi}_1 m}{F_0(C)} + E\left[\frac{\bar{\xi}_1 m}{\hat{F}_n(C)} - \frac{\bar{\xi}_1 m}{F_0(C)}\right] \\ &= E\frac{\bar{\xi}_1 m}{F_0(C)} - \frac{1}{n}\sum_{i=1}^{n}E\frac{\bar{\xi}_1 m}{F_0(C)}b_i(C) + o_p(n^{-1/2}).\end{aligned}$$

Combining all the above results and Lemma A1, one obtains that uniformly for $\gamma \in \Gamma$,

$$\begin{aligned}\sup_{\gamma\in\Gamma}\Bigg| & \frac{1}{n}U_{nB}(\gamma) - E\Big[\bar{\xi}\exp\big(\bar{\xi}'\gamma_0\big) - \bar{\xi}\exp\big(\bar{\xi}'\gamma\big)\Big] \\ & -\frac{1}{n}U_{nB}(\gamma_0) + \frac{1}{n}\sum_{i=1}^{n} E\bar{\xi}_1\frac{m}{F_0(C)}b_i(C)\Bigg| \\ & = o_p(n^{-1/2}),\end{aligned}$$

which yields that (3.12) holds.

By similar ideas from Theorem 3.1, one can obtain that

$$\begin{aligned}&\sqrt{n}(\hat{\gamma}_n - \gamma_0) \\ = \;& D_B(\gamma_0)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[\frac{\bar{\xi}_{i1}m_i}{F_0(C_i)} - \bar{\xi}_{i2}\exp\Big(-\frac{1}{2}\beta_{z0}'\Sigma\beta_{z0}\Big)\exp\big(\bar{\xi}_{i1}'\gamma_0\big) - E\frac{\bar{\xi}_1 m}{F_0(C)}b_i(C)\Big] \\ & + o_p(1)\end{aligned}$$

and thus

$$\sqrt{n}(\hat{\gamma}_n - \gamma_0) \rightarrow N(0, D_B^{-1}(\beta_0)\Sigma_B D_B^{-1}(\beta_0)) \;\text{ in distribution,}$$

where

$$\Sigma_B = \mathrm{Cov}\bigg(\Big[\bar{\xi}_{i1}\frac{m_i}{F_0(C_i)} - \bar{\xi}_{i2}\exp\Big(-\frac{1}{2}\beta_{z0}'\Sigma\beta_{z0}\Big)\exp\big(\bar{\xi}_{i1}'\gamma_0\big) - E\bar{\xi}_1\frac{m}{F_0(C)}b_i(C)\Big]\bigg).$$

$\square$

Detailed Derivations for Lemma A1, Theorems 3.1 and 3.2.

Various expectations need to be calculated in order to derive the three equalities in Lemma A1. We will present them one by one in the following.

Since

$$E[\exp(e_i'\beta_z)] = \exp\left(\beta_z\mu + \frac{1}{2}\beta_z'\Sigma\beta_z\right) = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right),$$

we get (i):

$$\begin{aligned} E[\exp(W_i'\beta_z)|O_i, Z_i] &= \exp(Z_i'\beta_z) \cdot E[\exp(e_i'\beta_z)] \\ &= \exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right). \end{aligned}$$

For (ii), because

$$\begin{aligned} E[e_i\exp(e_i'\beta_z)] &= \int e_i\exp(e_i'\beta_z) \cdot \frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}} \cdot \exp\left(-\frac{1}{2}e_i'\Sigma^{-1}e_i\right)de_i \\ &= \int e_i\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\exp\left[-\frac{(e_i-\Sigma\beta_z)^{\otimes 2}}{2\Sigma} + \frac{(\Sigma\beta_z)^{\otimes 2}}{2\Sigma}\right]de_i \\ &= \exp\left[\frac{1}{2}\beta_z'\Sigma\beta_z\right] \cdot \Sigma\beta_z, \end{aligned}$$

we obtain (ii) as

$$
\begin{aligned}
&E[(W_i - \Sigma\beta_z)\exp(W_i'\beta_z)|O_i, Z_i] \\
&= E[W_i \exp(W_i'\beta_z)|O_i, Z_i] - \Sigma\beta_z E[\exp(W_i'\beta_z)|O_i, Z_i] \\
&= E[Z_i \exp(W_i'\beta_z)|O_i, Z_i] + E[e_i \exp(W_i'\beta_z)|O_i, Z_i] - \Sigma\beta_z \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&= Z_i \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) + \exp(Z_i\beta_z)E[e_i\exp(e_i'\beta_z)] \\
&\quad -\Sigma\beta_z \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&= Z_i \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) + \Sigma\beta_z \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&\quad -\Sigma\beta_z \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&= Z_i \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z).
\end{aligned}
$$

In order to derive (iii), we need the following expectations:

$$
\begin{aligned}
E[e_i^{\otimes 2}\exp(e_i'\beta_z)] &= \int e_i^{\otimes 2}\exp(e_i'\beta_z)\cdot\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\cdot\exp\left(-\frac{1}{2}e_i'\Sigma^{-1}e_i\right)de_i \\
&= \int e_i^{\otimes 2}\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\exp\left[-\frac{(e_i-\Sigma\beta_z)^{\otimes 2}}{2\Sigma}+\frac{(\Sigma\beta_z)^{\otimes 2}}{2\Sigma}\right]de_i \\
&= \exp\left[\frac{1}{2}\beta_z'\Sigma\beta_z\right]\cdot(\Sigma+(\Sigma\beta^2)^{\otimes 2}),
\end{aligned}
$$

and

$$
\begin{aligned}
&E[(Z_i^{\otimes 2} + 2Z_i'e_i + e_i^{\otimes 2})\exp(W_i'\beta_z)|O_i, Z_i] \\
&= Z_i^{\otimes 2}\exp(Z_i'\beta_z)E[\exp(e_i'\beta_z)] + 2Z_i\exp(Z_i'\beta_z)E[e_i\exp(e_i'\beta_z)] \\
&\quad + \exp(Z_i\beta_z)E[e_i^{\otimes 2}\exp(e_i'\beta_z)] \\
&= Z_i^{\otimes 2}\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) + 2Z_i\exp(Z_i'\beta_z)\Sigma\beta_z\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) \\
&\quad + \exp(Z_i\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)(\Sigma + (\Sigma\beta_z)^{\otimes 2}) \\
&= (Z_i^{\otimes 2} + \Sigma)\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) + 2Z_i\Sigma\beta_z\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) \\
&\quad +(\Sigma\beta_z)^{\otimes 2}\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right).
\end{aligned}
$$

Then, we can obtain (iii)

$$
\begin{aligned}
&E[(W_i - \Sigma\beta_z)^{\otimes 2}\exp(W_i'\beta_z)|O_i, Z_i] \\
&= E[(W_i^{\otimes 2} - 2W_i'\Sigma\beta_z + \Sigma^{\otimes 2}\beta_z^{\otimes 2})\exp(W_i'\beta_z)|O_i, Z_i] \\
&= E[(Z_i^{\otimes 2} + 2Z_i'e_i + e_i^{\otimes 2})\exp(W_i'\beta_z)|O_i, Z_i] - 2\Sigma\beta_z E[W_i\exp(W_i'\beta_z)|O_i, Z_i] \\
&\quad +(\Sigma\beta_z)^{\otimes 2}E[\exp(W_i'\beta_z)|O_i, Z_i] \\
&= E[(Z_i^{\otimes 2} + 2Z_i'e_i + e_i^{\otimes 2})\exp(W_i'\beta_z)|O_i, Z_i] - 2\Sigma\beta_z(Z_i + \Sigma\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&\quad +(\Sigma\beta_z)^{\otimes 2}\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) \\
&= E[(Z_i^{\otimes 2} + 2Z_i'e_i + e_i^{\otimes 2})\exp(W_i'\beta_z)|O_i, Z_i] - 2\Sigma\beta_z'Z_i\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z) \\
&\quad -(\Sigma\beta_z)^{\otimes 2}\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) \\
&= (Z_iZ_i' + \Sigma)\exp(Z_i'\beta_z)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right).
\end{aligned}
$$

Next, we prove that $n^{-1}U_{nA}(\beta_0) \to 0$ a.s. as $n \to \infty$. Similar derivations can be applied for Theorem 3.2. Referring to the formula

$$n^{-1}U_{nA}(\beta) = n^{-1}\sum_{i=1}^{n}\int\left[\xi_{i1} - \frac{\sum_{j=1}^{n} Y_j(t)\exp\left(\xi_{j1}'\beta\right)\xi_{j2}}{\sum_{j=1}^{n} Y_j(t)\exp\left(\xi_{j1}'\beta\right)}\right]dN_i(t),$$

we separately derive limits for each of the component.

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp(X_i'\beta_x + W_i'\beta_z) \\
&\quad\to\quad E[Y(t)\exp(X'\beta_x + W'\beta_z)] \\
&\quad=\quad E[Y(t)\exp(X'\beta_x)\exp(W'\beta_z)] \\
&\quad=\quad E\left[Y(t)\exp(X'\beta_x)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z'\beta_z)\right] \\
&\quad=\quad \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)E\left[Y(t)\exp(X'\beta_x + Z'\beta_z)\right]
\end{aligned}$$

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp(X_i'\beta_x + W_i'\beta_z)\begin{pmatrix} X_i \\ \\ W_i - \Sigma\beta_z \end{pmatrix} \\
&\quad\to\quad E\left[Y(t)\exp(X'\beta_x + W'\beta_z)\begin{pmatrix} X \\ \\ W - \Sigma\beta_z \end{pmatrix}\right] \\
&\quad=\quad E\left[\begin{pmatrix} Y(t)\exp(X'\beta_x)\exp(W'\beta_z)X \\ \\ Y(t)\exp(X'\beta_x)\exp(W'\beta_z)(W - \Sigma\beta_z) \end{pmatrix}\right] \\
&\quad=\quad E\left[\begin{pmatrix} Y(t)\exp(X'\beta_x)\exp(Z'\beta_z)\exp\left(\tfrac{1}{2}\beta_z'\Sigma\beta_z\right)X \\ \\ Y(t)\exp(X'\beta_x)\exp(Z'\beta_z)\exp\left(\tfrac{1}{2}\beta_z'\Sigma\beta_z\right)Z \end{pmatrix}\right] \\
&\quad=\quad \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)E\left[Y(t)\exp(X'\beta_x + Z'\beta_z)\begin{pmatrix} X \\ \\ Z \end{pmatrix}\right]
\end{aligned}
$$

Then, we get

$$
\begin{aligned}
&\frac{1}{n}U_{nA}(\beta) \\
\to\ & E\int\left[\begin{pmatrix} X \\ W \end{pmatrix} - \frac{E\left[Y(t)\exp(X'\beta_x + W'\beta_z)\begin{pmatrix} X \\ W - \Sigma\beta_z \end{pmatrix}\right]}{E[Y(t)\exp(X'\beta_x + W'\beta_z)]}\right] dN(t) \\
=\ & E\int\left[\begin{pmatrix} X \\ W \end{pmatrix} - \frac{\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)E\left[Y(t)\exp(X'\beta_x + Z'\beta_z)\begin{pmatrix} X \\ Z \end{pmatrix}\right]}{\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)E\left[Y(t)\exp(X'\beta_x + Z'\beta_z)\right]}\right] dN(t) \\
=\ & E\int\left[\begin{pmatrix} X \\ W \end{pmatrix} - \frac{E\left[Y(t)\exp(X'\beta_x + Z'\beta_z)\begin{pmatrix} X \\ Z \end{pmatrix}\right]}{E\left[Y(t)\exp(X'\beta_x + Z'\beta_z)\right]}\right] dN(t) \\
\to\ & 0
\end{aligned}
$$

$\square$

Chapter 4: Regression Analysis of Panel Count Data with Measurement Errors

4.1 Introduction

In this chapter, we will be analyzing the second data type for recurrent event processes, panel count data.

Panel count data are frequently encountered in long-term studies that concern the occurrence rate of a recurrent event, such as prospective cohort studies, population-based epidemiological studies, tumorigenicity experiments, etc.. In such studies, study subjects are examined or observed only at discrete time points (Kalbfleisch and Lawless, 1985; Sun and Zhao, 2013). The observation process includes a sequence of discrete observation times; and the observation times may or may not be independent of the underlying point process generating the observed panel count data. As indicated by Andersen et al. (1993), it is common and convenient to characterize the occurrences of recurrent events by counting processes and to model the intensity function of the counting process. However, for anlyzing panel count data, it is usually more convenient to work directly on the mean function of the counting processes.

Regression analysis of panel count data with accurate recording of all the covariates have been discussed by many researchers. Among others, Sun and Wei (2000) proposed regression analysis methods on estimating the mean functions for panel count data with both independent and dependent observation processes; and the asymptotic properties for both cases were established as well. However, in many applications, some of the covariates may be recorded with some measurenment errors, and it is inappropriate to simply assume those measurement errors are ignorable.

Kim (2007) studied a measurement error model in which the continuous covariates may be measured with errors. He used a pseudo-likelihood and iterative procedure for the estimation procedure. Our proposed method will directly use an estimating equation based approach and the procedure is easy to implement.

The remainder of this chapter is organized as follows. Section 4.2 introduces the notations and models for both independent and dependent observation processes as well as their estimating procedures. In Section 4.3, simulation studies are conducted to evaluate the performance of the proposed methods. We apply our methods to analyze panel count data from a bladder cancer study in Section 4.4. Conclusions and discussions are given in Section 4.5.

4.2 Notation and Models

Consider a study involving n subjects who may experience recurrent events. All the subjects are only monitored occasionally. Denote X_i and Z_i as the p_1 and

p_2 dimensional vectors of covariates, C_i as the censoring time for $i = 1, \ldots, n$. Let $N_i^*(t)$ be the cumulative number of events that have occurred before time t for $i = 1, \ldots, n$ and $0 \leq t \leq \tau$, where τ denotes the end of the study. For subjects who are only observed discretely, denote the observation time points for the ith subject as $T_{i,j}, j = 1, ..., m_i$, where m_i is the total number of observation times. Let the number of observation times before time t be $H_i(t) = \sum_{j=1}^{m_i} I(T_{i,j} \leq \min(t, C_i))$ where $I(\cdot)$ is an indicator function. Suppose that X_i's are completely observed and Z_i's are not available, but we can observe the corresponding surrogate W_i's such that

$$W_i = Z_i + \epsilon_i, \tag{4.1}$$

where $\epsilon_i \sim N(0, \Sigma)$ with a known covariance matrix Σ. Then the observed data consist of $\{O_i, W_i,\ i = 1, \ldots, n\}$, where

$$O_i = (T_{i,1}, \ldots, T_{i,m_i}, m_i, C_i, N_i(T_{i,1}), \ldots, N_i(T_{i,m_i}), X_i)$$

and $N_i(t) = N_i^*(\min(t, C_i))$.

Following Sun and Wei (2000) and He et al. (2008), we assume that the conditional mean of $N_i^*(t)$ given X_i and Z_i takes the form

$$E(N_i^*(t)|X_i, Z_i) = \mu_0(t) \exp(\beta_{x0}' X_i + \beta_{z0}' Z_i)\,, \tag{4.2}$$

where $(\beta_{x0}', \beta_{z0}')'$ is a vector of regression parameters of interests, and $\mu_0(t)$ is a completely unspecified baseline mean function.

4.2.1 Independent Observation Process

In this subsection, we discuss the estimation of $(\beta'_{x0}, \beta'_{z0})'$ when the observational processes $H_i(\cdot)$ are independent of X_i, Z_i and $N_i^*(\cdot)$. For $i = 1, \ldots, n$, we define

$$\bar{N}_i = \int_0^\tau I(C_i \geq t) N_i(t) dH_i(t) = \sum_{j=1}^{m_i} N_i(T_{i,j}) I(T_{i,j} \leq C_i).$$

Following Sun and Wei (2000) and He et al. (2008), one obtains that

$$E[\bar{N}_i | X_i, Z_i] = \exp(X_i' \beta_{x0} + Z_i' \beta_{z0} + \xi_0), \tag{4.3}$$

where $\xi_0 = \log \int_0^\tau EI(C_i \geq t) \mu_0(t) dEH_i(t)$, which results in the estimating equation

$$U_n^*(\beta) = \sum_{i=1}^n V_i [\bar{N}_i - \exp(X_i' \beta_x + Z_i' \beta_z + \xi)], \tag{4.4}$$

where $V_i = (X_i', Z_i', 1)'$.

When the covariates Z_i's are completely observed, the estimating equation (4.4) can be used for the estimation of β and the asymptotic properties can be derived (He et al., 2008). By replacing Z_i's with the observed surrogate W_i's, the resulting estimator would be biased.

Let 0_d denote a $d-$ dimensional zero vector and let $0_{d \times d}$ denote $d \times d$ zero matrix. For a vector a, $a^{\otimes 2} = aa'$. Following Lemma B1 in Section 4.6.2, we propose the following estimating equation

$$U_{n1}(\beta) = \sum_{i=1}^{n} \left[\bar{N}_i V_{i1} - (V_{i1} - V_{i2}) \exp\left(-\frac{1}{2} \beta_z' \Sigma \beta_z \right) \exp(V_{i1}' \beta) \right], \tag{4.5}$$

where $V_{i1} = (X_i', W_i', 1)'$, $V_{i2} = (0_{p_1}', \beta_z' \Sigma, 0)'$, $\beta = (\beta_x', \beta_z', \xi)'$.

Denote $\hat{\beta}_{n1}$ as the solution to $U_{n1}(\beta) = 0$. We anticipate that this estimator is consistent and asymptotically normally distributed, which is supported by the empirical results in Section 4.3. The bootstrap approach will be used for the variance estimation procedure. Detailed derivation for the estimation of the variance is still under investigation.

4.2.2 Dependent Observation Process

In this subsection, we discuss the estimation of $(\beta_{x0}', \beta_{z0}')'$ when the observation processes H_i are conditionally independent of $N_i^*(\cdot)$ given X_i, Z_i. For $i = 1, \ldots, n$, we assume that the conditional mean of $H_i(\cdot)$ takes the form

$$E[H_i(t) | X_i, Z_i] = \lambda_0(t) \exp(X_i' \gamma_{x0} + Z_i' \gamma_{z0}), \tag{4.6}$$

where $\lambda_0(t)$ is a completely unspecified baseline function. Then, the estimate $\hat{\gamma} = (\hat{\gamma}_x', \hat{\gamma}_z')'$ of $\gamma = (\gamma_{x0}', \gamma_{z0}')'$ can be obtained by using the estimating equation (3.5)

$$U_{n\gamma}(\gamma) = \sum_{i=1}^{n} \int \left[V_{i0} - \frac{\sum_{j=1}^{n} Y_j(t) \exp\left(V_{j0}' \gamma\right) (V_{j0} - (0_{p_1}', \gamma_z' \Sigma)')}{\sum_{j=1}^{n} Y_j(t) \exp\left(V_{j0}' \gamma\right)} \right] dN_i(t),$$

where $V_{i0} = (X_i', W_i')'$.

Next, in subsection 4.2.1, we have obtained that

$$E[\bar{N}_i|X_i, Z_i] = \exp(X_i'\beta_{x0} + Z_i'\beta_{z0} + \xi_0),$$

where $\xi_0 = \log \int_0^\tau EI(C_i \geq t)\mu_0(t)dEH_i(t)$. Since the observation processes H_i's are no longer independent of $N_i^*(\cdot)$, we have

$$\begin{aligned} E[\bar{N}_i(t)|X_i, Z_i] &= \int_0^\tau P(t \leq C_i)\mu_0(t)\exp(X_i\beta_{x0} + Z_i\beta_{z0})dEH_i(t) \\ &= \exp[X_i(\gamma_{x0} + \beta_{x0}) + Z_i(\gamma_{z0} + \beta_{z0}) + \xi_{01}], \end{aligned} \tag{4.7}$$

where $\xi_{01} = \log \int_0^\tau P(t \leq C_i)\mu_0(t)\lambda_0(t)dt$. By plugging $\hat{\gamma} = (\hat{\gamma}_x', \hat{\gamma}_z')'$, the estimating equation is

$$U_n^*(\beta) = \sum_{i=1}^n V_i[\bar{N}_i - \exp(X_i'(\hat{\gamma}_x + \beta_x) + Z_i'(\hat{\gamma}_z + \beta_z) + \xi)], \tag{4.8}$$

where $V_i = (X_i', Z_i', 1)'$.

When the covariates Z_i's are completely observed, the estimating equation (4.8) can be used for the estimation of β and the asymptotic properties were established in Sun and Wei (2000), using standard procedures. Replacing the Z_i's with the observed surrogate W_i's directly would result in a biased estimator.

Following similar ideas to those in Section 4.2.1 and Lemma B1 in Section 4.6.2, when the covariates Z_i's are not observed, we propose the following estimating equation

$$U_{n2}(\beta) = \sum_{i=1}^{n} \left[\bar{N}_i V_{i1} - (V_{i1} - V_{i2}) \exp\left(-\frac{1}{2} (\hat{\gamma}_z + \beta_z)' \Sigma (\hat{\gamma}_z + \beta_z) \right) \exp(V_{i0}' \hat{\gamma}) \exp(V_{i1}' \beta) \right], \tag{4.9}$$

where $V_{i1} = (X_i', W_i', 1)'$, $V_{i2} = (0_{p_1}', (\hat{\gamma}_z + \beta_z)'\Sigma, 0)'$ and $\beta = (\beta_x', \beta_z', \xi)'$.

Denote $\hat{\beta}_{n2}$ as the solution to $U_{n2}(\beta) = 0$. Again, this estimator seems consistent and asymptotically normally distributed based on simulation studies in Section 4.3. Bootstrap approach will be used for the variance estimation procedure. Detailed derivation for the variance estimation is still under development.

Similar approach for the estimation of Σ in Chapter 3 will be applied.

4.3 Simulation

Simulation studies were conducted to evaluate the performance of both proposed methods under various situations while the main focus is the estimation of γ's and β's with presence of measurement errors in some of the covariates. We asume that the covariate without measurement errors X_i's follow a Bernoulli distribution with success probability 0.5; and the covariate with measurement errors W_i be $Z_i + e_i$, where Z_i's follow a normal distribution with mean 0 and variance 1; and e_i's follow a normal distribution with mean 0 and variance $\Sigma = \sigma^2$. The censoring times C_i's are generated from a uniform distribution over $(\tau/3, \tau)$, where $\tau = 1$. The observation times $(T_{i,1}, \ldots, T_{i,m_i})$ are defined as

$$T_{i,j+1} = T_{i,j} - \log(U_{i,j+1})\{\exp(V_{i0}'\gamma) d\lambda_0\}^{-1},$$

censored by C_i, where $T_{i,0} = 0, U_{i,j} \sim U(0,1)$, and m_i is the number of observation times for subject i, $i = 1, \ldots, n$.

Given m_i and $(T_{i,1}, T_{i,2}, \ldots, T_{i,m_i})$, we generate panel count data $N_i(T_{i,j})$ from the mixed Poisson processes as

$$N_i(T_{i,j}) = N_i^*[\mu(T_{i,1})] + N_i^*[\mu(T_{i,2}) - \mu(T_{i,1})] + \cdots + N_i^*[\mu(T_{i,j}) - \mu(T_{i,j-1})],$$

$j = 1, \ldots, m_i, i = 1, \ldots, n$. In the above, $N_i^*[\mu(t)]$ denotes the random number generated from the Poisson distribution with mean $\exp(V_{i1}\beta_0)$, and $\mu_0(t) = \theta t$, where $\theta = 1$ is applied in the simulation. The results given below are based on $n = 200$ and 500, $\sigma = 0.1, 0.3, 0.5$ for the independent case; and $n = 500$, $\sigma = 0.05, 0.15, 0.25$ for the dependent case, running 500 replications. For detailed tables with more σ values, please refer to Section 4.6.1.

Table 4.1: Simulation Results for Independent Observation Processes, n=200

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.1	Proposed	(1,1)	(-0.78%, **-0.62%**)	(0.179, 0.109)	(0.194, 0.129)	(0.94, 0.92)
		(1,-1)	(-0.34%, **-0.38%**)	(0.179, 0.110)	(0.191, 0.128)	(0.92, 0.90)
		(-1,1)	(-0.16%, **0.14%**)	(0.188, 0.120)	(0.193, 0.130)	(0.94, 0.94)
	Sun and Wei (2000)	(1,1)	(-1.52%, **-2.66%**)	(0.179, 0.104)	(0.190, 0.124)	(0.94, 0.86)
		(1,-1)	(-0.58%, **-1.47%**)	(0.179, 0.106)	(0.202, 0.126)	(0.92, 0.87)
		(-1,1)	(1.14%, **-2.78%**)	(0.186, 0.109)	(0.204, 0.121)	(0.92, 0.88)
0.3	Proposed	(1,1)	(1.51%, **1.68%**)	(0.190, 0.124)	(0.213, 0.138)	(0.93, 0.92)
		(1,-1)	(0.96%, **1.12%**)	(0.194, 0.124)	(0.217, 0.149)	(0.94, 0.88)
		(-1,1)	(1.40%, **1.33%**)	(0.199, 0.129)	(0.207, 0.151)	(0.93, 0.91)
	Sun and Wei (2000)	(1,1)	(-0.69%, **-9.17%**)	(0.187, 0.103)	(0.208, 0.120)	(0.92, 0.75)
		(1,-1)	(0.98%, **-8.89%**)	(0.186, 0.103)	(0.208, 0.120)	(0.91, 0.77)
		(-1,1)	(-1.47%, **-9.01%**)	(0.196, 0.109)	(0.214, 0.126)	(0.92, 0.78)
0.5	Proposed	(1,1)	(1.30%, **3.80%**)	(0.232, 0.204)	(0.242, 0.202)	(0.95, 0.88)
		(1,-1)	(-0.19%, **1.82%**)	(0.227, 0.197)	(0.226, 0.189)	(0.95, 0.90)
		(-1,1)	(1.40%, **4.09%**)	(0.236, 0.203)	(0.247, 0.198)	(0.95, 0.87)
	Sun and Wei (2000)	(1,1)	(-0.23%, **-21.08%**)	(0.197, 0.098)	(0.216, 0.111)	(0.94, 0.43)
		(1,-1)	(0.30%, **-20.67%**)	(0.198, 0.102)	(0.208, 0.123)	(0.94, 0.41)
		(-1,1)	(0.99%, **-20.40%**)	(0.209, 0.106)	(0.232, 0.140)	(0.92, 0.45)

Table 4.2: Simulation Results for Independent Observation Processes, n=500

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.1	Proposed	(1,1)	(0.15%, **-0.68%**)	(0.117, 0.075)	(0.123, 0.089)	(0.94, 0.90)
		(1,-1)	(0.21%, **-0.40%**)	(0.119, 0.073)	(0.118, 0.083)	(0.95, 0.91)
		(-1,1)	(-0.88%, **0.10%**)	(0.124, 0.078)	(0.127, 0.083)	(0.94, 0.92)
	Sun and Wei (2000)	(1,1)	(0.46%, **-1.36%**)	(0.085, 0.055)	(0.087, 0.060)	(0.94, 0.89)
		(1,-1)	(0.44%, **-1.74%**)	(0.086, 0.054)	(0.090, 0.061)	(0.94, 0.90)
		(-1,1)	(-0.13%, **-1.33%**)	(0.089, 0.055)	(0.096, 0.065)	(0.94, 0.90)
0.3	Proposed	(1,1)	(0.31%, **0.20%**)	(0.091, 0.063)	(0.091, 0.071)	(0.95, 0.92)
		(1,-1)	(0.96%, **-0.35%**)	(0.091, 0.063)	(0.096, 0.070)	(0.94, 0.92)
		(-1,1)	(-0.33%, **-0.07%**)	(0.095, 0.064)	(0.096, 0.072)	(0.95, 0.90)
	Sun and Wei (2000)	(1,1)	(0.31%, **-8.24%**)	(0.090, 0.055)	(0.092, 0.068)	(0.95, 0.58)
		(1,-1)	(-0.03%, **-9.01%**)	(0.090, 0.055)	(0.098, 0.063)	(0.92, 0.54)
		(-1,1)	(0.57%, **-8.45%**)	(0.093, 0.056)	(0.098, 0.064)	(0.94, 0.57)
0.5	Proposed	(1,1)	(0.33%, **0.63%**)	(0.142, 0.108)	(0.147, 0.120)	(0.95, 0.91)
		(1,-1)	(0.67%, **0.55%**)	(0.102, 0.078)	(0.106, 0.085)	(0.93, 0.93)
		(-1,1)	(0.24%, **0.79%**)	(0.105, 0.079)	(0.109, 0.089)	(0.95, 0.93)
	Sun and Wei (2000)	(1,1)	(-0.07%, **-20.45%**)	(0.096, 0.053)	(0.099, 0.056)	(0.93, 0.10)
		(1,-1)	(-0.05%, **-19.77%**)	(0.096, 0.055)	(0.099, 0.063)	(0.94, 0.12)
		(-1,1)	(-0.26%, **-20.45%**)	(0.099, 0.054)	(0.105, 0.060)	(0.94, 0.10)

Tables 4.1 and 4.2 present the simulation results for independent observation processes for $n = 200$ and $n = 500$, respectively. The two tables include the estimated relative bias (Relative Bias), the average of the bootstrap sample standard deviations (BSE), the sample standard deviation of $\hat{\beta}$ (SSE), and the empirical 95% coverage probability (CP) for β. It can be seen that when the sample size is small, the point estimates are not as stable as those when the sample size is large. However, our estimators are still better than the one in Sun and Wei (2000) by ignoring the measurement errors. Moreover, the BSE's are close to the SSE's, suggesting that the bootstrap variance estimation is reasonable.

Figure 4.1 displays the Normal Q-Q plot of $\hat{\beta}_z$ for the setting of $n = 200, \sigma = 0.3$, which indicates that our estimators approaximately follow a normal distribution. The plots for other setups are similar.

Table 4.3 shows the simulation results regarding γ. Our γ estimate is stable

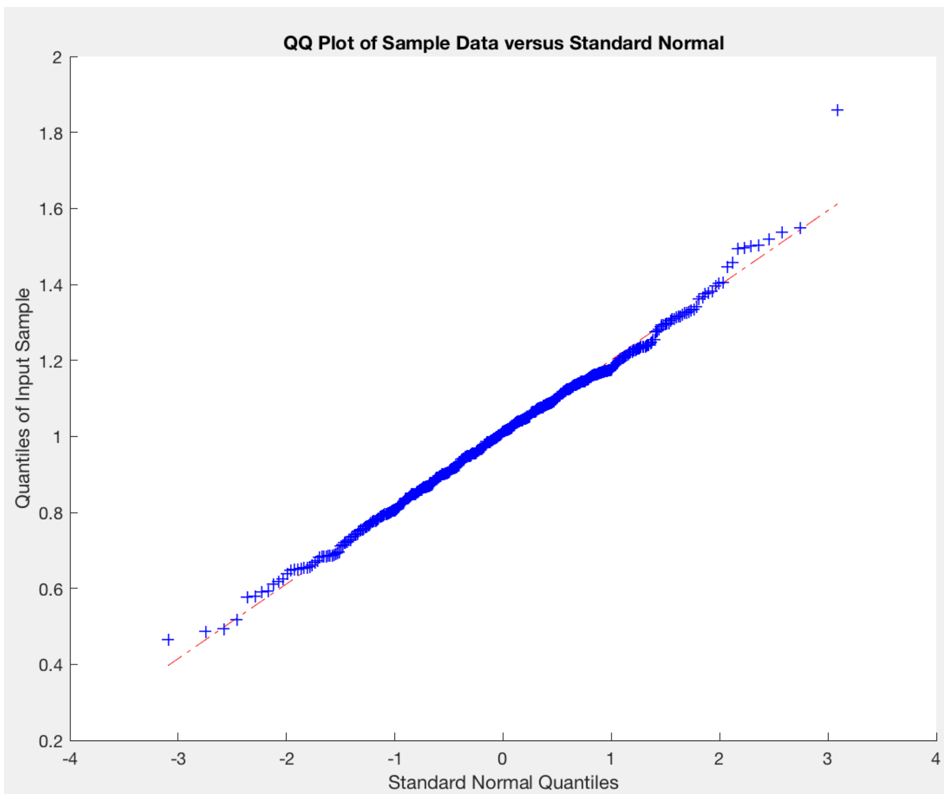

Figure 4.1: Independent Case: QQ Plot for $\hat{\beta}_z$ with $n = 200, \sigma = 0.3$

Table 4.3: Simulation Results of γ for Dependent Observation Processes, $n = 500$

σ	Methods	γ	Relative Bias	SSE	SSE	CP
0.05	Proposed	(1,1)	(0.69%, **0.02%**)	(0.070, 0.031)	(0.069, 0.033)	(0.95, 0.92)
		(1,-1)	(-0.66%, **-0.74%**)	(0.071, 0.032)	(0.073, 0.031)	(0.94, 0.94)
	Sun and Wei (2000)	(1,1)	(-0.68%, **-1.19%**)	(0.071, 0.032)	(0.073, 0.033)	(0.94, 0.91)
		(1,-1)	(0.70%, **-1.13%**)	(0.071, 0.032)	(0.071, 0.032)	(0.95, 0.93)
0.15	Proposed	(1,1)	(-0.19%, **-0.70%**)	(0.073, 0.033)	(0.073, 0.038)	(0.95, 0.90)
		(1,-1)	(-0.49%, **-0.62%**)	(0.074, 0.033)	(0.076, 0.038)	(0.93, 0.91)
	Sun and Wei (2000)	(1,1)	(-0.20%, **-3.28%**)	(0.073, 0.033)	(0.075, 0.036)	(0.94, 0.78)
		(1,-1)	(-0.44%, **-2.93%**)	(0.073, 0.033)	(0.074, 0.036)	(0.96, 0.80)
0.25	Proposed	(1,1)	(-0.48%, **-0.41%**)	(0.079, 0.036)	(0.078, 0.045)	(0.95, 0.91)
		(1,-1)	(-0.18%, **-0.10%**)	(0.078, 0.036)	(0.082, 0.041)	(0.94, 0.90)
	Sun and Wei (2000)	(1,1)	(-0.67%, **-6.55%**)	(0.078, 0.035)	(0.075, 0.038)	(0.96, 0.51)
		(1,-1)	(-1.26%, **-6.55%**)	(0.078, 0.034)	(0.079, 0.039)	(0.94, 0.52)

based on the simulation results. For the estimation of β with dependent observation processes, Tables 4.4 and 4.5 show the simulation results for $\gamma = (1, 1)$ and $\gamma = (1, -1)$, respectively. We applied bootstrap to estimate the standard errors. As σ gets larger, the estimation of β_z gets more off, which is reasonable since the estimation of β is a two-step procedure using the estimated γ.

Table 4.4: Simulation Results of β for $\gamma = (1, 1)$, $n = 500$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.05	Proposed	(1,1)	(0.99%, **-0.66%**)	(0.206, 0.131)	(0.206, 0.156)	(0.93, 0.93)
		(1,-1)	(1.23%, **-0.84%**)	(0.188, 0.105)	(0.174, 0.105)	(0.95, 0.96)
		(-1,1)	(-0.66%, **0.39%**)	(0.213, 0.130)	(0.241, 0.157)	(0.93, 0.92)
	Sun and Wei (2000)	(1,1)	(0.47%, **-1.98%**)	(0.206, 0.130)	(0.236, 0.164)	(0.92, 0.85)
		(1,-1)	(1.09%, **-1.23%**)	(0.187, 0.104)	(0.183, 0.111)	(0.95, 0.91)
		(-1,1)	(0.20%, **-1.06%**)	(0.215, 0.131)	(0.228, 0.168)	(0.93, 0.84)
0.15	Proposed	(1,1)	(0.48%, **1.04%**)	(0.225, 0.152)	(0.234, 0.176)	(0.96, 0.90)
		(1,-1)	(-0.66%, **-0.73%**)	(0.186, 0.106)	(0.176, 0.109)	(0.97, 0.93)
		(-1,1)	(-1.07%, **-1.22%**)	(0.230, 0.151)	(0.255, 0.171)	(0.92, 0.92)
	Sun and Wei (2000)	(1,1)	(-0.89%, **-3.67%**)	(0.215, 0.137)	(0.236, 0.175)	(0.93, 0.83)
		(1,-1)	(1.27%, **-3.46%**)	(0.185, 0.102)	(0.172, 0.108)	(0.96, 0.82)
		(-1,1)	(-0.87%, **-2.91%**)	(0.226, 0.136)	(0.249, 0.170)	(0.94, 0.84)
0.25	Proposed	(1,1)	(3.01%, **6.15%**)	(0.275, 0.218)	(0.288, 0.232)	(0.97, 0.94)
		(1,-1)	(1.47%, **-1.24%**)	(0.188, 0.107)	(0.178, 0.112)	(0.95, 0.94)
		(-1,1)	(0.72%, **4.72%**)	(0.271, 0.202)	(0.275, 0.235)	(0.94, 0.92)
	Sun and Wei (2000)	(1,1)	(0.77%, **-8.98%**)	(0.239, 0.148)	(0.258, 0.189)	(0.95, 0.80)
		(1,-1)	(1.63%, **-6.89%**)	(0.167, 0.100)	(0.171, 0.110)	(0.96, 0.84)
		(-1,1)	(0.26%, **-8.88%**)	(0.250, 0.148)	(0.267, 0.166)	(0.96, 0.81)

The performance of the normal approximation of $\hat{\beta}_z$ for the dependent obser-

Table 4.5: Simulation Results of β for $\gamma = (1, -1)$, $n = 500$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.05	Proposed	(1,1)	(1.35%, **-0.51%**)	(0.188, 0.104)	(0.175, 0.104)	(0.95, 0.95)
		(1,-1)	(1.07%, **0.15%**)	(0.207, 0.130)	(0.214, 0.135)	(0.93, 0.94)
		(-1,1)	(-0.98%, **0.57%**)	(0.227, 0.133)	(0.216, 0.142)	(0.96, 0.94)
	Sun and Wei (2000)	(1,1)	(1.19%, **-1.72%**)	(0.187, 0.103)	(0.176, 0.103)	(0.95, 0.92)
		(1,-1)	(-0.07%, **-2.31%**)	(0.204, 0.127)	(0.234, 0.160)	(0.93, 0.86)
		(-1,1)	(-0.06%, **-1.63%**)	(0.228, 0.134)	(0.235, 0.145)	(0.95, 0.90)
0.15	Proposed	(1,1)	(1.04%, **1.55%**)	(0.189, 0.105)	(0.175, 0.110)	(0.96, 0.92)
		(1,-1)	(1.24%, **1.89%**)	(0.217, 0.142)	(0.234, 0.173)	(0.95, 0.89)
		(-1,1)	(0.08%, **2.03%**)	(0.230, 0.135)	(0.231, 0.144)	(0.94, 0.93)
	Sun and Wei (2000)	(1,1)	(2.91%, **-3.57%**)	(0.184, 0.100)	(0.179, 0.110)	(0.95, 0.87)
		(1,-1)	(2.53%, **-3.20%**)	(0.218, 0.134)	(0.250, 0.167)	(0.93, 0.86)
		(-1,1)	(-0.19%, **-2.99%**)	(0.228, 0.133)	(0.228, 0.138)	(0.95, 0.90)
0.25	Proposed	(1,1)	(0.94%, **6.25%**)	(0.188, 0.107)	(0.172, 0.120)	(0.97, 0.86)
		(1,-1)	(2.85%, **9.67%**)	(0.210, 0.163)	(0.259, 0.181)	(0.94, 0.76)
		(-1,1)	(1.65%, **9.15%**)	(0.230, 0.183)	(0.231, 0.127)	(0.95, 0.83)
	Sun and Wei (2000)	(1,1)	(0.94%, **-10.68%**)	(0.186, 0.098)	(0.180, 0.105)	(0.95, 0.74)
		(1,-1)	(0.74%, **-12.05%**)	(0.243, 0.152)	(0.266, 0.180)	(0.95, 0.73)
		(-1,1)	(-0.61%, **-10.03%**)	(0.227, 0.127)	(0.234, 0.131)	(0.94, 0.71)

vation processes is shown in Figure 4.2 where the true value of β_z equals to 1. The plot indicates that our estimator approximately follows a normal distribution.

4.4 An Example

We apply the proposed estimation procedure to a bladder cancer study conducted by the Veterans Administration Cooperative Urological Reserach Group (Byar et al., 1977; Byar, 1980; Sun and Wei, 2000). This is a follow-up study for patients with superficial bladder tumors. At the beginning of the study, tumors were removed transurethrally; and patients were randomly assigned into one of the three treatment groups: placebo, thiotepa, and pyridoxine. Many patients had multiple recurrences of tumors during the study and new tumors were recorded and removed at each visit. For each patient, the recurrence time, if any, was measured from the beginning of treatment. Following Sun and Wei (2000), we focus on

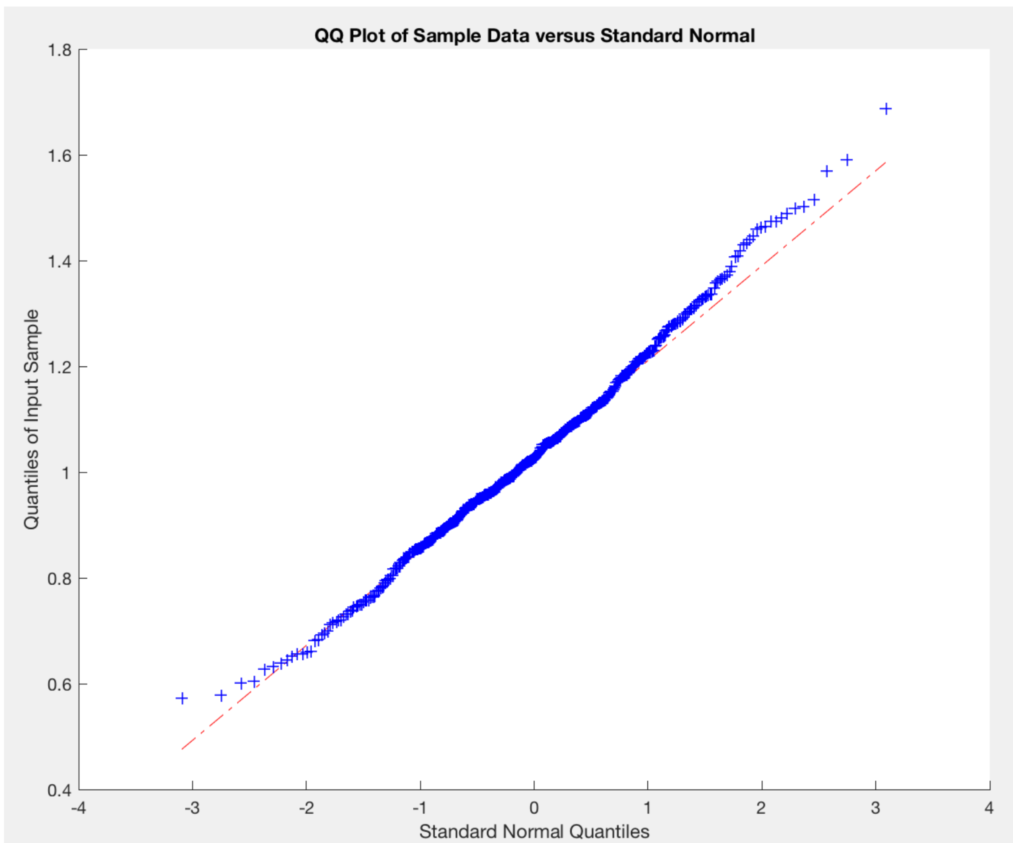

Figure 4.2: Dependent Case: QQ Plot for $\hat{\beta}_z$ with $\gamma = (1, 1), \sigma = 0.15$

85 bladder cancer patients in the placebo and thiotepa groups and analyze three baseline covatiates, including the treatment indicator, the number of initial tumors before entering the trial, and the size of the largest initial tumor.

For patient i, define X_{i1} to be the indicator of the treatment (0 for placebo; 1 for thiotepa), X_{i2} be number of initial tumors, and Z_i to be the size of the largest initial tumor, $i = 1, \ldots, 85$. We assume that the visit process of the bladder tumor can be described by model (4.6) depending on the covariates; and the occurrence processes of the bladder tumor can be described by models (4.3) and (4.7) for independent and dependent observation processes, respectively. X_i's were accuratly recorded, but for Z_i's, we consider adding a hypothetical normally distributed error term with mean 0 and variance 0.09. We applied both estimating equations (4.5) and (4.9) to estimate the regression parameters under the scenaios with independent and dependent observation processes, respectively.

For the independent case, we obtained $\hat{\beta} = (-1.0011, 0.2281, -0.0066)'$, with the corresponding standard errors $(0.4217, 0.0906, 0.1204)'$. The results suggest that the treatment and number of initial tumors had significant effects. In particular, the thiotepa treatment effectively reduced the recurrence rate of bladder tumors; and the number of initial tumors was positively associated with the tumor recurrent rate. Our conclusions are similar to those in Sun and Wei (2000).

For the dependent observation process, we obtained $\hat{\gamma} = (0.4999, -0.0053, 0.0268)'$, with the corresponding standard errors $(0.1209, 0.0408, 0.0383)'$, indicating a covariate dependent observation process. Patients seemed to visit the clinics more often in the thiotepa treatment group compared to those in the placebo group. With respect to the tumor recurrence process, $\hat{\beta} = (-1.5006, 0.2336, -0.0325)'$ with the corresponding standard errors $(0.4396, 0.0970, 0.1265)'$. Similar conclusions can be made on the effects of the thiotepa treatment and the number of initial tumors as those under the independent observation process. Both results are comparable with those in Sun and Wei (2000) and others.

4.5 Conclusion and Discussion

In the previous sections, we discussed the estimating procedures under a semi-parametric regression model for panel count data with measurement errors with both independent and dependent observation processes. Based on the simulation results, our estimates seem unbiased and asymptotically normally distributed.

In this chapter, the recurrent event process and the observation process are assumed to be independent given covariates. In addition, the proposed estimation procedures require the assumption of independent censoring. Developing a joint modeling approach to incorporate informative observation and censoring processes would be an important direction for future research.

4.6 Appendix

4.6.1 Detailed Tables in Simulation

Table 4.6: Detailed Simulation Results for Independent Observation Processes, n=200

σ	Methods	β	Relative Bias	SEE	SSE	CP
0.1	Proposed	(1,1)	(-0.7786%,**-0.6170%**)	(0.1793,0.1089)	(0.1937,0.1285)	(0.938,0.916)
		(1,-1)	(-0.3443%,**-0.3812%**)	(0.1788,0.1097)	(0.1912,0.1282)	(0.922,0.900)
		(-1,1)	(-0.1581%,**0.1352%**)	(0.1876,0.1194)	(0.1928,0.1301)	(0.944,0.940)
	Sun and Wei (2000)	(1,1)	(-1.5220%,**-2.6553%**)	(0.1792,0.1037)	(0.1896,0.1239)	(0.936,0.864)
		(1,-1)	(-0.5771%,**-1.4706%**)	(0.1792,0.1060)	(0.2016,0.1262)	(0.922,0.866)
		(-1,1)	(1.1425%,**-2.7816%**)	(0.1860,0.1088)	(0.2038,0.1212)	(0.924,0.880)
0.2	Proposed	(1,1)	(0.0573%,**-0.4872%**)	(0.1822,0.1101)	(0.1915,0.1288)	(0.942,0.898)
		(1,-1)	(-1.9597%,**-0.5119%**)	(0.1841,0.1112)	(0.2019,0.1344)	(0.920,0.902)
		(-1,1)	(1.4674%,**0.2105%**)	(0.1941,0.1188)	(0.2082,0.1356)	(0.936,0.902)
	Sun and Wei (2000)	(1,1)	(-0.5452%,**-4.9542%**)	(0.1827,0.1065)	(0.1957,0.1246)	(0.946,0.840)
		(1,-1)	(-0.5258%,**-4.1207%**)	(0.1815,0.1065)	(0.2024,0.1256)	(0.916,0.852)
		(-1,1)	(0.5056%,**-4.8661%**)	(0.1904,0.1081)	(0.2035,0.1287)	(0.922,0.844)
0.3	Proposed	(1,1)	(1.5061%,**1.6781%**)	(0.1904,0.1238)	(0.2125,0.1379)	(0.926,0.924)
		(1,-1)	(0.9629%,**1.1208%**)	(0.1935,0.1241)	(0.2165,0.1486)	(0.940,0.882)
		(-1,1)	(1.3962%,**1.3304%**)	(0.1993,0.1290)	(0.2074,0.1509)	(0.932,0.910)
	Sun and Wei (2000)	(1,1)	(-0.6865%,**-9.1691%**)	(0.1874,0.1026)	(0.2084,0.1204)	(0.922,0.754)
		(1,-1)	(0.9799%,**-8.8887%**)	(0.1857,0.1034)	(0.2075,0.1202)	(0.912,0.772)
		(-1,1)	(-1.4699%,**-9.0070%**)	(0.1957,0.1089)	(0.2138,0.1261)	(0.924,0.780)
0.4	Proposed	(1,1)	(1.5177%,**1.9298%**)	(0.2049,0.1487)	(0.2152,0.1652)	(0.940,0.900)
		(1,-1)	(1.5702%,**1.9082%**)	(0.2060,0.1510)	(0.2210,0.1617)	(0.942,0.898)
		(-1,1)	(0.2259%,**1.1685%**)	(0.2121,0.1527)	(0.2214,0.1615)	(0.950,0.888)
	Sun and Wei (2000)	(1,1)	(-0.4485%,**-14.0378%**)	(0.1926,0.1036)	(0.2108,0.1253)	(0.918,0.604)
		(1,-1)	(1.7291%,**-15.2704%**)	(0.1938,0.1025)	(0.2169,0.1136)	(0.912,0.600)
		(-1,1)	(-0.6813%,**-14.7216%**)	(0.2010,0.1055)	(0.2097,0.1326)	(0.964,0.588)
0.5	Proposed	(1,1)	(1.3040%,**3.8038%**)	(0.2317,0.2039)	(0.2420,0.2017)	(0.946,0.882)
		(1,-1)	(-0.1917%,**1.8197%**)	(0.2270,0.1966)	(0.2263,0.1885)	(0.952,0.900)
		(-1,1)	(1.3985%,**4.0932%**)	(0.2358,0.2034)	(0.2472,0.1984)	(0.946,0.870)
	Sun and Wei (2000)	(1,1)	(-0.2331%,**-21.0808%**)	(0.1970,0.0984)	(0.2158,0.1113)	(0.942,0.434)
		(1,-1)	(0.2984%,**-20.6654%**)	(0.1982,0.1015)	(0.2083,0.1228)	(0.938,0.414)
		(-1,1)	(0.9879%,**-20.4018%**)	(0.2091,0.1061)	(0.2322,0.1402)	(0.920,0.446)

Table 4.7: Detailed Simulation Results for Independent Observation Processes, n=500

σ	Methods	β	Relative Bias	SEE	SSE	CP
0.1	Proposed	(1,1)	(0.1493%,**-0.6832%**)	(0.1170,0.0746)	(0.1227,0.0889)	(0.944,0.902)
		(1,-1)	(0.2149%,**-0.3999%**)	(0.1188,0.0733)	(0.1177,0.0825)	(0.946,0.906)
		(-1,1)	(-0.8781%,**0.1026%**)	(0.1241,0.0775)	(0.1266,0.0826)	(0.942,0.916)
	Sun and Wei (2000)	(1,1)	(0.4624%,**-1.3631%**)	(0.0854,0.0548)	(0.0868,0.0603)	(0.936,0.888)
		(1,-1)	(0.4408%,**-1.7434%**)	(0.0858,0.0544)	(0.0897,0.0606)	(0.938,0.902)
		(-1,1)	(-0.1347%,**-1.3275%**)	(0.0893,0.0547)	(0.0958,0.0645)	(0.940,0.896)
0.2	Proposed	(1,1)	(-0.7339%,**-0.2172%**)	(0.1202,0.0759)	(0.1203,0.0883)	(0.934,0.898)
		(1,-1)	(-0.3657%,**-0.6457%**)	(0.1225,0.0776)	(0.1302,0.0820)	(0.930,0.902)
		(-1,1)	(0.2003%,**-0.1487%**)	(0.1254,0.0800)	(0.1293,0.0869)	(0.940,0.926)
	Sun and Wei (2000)	(1,1)	(-0.1220%,**-3.9465%**)	(0.0887,0.0556)	(0.0909,0.0622)	(0.942,0.840)
		(1,-1)	(0.5252%,**-4.2322%**)	(0.0870,0.0559)	(0.0834,0.0592)	(0.960,0.816)
		(-1,1)	(-0.9616%,**-4.8557%**)	(0.0904,0.0558)	(0.0943,0.0649)	(0.954,0.780)
0.3	Proposed	(1,1)	(0.3117%,**0.2008%**)	(0.0908,0.0630)	(0.0914,0.0711)	(0.950,0.920)
		(1,-1)	(0.9629%,**-0.3544%**)	(0.0908,0.0631)	(0.0960,0.0702)	(0.936,0.920)
		(-1,1)	(-0.3331%,**-0.0723%**)	(0.0946,0.0640)	(0.0964,0.0724)	(0.950,0.898)
	Sun and Wei (2000)	(1,1)	(0.3126%,**-8.2364%**)	(0.0900,0.0550)	(0.0921,0.0676)	(0.948,0.582)
		(1,-1)	(-0.0339%,**-9.0132%**)	(0.0895,0.0546)	(0.0982,0.0631)	(0.922,0.540)
		(-1,1)	(0.5705%,**-8.4499%**)	(0.0926,0.0560)	(0.0978,0.0636)	(0.936,0.574)
0.4	Proposed	(1,1)	(0.3280%,**0.0114%**)	(0.0951,0.0677)	(0.0948,0.0765)	(0.956,0.922)
		(1,-1)	(-0.2009%,**-0.3720%**)	(0.0510,0.0254)	(0.0536,0.0271)	(0.938,0.910)
		(-1,1)	(-0.1442%,**-0.1064%**)	(0.0987,0.0691)	(0.0996,0.0728)	(0.936,0.932)
	Sun and Wei (2000)	(1,1)	(0.4824%,**-13.7046%**)	(0.0918,0.0547)	(0.1005,0.0612)	(0.932,0.284)
		(1,-1)	(-0.0716%,**-14.2258%**)	(0.0920,0.0573)	(0.0980,0.0605)	(0.934,0.274)
		(-1,1)	(-0.2010%,**-14.0629%**)	(0.0950,0.0553)	(0.1005,0.0655)	(0.926,0.286)
0.5	Proposed	(1,1)	(0.3313%,**0.6319%**)	(0.1415,0.1078)	(0.1466,0.1200)	(0.946,0.910)
		(1,-1)	(0.6736%,**0.5521%**)	(0.1024,0.0778)	(0.1061,0.0850)	(0.932,0.932)
		(-1,1)	(0.2432%,**0.7855%**)	(0.1052,0.0786)	(0.1085,0.0887)	(0.946,0.928)
	Sun and Wei (2000)	(1,1)	(-0.0735%,**-20.4502%**)	(0.0957,0.0530)	(0.0985,0.0559)	(0.930,0.100)
		(1,-1)	(-0.0464%,**-19.7650%**)	(0.0963,0.0547)	(0.0991,0.0629)	(0.942,0.122)
		(-1,1)	(-0.2635%,**-20.4482%**)	(0.0986,0.0540)	(0.1049,0.0604)	(0.936,0.104)

Table 4.8: Detailed Simulation Results of $\gamma =$ for Dependent Observation Processes, $n = 500$

σ	Methods	γ	Relative Bias	BSE	SSE	CP
0.05	Proposed	(1,1)	(0.6913%,**-0.0248%**)	(0.0702,0.0314)	(0.0688,0.0326)	(0.952,0.920)
		(1,-1)	(-0.6643%,**-0.7381%**)	(0.0707,0.0316)	(0.0734,0.0314)	(0.940,0.940)
	Sun and Wei (2000)	(1,1)	(-0.6760%,**-1.1926%**)	(0.0709,0.0317)	(0.0734,0.0328)	(0.944,0.914)
		(1,-1)	(-0.7014%,**-1.1323%**)	(0.0705,0.0316)	(0.0705,0.0323)	(0.946,0.932)
0.10	Proposed	(1,1)	(-0.9327%,**-0.6041%**)	(0.0719,0.0323)	(0.0729,0.0329)	(0.940,0.916)
		(1,-1)	(-0.3627%,**-0.7575%**)	(0.0714,0.0320)	(0.0731,0.0360)	(0.942,0.902)
	Sun and Wei (2000)	(1,1)	(-0.7738%,**-1.8235%**)	(0.0714,0.0318)	(0.0718,0.0330)	(0.946,0.896)
		(1,-1)	(-1.0992%,**-1.7769%**)	(0.0718,0.0323)	(0.0692,0.0331)	(0.946,0.898)
0.15	Proposed	(1,1)	(-0.1903%,**-0.7042%**)	(0.0733,0.0329)	(0.0734,0.0380)	(0.950,0.898)
		(1,-1)	(-0.4867%,**-0.6219%**)	(0.0736,0.0331)	(0.0764,0.0381)	(0.932,0.910)
	Sun and Wei (2000)	(1,1)	(-0.1978%,**-3.2756%**)	(0.0734,0.0327)	(0.0751,0.0357)	(0.942,0.782)
		(1,-1)	(-0.4437%,**-2.9639%**)	(0.0731,0.0382)	(0.0735,0.0361)	(0.956,0.796)
0.20	Proposed	(1,1)	(-0.0060%,**-0.3781%**)	(0.0760,0.0345)	(0.0773,0.0385)	(0.946,0.902)
		(1,-1)	(-0.4753%,**-0.6663%**)	(0.0756,0.0341)	(0.0719,0.0398)	(0.954,0.898)
	Sun and Wei (2000)	(1,1)	(-0.9945%,**-4.8764%**)	(0.0750,0.0334)	(0.0788,0.0382)	(0.924,0.646)
		(1,-1)	(-0.4976%,**-4.5312%**)	(0.0752,0.0328)	(0.0746,0.0351)	(0.948,0.682)
0.25	Proposed	(1,1)	(-0.4772%,**-0.4129%**)	(0.0788,0.0356)	(0.0780,0.0456)	(0.948,0.910)
		(1,-1)	(-0.1755%,**-0.0954%**)	(0.0784,0.0356)	(0.0823,0.0414)	(0.940,0.902)
	Sun and Wei (2000)	(1,1)	(-0.6712%,**-6.4900%**)	(0.0782,0.0349)	(0.0753,0.0379)	(0.956,0.512)
		(1,-1)	(-1.2598%,**-6.5517%**)	(0.0776,0.0342)	(0.0788,0.0385)	(0.940,0.518)

Table 4.9: Detailed Simulation Results of β for $\gamma = (1, 1)$, $n = 500$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.05	Proposed	(1,1)	(0.9588%,**-0.6552**%)	(0.2056,0.1313)	(0.2056,0.1561)	(0.926,0.928)
		(1,-1)	(1.2327%,**-0.8417**%)	(0.1877,0.1046)	(0.1742,0.1049)	(0.954,0.966)
		(-1,1)	(-0.6598%,**0.3857**%)	(0.2134,0.1301)	(0.2412,0.1571)	(0.932,0.920)
	Sun and Wei (2000)	(1,1)	(0.4725%,**-1.9828**%)	(0.2062,0.1292)	(0.2359,0.1641)	(0.922,0.854)
		(1,-1)	(1.0884%,**-1.2271**%)	(0.1865,0.1041)	(0.1830,0.1106)	(0.952,0.906)
		(-1,1)	(0.2021%,**-1.0582**%)	(0.2146,0.1314)	(0.2284,0.1680)	(0.932,0.844)
0.10	Proposed	(1,1)	(1.1331%,**0.6431**%)	(0.2060,0.1310)	(0.2330,0.1620)	(0.932,0.898)
		(1,-1)	(1.3084%,**-1.1284**%)	(0.1879,0.1036)	(0.1738,0.1036)	(0.964,0.938)
		(-1,1)	(-1.1257%,**1.2009**%)	(0.2205,0.1386)	(0.2372,0.1639)	(0.944,0.892)
	Sun and Wei (2000)	(1,1)	(-0.9522%,**-1.9419**%)	(0.1974,0.1333)	(0.2121,0.1567)	(0.936,0.878)
		(1,-1)	(-0.3339%,**-2.0754**%)	(0.1863,0.1021)	(0.1752,0.1046)	(0.974,0.886)
		(-1,1)	(-0.8960%,**-1.9528**%)	(0.2202,0.1321)	(0.2410,0.1583)	(0.934,0.886)
0.15	Proposed	(1,1)	(0.4787%,**1.0416**%)	(0.2245,0.1521)	(0.2335,0.1763)	(0.960,0.902)
		(1,-1)	(-0.6611%,**-0.7260**%)	(0.1861,0.1059)	(0.1762,0.1086)	(0.968,0.930)
		(-1,1)	(-1.0735%,**1.2249**%)	(0.2300,0.1510)	(0.2549,0.1712)	(0.924,0.916)
	Sun and Wei (2000)	(1,1)	(-0.8855%,**-3.6691**%)	(0.2153,0.1365)	(0.2364,0.1749)	(0.934,0.834)
		(1,-1)	(1.2704%,**-3.4577**%)	(0.1853,0.1024)	(0.1724,0.1080)	(0.962,0.820)
		(-1,1)	(-0.8739%,**-2.9143**%)	(0.2262,0.1355)	(0.2492,0.1698)	(0.936,0.844)
0.20	Proposed	(1,1)	(1.7041%,**3.4319**%)	(0.2469,0.1813)	(0.2822,0.2069)	(0.938,0.908)
		(1,-1)	(2.2742%,**-0.4728**%)	(0.1892,0.1043)	(0.1800,0.1115)	(0.960,0.932)
		(-1,1)	(-0.3026%,**3.5538**%)	(0.2506,0.1735)	(0.2610,0.1896)	(0.956,0.930)
	Sun and Wei (2000)	(1,1)	(-0.6144%,**-6.5813**%)	(0.2224,0.1361)	(0.2223,0.1736)	(0.958,0.834)
		(1,-1)	(1.4522%,**-5.4214**%)	(0.1882,0.1010)	(0.1769,0.1036)	(0.968,0.808)
		(-1,1)	(-0.4873%,**-5.4593**%)	(0.2347,0.1391)	(0.2493,0.1588)	(0.944,0.866)
0.25	Proposed	(1,1)	(3.0082%,**6.1529**%)	(0.2750,0.2180)	(0.2884,0.2320)	(0.972,0.862)
		(1,-1)	(1.4722%,**-1.2413**%)	(0.1883,0.1065)	(0.1781,0.1115)	(0.954,0.936)
		(-1,1)	(0.7193%,**4.7237**%)	(0.2710,0.2018)	(0.2745,0.2351)	(0.944,0.884)
	Sun and Wei (2000)	(1,1)	(-0.7701%,**-8.9750**%)	(0.2386,0.1478)	(0.2581,0.1886)	(0.954,0.796)
		(1,-1)	(1.6277%,**-6.8879**%)	(0.1867,0.1001)	(0.1706,0.1100)	(0.958,0.838)
		(-1,1)	(0.2642%,**-8.8771**%)	(0.2495,0.1475)	(0.2672,0.1663)	(0.955,0.806)

Table 4.10: Detailed Simulation Results of β for $\gamma = (1, -1)$, $n = 500$

σ	Methods	β	Relative Bias	BSE	SSE	CP
0.05	Proposed	(1,1)	(0.3511%,**-0.5055%**)	(0.1883,0.1038)	(0.1747,0.1043)	(0.954,0.948)
		(1,-1)	(1.0685%,**0.1582%**)	(0.2071,0.1303)	(0.2138,0.1350)	(0.928,0.936)
		(-1,1)	(-0.9834%,**0.5744%**)	(0.2271,0.1326)	(0.2164,0.1415)	(0.962,0.936)
	Sun and Wei (2000)	(1,1)	(1.1875%,**-1.7248%**)	(0.1872,0.1033)	(0.1755,0.1031)	(0.946,0.922)
		(1,-1)	(-0.0655%,**-2.3055%**)	(0.2039,0.1266)	(0.2335,0.1603)	(0.930,0.862)
		(-1,1)	(-0.0620%,**-1.6338%**)	(0.2279,0.1340)	(0.2348,0.1450)	(0.946,0.902)
0.10	Proposed	(1,1)	(0.7512%,**1.0295%**)	(0.1883,0.1047)	(0.1772,0.1097)	(0.964,0.952)
		(1,-1)	(0.2372%,**1.1370%**)	(0.2107,0.1325)	(0.2390,0.1577)	(0.940,0.908)
		(-1,1)	(0.3349%,**0.9634%**)	(0.2286,0.1328)	(0.2202,0.1431)	(0.942,0.940)
	Sun and Wei (2000)	(1,1)	(-0.0010%,**-2.9727%**)	(0.1880,0.1021)	(0.1737,0.1074)	(0.958,0.900)
		(1,-1)	(-0.1884%,**-1.9335%**)	(0.2078,0.1304)	(0.2322,0.1605)	(0.932,0.844)
		(-1,1)	(1.5041%,**-2.9314%**)	(0.2278,0.1329)	(0.2130,0.1463)	(0.956,0.888)
0.15	Proposed	(1,1)	(1.0357%,**1.5548%**)	(0.1886,0.1052)	(0.1753,0.1101)	(0.964,0.922)
		(1,-1)	(1.2434%,**1.8937%**)	(0.2168,0.1418)	(0.2338,0.1726)	(0.946,0.892)
		(-1,1)	(0.0834%,**2.0296%**)	(0.2295,0.1350)	(0.2306,0.1437)	(0.942,0.934)
	Sun and Wei (2000)	(1,1)	(2.9117%,**-3.5696%**)	(0.1874,0.1003)	(0.1785,0.1097)	(0.954,0.870
		(1,-1)	(2.5333%,**-3.1994%**)	(0.2178,0.1340)	(0.2502,0.1667)	(0.934,0.860)
		(-1,1)	(-0.1851%,**-2.9873%**)	(0.2280,0.1325)	(0.2283,0.1380)	(0.950,0.900)
0.20	Proposed	(1,1)	(2.3356%,**3.5851%**)	(0.1868,0.1066)	(0.1961,0.1159)	(0.950,0.896)
		(1,-1)	(3.9029%,**-3.6185%**)	(0.2340,0.1494)	(0.2685,0.1691)	(0.946,0.890)
		(-1,1)	(0.5471%,**2.4405%**)	(0.2281,0.1364)	(0.2150,0.1452)	(0.972,0.914)
	Sun and Wei (2000)	(1,1)	(2.6750%,**-6.1378%**)	(0.1879,0.1012)	(0.1811,0.1062)	(0.964,0.902)
		(1,-1)	(0.7064%,**-5.8169%**)	(0.2298,0.1439)	(0.2428,0.1793)	(0.942,0.844)
		(-1,1)	(0.2154%,**-5.3519%**)	(0.2300,0.1316)	(0.2238,0.1375)	(0.954,0.826)
0.25	Proposed	(1,1)	(0.9424%,**6.2481%**)	(0.1881,0.1074)	(0.1723,0.1202)	(0.970,0.858)
		(1,-1)	(2.8470%,**9.6705%**)	(0.2410,0.1630)	(0.2586,0.1813)	(0.936,0.758)
		(-1,1)	(1.6476%,**9.1468%**)	(0.2296,0.1383)	(0.2308,0.1269)	(0.954,0.826)
	Sun and Wei (2000)	(1,1)	(0.9365%,**-10.6765%**)	(0.1862,0.0983)	(0.1796,0.1054)	(0.948,0.736)
		(1,-1)	(0.7440%,**-12.0503%**)	(0.2433,0.1518)	(0.2661,0.1798)	(0.950,0.728)
		(-1,1)	(-0.6130%,**-10.0312%**)	(0.2274,0.1270)	(0.2340,0.1309)	(0.944,0.712)

4.6.2 Consistency

To derive the consistency of the proposed estimator $\hat{\beta}_{1n}$, we need the following regularity conditions.

(C1). The true value β_0 belongs to a known open, convex, and bounded set $\mathcal{B}$ in R^p with $p = p_1 + p_2 + 1$.

(C2). The censoring time C_i is independent of $N_i^*(\cdot)$, $H_i(\dots)$, X_i, Z_i and ϵ_i. Also the measurement error ϵ_i is also independent of $N_i^*(\cdot)$, $H_i(\dots)$, X_i, Z_i and C_i.

(C3). $P(C \geq \tau) > 0$, where τ is the largest follow-up time.

(C4). The observational process $H_i(\cdot)$ is independent of X_i, Z_i and $N_i^*(\cdot)$.

(C4′). The observational process $H_i(\cdot)$ is conditional independent of $N_i^*(\cdot)$ given X_i, Z_i.

C5. The matrix $A(\beta) = E \exp(X_i'\beta_x + Z_i'\beta_z + \xi)V_iV_i'$ is positive definite for $\beta \in$ a neighborhood of β_0 where $V_i = (X_i', Z_i', 1)'$.

Lemma B1 Under the assumptions of C1-C4, one obtains that

$$E\left[\exp(W_i'\beta_z)|O_i, Z_i\right] = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z), \qquad (i)$$

$$E\left[(W_i - \Sigma\beta_z)\exp(W_i'\beta_z)|O_i, Z_i\right] = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) Z_i \exp(Z_i'\beta_z), \qquad (ii)$$

$$E\left[(W_i - \Sigma\beta_z)^{\otimes 2}\exp(W_i'\beta_z)|O_i, Z_i\right] = \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)[Z_iZ_i' + \Sigma]\exp(Z_i'\beta_z). \qquad (iii)$$

Proof. Note that

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\exp(W_i'\beta_z)\exp\left(-\frac{1}{2}(W_i-Z_i)'\Sigma^{-1}(W_i-Z_i)\right)\\
&= \exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z)\\
&\quad\times\frac{1}{\sqrt{2\pi}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(W_i-Z_i-\Sigma\beta_z)'\Sigma^{-1}(W_i-Z_i-\Sigma\beta_z)\right),
\end{aligned}$$

which is the normal density function with mean $Z_i + \Sigma\beta_z$ and covariance matrix Σ, muliplied by a term $\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp(Z_i'\beta_z)$. Refer Chapter 3 for the derivations of the three equalities.

The consistency of the proposed estimators $\hat{\beta}_{1n}$ and $\hat{\beta}_{2n}$ can be derived similarly as in Chapter 3.

Detailed Proof for the Independent Observation Process

$$
\begin{aligned}
&\frac{1}{n}U_{n1}(\beta) \\
\rightarrow\quad & E\left[\begin{pmatrix} X \\ W \\ 1 \end{pmatrix} \bar{N}_i\right] \\
& - \exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right) E\left[\exp(X'\beta_x + W'\beta_z + \xi)\begin{pmatrix} X \\ W - \Sigma\beta_z \\ 1 \end{pmatrix}\right] \\
=\quad & E\left[\begin{pmatrix} X \\ W \\ 1 \end{pmatrix} \bar{N}_i\right] \\
& - \exp\left(-\frac{1}{2}\beta_z'\Sigma\beta_z\right)\exp\left(\frac{1}{2}\beta_z'\Sigma\beta_z\right) E\left[\exp(X'\beta_x + Z'\beta_z + \xi)\begin{pmatrix} X \\ Z \\ 1 \end{pmatrix}\right] \\
=\quad & E\left[\begin{pmatrix} X \\ W \\ 1 \end{pmatrix} \bar{N}_i\right] - E\left[\exp(X'\beta_x + Z'\beta_z + \xi)\begin{pmatrix} X \\ Z \\ 1 \end{pmatrix}\right] \\
\rightarrow\quad & 0
\end{aligned}
$$

$\square$

Detailed Proof for the Dependent Observation Process

Let $\alpha_x = (\hat{\gamma}_x + \beta_x)'$, and $\alpha_z = (\hat{\gamma}_z + \beta_z)'$, then

$$
\begin{aligned}
&\frac{1}{n}U_{n2}(\beta) \\
\rightarrow\;& E\left[\begin{pmatrix} X \\ W \\ 1 \end{pmatrix}\bar{N}_i\right] \\
&- \exp\left(-\frac{1}{2}\alpha_z'\Sigma\alpha_z\right)E\left[\exp(X'\alpha_x + W'\alpha_z + \xi)\begin{pmatrix} X \\ W - \Sigma\alpha_z \\ 1 \end{pmatrix}\right] \\
=\;& E\left[\begin{pmatrix} X \\ W \\ 1 \end{pmatrix}\bar{N}_i\right] \\
&- \exp\left(-\frac{1}{2}\alpha_z'\Sigma\alpha_z\right)\exp\left(\frac{1}{2}\alpha_z'\Sigma\alpha_z\right)E\left[\exp(X'\alpha_x + Z'\alpha_z + \xi)\begin{pmatrix} X \\ Z \\ 1 \end{pmatrix}\right] \\
=\;& E\left[\begin{pmatrix} X \\ W \\ 1 \end{pmatrix}\bar{N}_i\right] - E\left[\exp(X'\alpha_x + Z'\alpha_z + \xi)\begin{pmatrix} X \\ Z \\ 1 \end{pmatrix}\right] \\
\rightarrow\;& 0
\end{aligned}
$$

$\square$

Chapter 5: Conclusions and Discussions

In the preceding chapters, we proposed some estimating equation based approaches for regression analysis of both recurrent event data and panel count data. A key feature for these methods is that they allow the presence of measurement errors on some of the covariates. We studied different situations and provided solutions for both types of data.

For recurrent event data, we proposed two approaches dealing with the issues with the presence of measurement errors in some of the covariates. In Model A, the estimating equation was developed following Lin et al. (2000). In Model B, we adopted the procedure in Wang et al. (2001). The resulting estimates were shown to be consistent and asymptotically normal. Simulation studies indicate that the proposed estimation procedures performed well for practical situations. An example from the study of gamma interferon in chronic granulomatous disease is provided.

As for panel count data, we considered both independent and dependent observation processes. When the observation process is independent of covariates, our method can provide unbiased estimates when sample size is at least 500. We obtained similar results when the observation process is dependent of covariates. Due to the two-step estimation approach for the dependent observation processes, a larger sample size and smaller measurement error variance would be needed to reduce bias. The proposed methods were applied to the panel count data from a bladder cancer study with hypothetical measurement errors.

Future research for recurrent event data may lie on the studies when there are more than one type of event of interest, resulting in multivariate recurrent event data. In addition, when the multiple types of recurrent events are dependent, it would be interesting to generalize our proposed methods and utilize a joint modeling approach to develop similar estimating equation based approaches.

For the regression analysis of panel count data discussed in Chapter 4, it was assumed that the recurrent event process and the observation process are independent of each other given covariates. Moreover, the proposed estimation procedures only work under independent censoring. It would be important for future research to develop a joint modeling method to incorporate informative observation and censoring processes.

Bibliography

[1] Andersen PK, Borgan O, Gill RD, Keiding N. *Statistical models based on counting processes*. Springer: New York. 1993.

[2] Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study. *Annals of Statistics* 1982, **10:** 1100-1120.

[3] Balakrishnan N, Zhao X. New multi-sample nonparametric tests for panel count data. *Annals of Statistics* 2000, **37:** 1112-1149.

[4] Balakrishnan N, Zhao X. A class of multi-sample nonparametric tests for panel count data. *Annals of the Institute of Statistical Mathematics* 2010, **63:** 135-156.

[5] Bartlett P. Theoretical Statistics. Lecture 8 & 14. Spring 2008. Class Notes.

[6] Bilias Y, Gu M, Ying Z. Towards a general asymptotic theory for Cox model with staggered entry. *Annals of Statistics* 1997, **25:** 662-682.

[7] Breslow NE. Covariance analysis of censored survival data. *Biometrics* 1974. **30:** 89-99.

[8] Cai J, Prentice RL. Estimating equations for hazard ratio parameters based on correlated failure time data. *Biometrika* 1995, **82:** 151-164.

[9] Cai J, Schaubel DE. Marginal means/rates models for multiple type recurrent event data. *Lifetime Data Analysis* 2004, **10:** 121-138.

[10] Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. *Measurement error in nonlinear models*. 2nd ed. Chapman & Hall/CRC, 2006.

[11] Chang SH, Wang MC. Conditional regression analysis for recurrence time data. *J. Am. Statist. Ass.* 1999, **94:** 1221-1230.

[12] Chen BE, Cook RJ, Lawless JF, Zhan M. Staitistical methods for multivariate interval censored recurrent events. *Statistics in Medicine* 2005, **24:** 671-691.

[13] Cheng SC, Wei LJ. Inferences for a semiparametric model with panel data. *Biometrika*, **87:** 89-97.

[14] Chow YS, Teicher H. *Probability Theory: Independence, Interchangeability, Martingales*. 2nd ed. Springer: New York. 1988.

[15] Clegg LX, Cai J, Sen PK. A Marginal Mixed Baseline Hazards Model for Multivariate Failure Time Data. *Biometrics* 1999, **55:** 805-812.

[16] Cook RJ, Lawless JF. Interim monitoring of longitudinal comparative studies with recurrent event responses. *Biometrics* 1996, **52:** 1311-1323.

[17] Cook RJ, Lawless JF. Marginal analysis of recurrent events and a terminal event. *Statistics in Medicine* 1997, **16:** 911-924.

[18] Cook RJ, Lawless JF. *The Statistical Analysis of Recurrent Events* 2007, Springer: New York.

[19] Cook RJ, Lawless JF, Nadeau JC. Robust tests for treatment comparisons based on recurrent event responses. *Biometrics* 1996, **52:** 557-571.

[20] Cox DR. Regression models and life-tables (with discussion). *J. R. Statist. Soc.* 1972, **B, 34:** 187-220.

[21] Farewell VT, Sprott DA. The use of a mixture model in the analysis of count data. *Biometrics* 1988, **44:** 1191-1194.

[22] Fleming TR, Harrington DP. *Counting Processes and Survival Analysis* 1991, Wiley-Interscience: New York.

[23] Gail MH, Santner TJ, Brown CC. An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics* 1980, **36:** 255-266.

[24] Han F. Advanced Theory of Statistical Inference. Lecture 1. Spring 2018. Class Notes.

[25] He X, Tong X, Sun J, Cook RJ. Regression analysis of multivariate panel count data. *Biostatistics* 2008, **9:** 234-248.

[26] Huang CY, Wang MC. Joint modeling and estimation for recurrent event processes and failure time data. *J. Am. Statist. Ass.* 2004, **99:** 1153-1165.

[27] Jiang W, Turnbull BW, Clark LC. Semiparametric regression models for repeated events with random effects and measurement error. *Journal of American Statistical Assocation* 1999, **94:** 111-124.

[28] Kalbfleisch JD, Lawless JF. The analysis of panel data under a Markov assumption. *J. Am. Statist. Ass.* 1985, **80:** 863-871.

[29] Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data.* Wiley: New York. 1980.

[30] Kim YJ. Analysis of panel count data with measurements erros in the covariates. *Journal of Statistics Computing and Computing* 2007, **77:2** 109-117.

[31] Kosorok MR. *Introduction to Empirical Processes and Semiparametric Inference*. Springer. 2008.

[32] Lawless JF. Regression methods for Poisson process data. *Journal of American Statistical Association* 1987, **82:** 808-815.

[33] Lawless JF, Nadeau C. Some simple robust methods for the analysis of recurrent events. *Technometrics* 1995, **37:** 158-168.

[34] Lawless JF, Nadeau C, Cook RJ. Analysis of mean and rate functions for recurrent events. In *Proc. 1st Seattle Symp. Biostatistics: Survival Analysis* (eds D. Y. Lin and T. R. Fleming). Springer: New York, 1997, pp. 37-49.

[35] Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986, **73:** 13-22.

[36] Lin DY. Goodness-of-fit analysis for the Cox regression model based on a class of parameter estimators. *J. Am. Statist. Ass.* 1991, **86:** 725-728.

[37] Lin DY, Sun W, Ying Z. Nonparametric estimation of the gap time distributions for serial events with censored data. *Biometrika* 1999, **86:** 59-70.

[38] Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. *J. Am. Statist. Ass.* 1989, **84:** 1074-1078.

[39] Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *Journal of Royal Statistical Society* 2000, **62:4:** 711-730.

[40] Lin DY, Wei LJ, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* 1993, **80:** 557-572.

[41] Lin DY, Ying Z. Cox regression with incomplete covariate measurement. *Journal of American Statistical Association* 1993, **88:** 1341-1349.

[42] Nakamura T. Proportional hazards model with covariates subject to measurement error. *Biometrics* 1992, **48:** 829-838.

[43] Prentice RL. Covariate measurement errors and parameter estimation in failure time regression models. *Biometrika* 1982, **69:** 331-342.

[44] Pepe MS, Fleming TR. A nonparametric method for dealing with mismeasured covariate data. *Journal of American Statistical Association* 1991, **86:** 108-113.

[45] Ross SM. *Stochastic Processes* 1983, Wiley: New York.

[46] Sun J, Fang HB. A nonparametric test for panel count data. *Biometrika* 2003, **90:** 199-208.

[47] Sun J, Kalbfleisch JD. The analysis of current status data on point processes. *J. Am. Statist. Ass.* 1993, **88:** 1449-1454.

[48] Sun J, Kalbfleisch JD. Estimation of the mean function of point processes based on panel count data.*Statistica Sinica* 1995, **5:** 279-290.

[49] Sun J, Wei LJ. Regression analysis of panel count data with covariate-dependent observation and censoring times. *J. R. Statist. Soc, Series B* 2000, **62:** 293-302.

[50] Sun J, Zhao X. *The Statistical Analysis of Panel Count Data* 2013, Springer: New York.

[51] Thall PF. Mixed Poisson likelihood regression models for longitudinal interval count data. *Biometrics* 1988, **45:** 1039.

[52] Thall PF, Lachin JM. Analysis of recurrent events: Nonparametric methods for random-interval count data. *J. Am. Statist. Ass.* 1988, **83:** 339-347.

[53] Turnbull BW, Jiang W, Clark LC. Regression models for recurrent event data: parametric random effects models with measurement error. *Stat Med.*, 1997 **30:16:** 853-864.

[54] Wang CY, Hsu l, Feng ZD, Prentice RL. Regression calibration in failure time regression. *Biometrics* 1997, **53:** 131-145.

[55] Wang MC, Chang SH. Nonparametric estimation of a recurrent survival function. *J. Am. Statist. Ass.* 1999, **94:** 146-153.

[56] Wang MC, Jewell NP, Tsai WY. Asymptotic properties of the product limit estimate under random truncation. *The Annals of Staitistics* 1986, **14:** 1597-1605.

[57] Wang MC, Qin J, Chiang CT. Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association* 2001, **96:** 1057-1065.

[58] Wang MC, Wells MT. Nonparametric estimation of successive duration times under dependent censoring. *Biometrika* 1998, **85:** 561-572.

[59] Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Am. Statist. Ass.* 1989, **84:** 1065-1073.

[60] Wellner JA, Zhang Y. Two estimators of the mean of a counting process with panel count data. *Annals of Statistics* 2000, **28:** 779-814.

[61] Woodroofe M. Estimating a distribution function with truncated data. *Annals of Statistics* 1985, **13:** 163-177.

[62] Yi GY. *Survival Data with Measurement Error* 2017, Springer.

[63] Yu H, Cheng YJ, Wang CY. Semiparametric regression estimation for recurrent event data with errors in covariate under informative censoring. *The International Journal of Biostatistics* 2016, **12(2)**.

[64] Zhang Y. A semiparametric pseudolikelihood estimation method for panel count data. *Biometrika* 2002, **89:** 39-48.

[65] Zhou H, Pepe MS. Auxiliary covariate data in failure time regression analysis. *Biometrika* 1995, **82:** 139-149.

[66] Zhou H, Wang CY. Failure time regression with continuous covariates measured with error. *Journal of Royal Statistical Society* 2000, **B,62:** 657-665.