Far East J. Theo. Stat. **13**(2) (2004), 233-256
(A special volume on Biostatistics and related areas)

A UNIFORM SEMIPARAMETRIC APPROACH FOR LONGITUDINAL DATA ANALYSIS

XIAN ZHOU, JIANGUO SUN and LIUQUAN SUN

Department of Applied Mathematics
Hong Kong Polytechnic University, Hong Kong
e-mail: maxzhou@polyu.edu.hk

Department of Statistics, University of Missouri
222 Math Sciences Building, Columbia, MO 65211, U. S. A.
e-mail: tsun@stat.missouri.edu

Institute of Applied Mathematics, Academia Sinica
Beijing 100080, P. R. China
e-mail: slq@mail.amt.ac.cn

Abstract

Longitudinal data commonly occur in medical follow-up studies and epidemiological experiments. They usually include repeated measurements of the response variable and covariates at a set of distinct, irregularly spaced time points for each subject. One of the difficulties for the analysis of such data is that the set of observation times may vary from subject to subject. For their analysis, a number of methods have been proposed, but most of them were developed under specific models. In this paper, a class of general and uniform models is presented for semiparametric analysis of longitudinal data. For inference about regression parameters, a class of consistent and asymptotically normal estimators is proposed. Extensive simulation

2000 Mathematics Subject Classification: Primary 62G05; Secondary 62F12.

Key words and phrases: counting process, estimating function, longitudinal study, repeated measurements.

This work was partially supported by the Hong Kong Polytechnic University Internal Research Grant No. A-PC37 and the National Natural Science Foundation of China.

Received November 5, 2003

studies are conducted and an example with data from an AIDS clinical trial is presented to illustrate the proposed methodology.

1. Introduction

Longitudinal data commonly occur in medical follow-up studies and observational investigations and one major difficulty for their analysis is that the response variable is often repeatedly measured at irregular and possibly subject-specific observation times. To address the difficulty, various models and methods have been proposed (Diggle et al. [3]; Laird and Ware [5]; Liang and Zeger [7]; Lin and Carroll [8]; Lin and Ying [9]). For example, Laird and Ware [5] proposed simple linear random effects models that are now commonly used in many fields. Following Laird and Ware [5], many authors have developed various more general models such as semiparametric and nonparametric random effects models. In particular, Lin and Ying [9] recently studied a general semiparametric mean function model given by

$$E\{X(t) \mid Z\} = \mu_0(t) + \beta'(t)Z, \tag{1.1}$$

for regression analysis of longitudinal data, where $X(t)$ is the response variable, Z is a vector of covariates, $\mu_0(t)$ is an unspecified smooth function of time t and β is a vector of unknown regression parameters.

A special case of longitudinal data is panel count data where the response variable is the outcome of a point process. Panel count data occur if a recurrent event is under study and only the number of occurrences of the event is known at each observation time. The analysis of panel data has recently attracted considerable attention (e.g., Thall and Lachin [17]; Cheng and Wei [2]; Sun and Wei [16]; Zhang [18]). In these analyses, a common way is to model the mean function of the response variable in some semiparametric form. In particular, several authors have investigated the so-called proportional mean model

$$E\{X(t) \mid Z\} = \mu_0(t)e^{\beta' Z} \tag{1.2}$$

(Cheng and Wei [2]; Sun and Wei [16]; Zhang [18]). For example, Zhang [18] discussed this model and proposed a semiparametric pseudolikelihood estimation method for inference.

Like the proportional hazards model in failure time data analysis, the models (1.1) and (1.2) may not be flexible enough to fit the data adequately in many realistic situations. To address this problem, some more general models have been considered. For example, Lin et al. [12] generalized model (1.2) to

$$E\{X(t) \mid Z\} = g(\mu_0(t)e^{\beta' Z}) \tag{1.3}$$

for regression analysis of panel count data, where $g(\cdot)$ is a specified function, assumed to be twice continuously differentiable and strictly increasing. In this paper, we present a class of general and uniform models for regression analysis of longitudinal data that includes most of the existing models as special cases, and present an example of its application in the analysis of AIDS data.

We will begin with introducing notations and describing the models that will be used throughout the paper in Section 2. Section 3 presents a class of estimating equations for the estimation of the regression parameters and the baseline mean function under the model. We will first discuss the situation where observation times are independent of covariates, and then deal with the case where they may depend on covariates. Results from simulation studies to evaluate the proposed methods are reported in Section 4. In Section 5 we apply the methodology to a set of longitudinal data from an AIDS clinical trial, followed by some concluding remarks in Section 6. Finally, the asymptotic properties of the proposed estimators and their proofs are provided in Appendix.

2. A Class of Uniform Semiparametric Models

Consider a longitudinal study consisting of n subjects. For the ith subject, let $X_i(t)$ denote the univariate response variable, Z_i represent a $p \times 1$ vector of covariates, and C_i be the censoring (follow-up) time on the subject. Also let $N_i(t)$ denote the number of times that subject i are observed by no later than $t \in [0, \tau]$, where τ is a constant. We assume that $\{X_i(\cdot), Z_i, C_i, N_i(\cdot)\}$, $i = 1, \ldots, n$, are independent and identically distributed (i.i.d.). For the analysis, we will consider the following semiparametric model:

$$E\{X_i(t)\,|\,Z_i\} = \Phi(\mu_0(t),\, \beta_0,\, Z_i), \tag{2.1}$$

which will be referred to as a *semiparametric transformation model*, where $\Phi(u, v, w)$ is a known function, which is assumed to be strictly increasing in u and twice continuously differentiable with respect to u and v, $\mu_0(t)$ is an unknown baseline mean function, and β_0 is a vector of unknown regression parameters.

The model in (2.1) specifies only the mean function structure of $X_i(\cdot)$ and leaves the stochastic structure of the process $X_i(\cdot)$ totally free. It defines a very rich family of models through the link function Φ and includes many existing models as special cases. For example, model (2.1) reduces to model (1.1) if $\Phi(u, v, w) = u + v'w$, or to the proportional mean model (1.2) when $\Phi(u, v, w) = u\exp(v'w)$. Model (1.3) is also a special case of (2.1) with $\Phi(u, v, w) = g(ue^{v'w})$. In addition, the general additive-multiplicative means model

$$E\{X_i(t)\,|\,Z_i\} = g(\beta_{10}'Z_{1i}) + \mu_0(t)h(\beta_{20}'Z_{2i})$$

is another special case of model (2.1) with $\Phi(u, v, w) = g(v_1'w_1) + uh(v_2'w_2)$, where $\beta_0 = (\beta_{10}', \beta_{20}')'$, $Z_i = (Z_{1i}', Z_{2i}')'$, $v = (v_1', v_2')'$ and $w = (w_1', w_2')'$. Chen and Little [1] discussed a model similar to model (2.1) in the context of survival analysis.

In the following, we will mainly consider the inference about β_0. For this, we assume that C_i may depend on Z_i. But conditional on Z_i, C_i is independent of $X_i(\cdot)$, and for technical reasons, we assume that $\Pr(C_i \geq \tau\,|\,Z_i) > 0$.

3. Estimation Procedures

3.1. Independent observation times

We first consider the situations where $N_i(\cdot)$ is independent of $\{X_i(\cdot), Z_i, C_i\}$. Note that the observed data for the ith subject consist of observations on $X_i(\cdot)$ at the time points where $N_i(\cdot)$ jumps before the censoring time C_i and Z_i. Define $EN_i(t) = \Lambda(t)$ and $Y_i(t) = I(C_i \geq t)$, where $I(\cdot)$ is the indicator function. Also define

$$X_i^*(t) = \int_0^t Y_i(s) X_i(s) dN_i(s)$$

and

$$M_i(t; \beta, \Lambda) = X_i^*(t) - \int_0^t Y_i(s) \Phi(\mu_0(s), \beta, Z_i) d\Lambda(s).$$

Under model (2.1), $M_i(t; \beta_0, \Lambda)$ $(i = 1, ..., n)$ are zero-mean stochastic processes.

For the unknown mean function $\Lambda(t)$, a natural estimator is given by the Nelson-Aalen estimator

$$\hat{\Lambda}(t) = \sum_{i=1}^{n} \int_0^t \frac{Y_i(s) dN_i(s)}{\sum_{i=1}^{n} Y_i(s)}$$

(Lawless et al. [6]). Next, as $EM_i(t; \beta_0, \hat{\Lambda}) = 0$, it is natural to estimate $\mu_0(t)$, based on a given β, by the following equation:

$$\sum_{i=1}^{n} \left\{ X_i^*(t) - \int_0^t Y_i(s) \Phi(\mu_0(s), \beta, Z_i) d\hat{\Lambda}(s) \right\} = 0, \quad 0 \leq t \leq \tau. \qquad (3.1)$$

Denote the above estimator by $\hat{\mu}_0(t; \beta)$. Then consider the estimation of β_0. Following the idea used in Cheng and Wei [2] and Lin et al. [12], we propose the following class of estimating functions:

$$U_a(\beta) = \sum_{i=1}^{n} \int_0^\tau Q(t) Z_i \{ dX_i^*(t) - Y_i(t) \Phi(\hat{\mu}_0(t; \beta), \beta, Z_i) d\hat{\Lambda}(t) \}$$

for β_0, where $Q(t)$ is a known weight function that can have various forms. A simple but natural choice is $Q(t) = 1$, which will be referred to as the log-rank weight function below. Another choice, commonly used in survival analysis, is $Q(t) = n^{-1} \sum_{i=1}^{n} Y_i(t)$, which will be referred to as the Gehan weight function.

Let $\hat{\beta}_a$ denote the solution to $U_a(\beta) = 0$ and $\hat{\mu}_0(t) = \hat{\mu}_0(t; \hat{\beta}_a)$ the corresponding estimator of the baseline mean function $\mu_0(t)$. It can be shown by following the discussion in Lin et al. [12] that $\hat{\mu}_0(t; \hat{\beta}_a)$ and $\hat{\beta}_a$ always exist and are unique for large n. It is also easy to obtain $\hat{\beta}_a$ and $\hat{\mu}_0(t)$ numerically. Specifically, let $0 = t_1 < t_2 < \cdots < t_K < t_{K+1} = \tau$ be the set of all distinct observed time points where the counting processes $N_i(\cdot)$'s jump before the censoring time C_i and τ $(i = 1, \ldots, n)$. It is clear from (3.1) that $\hat{\mu}_0(t; \beta)$ is a step function that jumps only at $\{t_i, i = 0, \ldots, K+1\}$, and $\{\hat{\mu}_0(t_i; \beta), i = 0, \ldots, K+1\}$ are identical to $\{\hat{\mu}_0(t; \beta), 0 \le t \le \tau\}$. Thus, equation (3.1) and $U_a(\beta) = 0$ are respectively equivalent to

$$\sum_{i=1}^{n} \{\Delta X_i^*(t_l) - Y_i(t_l)\Phi(\mu_0(t_l), \beta, Z_i)\Delta\hat{\Lambda}(t_l)\} = 0, \quad l = 0, \ldots, K+1 \tag{3.2}$$

and

$$\sum_{i=1}^{n}\sum_{l=0}^{K+1} Q(t_l)Z_i\{\Delta X_i^*(t_l) - Y_i(t_l)\Phi(\mu_0(t_l), \beta, Z_i)\Delta\hat{\Lambda}(t_l)\} = 0, \tag{3.3}$$

where $\Delta X_i^*(t)$ and $\Delta\hat{\Lambda}(t)$ denote the jump values of $X_i^*(t)$ and $\hat{\Lambda}(t)$ at t, respectively. The estimating equations (3.2) and (3.3) can be solved by any standard root finding method or the standard Newton-Raphson algorithm.

In Appendix, we will first prove that $n^{-1/2}U_a(\beta_0)$ is asymptotically normal with mean zero and a covariance matrix that can be consistently estimated by

$$\hat{\Sigma}_a = n^{-1}\sum_{i=1}^{n}\left[\int_0^{\tau} Q(t)\{Z_i - \overline{Z}(t; \hat{\beta}_a)\}d\hat{M}_i(t) - \int_0^{\tau}\frac{R(t; \hat{\beta}_a)Y_i(t)d\hat{M}_i^*(t)}{\sum_{j=1}^{n}Y_j(t)}\right]^{\otimes 2},$$

where $v^{\otimes 2} = vv'$ for a column vector v, $\hat{M}_i^*(t) = N_i(t) - \hat{\Lambda}(t)$,

$$\hat{M}_i(t) = X_i^*(t) - \int_0^t Y_i(s)\Phi(\hat{\mu}_0(s), \hat{\beta}_a, Z_i)d\hat{\Lambda}(s),$$

$$R(t; \beta) = \sum_{i=1}^n Q(t)\{Z_i - \overline{Z}(t; \beta)\}Y_i(t)\Phi(\hat{\mu}_0(t; \beta), \beta, Z_i),$$

$$\overline{Z}(t; \beta) = \frac{\sum_{i=1}^n Y_i(t)\phi_1(\hat{\mu}_0(t; \beta), \beta, Z_i)Z_i}{\sum_{i=1}^n Y_i(t)\phi_1(\hat{\mu}_0(t; \beta), \beta, Z_i)},$$

and $\phi_1(u, v, w) = \partial\Phi(u, v, w)/\partial u$. Following this result, we can show that $n^{1/2}(\hat{\beta}_a - \beta_0)$ is asymptotically normal with a zero mean and a covariance matrix that can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}_a\hat{A}'^{-1}$, where

$$\hat{A} = n^{-1}\sum_{i=1}^n \int_0^\tau Q(t)Y_i(t)\{Z_i - \overline{Z}(t; \hat{\beta}_a)\}\phi_2(\hat{\mu}_0(t), \hat{\beta}_a, Z_i)d\hat{\Lambda}(t),$$

and $\phi_2(u, v, w) = \partial\Phi(u, v, w)/\partial v'$. The proof is given in Appendix.

3.2. Dependent observation times

Now we turn to the case in which the observation frequencies $N_i(\cdot)$ may depend on Z_i, but conditional on Z_i, $N_i(\cdot)$ is independent of $\{X_i(\cdot), C_i\}$. Assume that, conditional on Z_i, the mean function of the $N_i(\cdot)$ has the form

$$E\{N_i(t) | Z_i\} = \Lambda_0(t)\exp(\gamma_0' Z_i), \tag{3.4}$$

where $\Lambda_0(t)$ is an unspecified baseline mean function and γ_0 is a p-vector of unknown regression parameters. Define

$$\widetilde{M}_i(t; \beta, \gamma, \Lambda_0) = X_i^*(t) - \int_0^t Y_i(s)\Phi(\mu_0(s), \beta, Z_i)\exp(\gamma' Z_i)d\Lambda_0(s).$$

Under models (2.1) and (3.4), $\widetilde{M}_i(t;\ \beta_0,\ \gamma_0,\ \Lambda_0)$ $(i = 1, \ldots, n)$ are zero-mean stochastic processes.

For the current situation, given γ, $\Lambda_0(t)$ can be estimated by

$$\hat{\Lambda}_0(t;\ \gamma) = \sum_{i=1}^{n} \int_0^t \frac{Y_i(s) dN_i(s)}{\sum_{i=1}^{n} Y_i(s) \exp(\gamma' Z_i)}$$

(Lawless et al. [6]). Given β and γ, we can estimate $\mu_0(t)$ by the solution to

$$\sum_{i=1}^{n} \left\{ X_i^*(t) - \int_0^t Y_i(s) \Phi(\mu_0(s),\ \beta,\ Z_i) \exp(\gamma' Z_i) d\hat{\Lambda}_0(s;\ \gamma) \right\} = 0$$

for $0 \le t \le \tau$. Denote the above estimator by $\tilde{\mu}_0(t;\ \beta,\ \gamma)$. Motivated by U_a, to estimate β_0 under the current situation, we propose to use the class of estimating equations $U_b(\beta,\ \gamma) = 0$, where

$$U_b(\beta,\ \gamma)$$

$$= \sum_{i=1}^{n} \int_0^{\tau} Q(t) Z_i \{ dX_i^*(t) - Y_i(t) \Phi(\tilde{\mu}_0(t;\ \beta,\ \gamma),\ \beta,\ Z_i) \exp(\gamma' Z_i) d\hat{\Lambda}_0(t;\ \gamma) \}.$$

For the nuisance parameter γ in U_b, one can estimate it by using the following partial likelihood sore equation:

$$U(\gamma) = \sum_{i=1}^{n} \int_0^{\tau} Y_i(t) \{ Z_i - \overline{Z}^*(t;\ \gamma) \} dN_i(t) = 0$$

(Sun and Wei [16]), where

$$\overline{Z}^*(t;\ \gamma) = \frac{\sum_{i=1}^{n} Y_i(t) \exp(\gamma' Z_i) Z_i}{\sum_{i=1}^{n} Y_i(t) \exp(\gamma' Z_i)}.$$

Let $\hat{\beta}_b$ and $\hat{\gamma}$ denote the estimators given by $U_b(\beta,\ \gamma) = 0$ and $U(\gamma) = 0$. Given these, the baseline mean function $\mu_0(t)$ can be estimated by $\tilde{\mu}_0(t) = \tilde{\mu}_0(t;\ \hat{\beta}_b,\ \hat{\gamma})$. As discussed in Section 3.1, $\tilde{\mu}_0(t;\ \beta,\ \gamma)$ and $\hat{\beta}_b$

always exist and are unique for large n and it is also easy to obtain $\hat{\beta}_b$ and $\tilde{\mu}_0(t)$ numerically in place of $\hat{\beta}_a$ and $\hat{\mu}_0(t)$.

To investigate the asymptotic property of $\hat{\beta}_b$, again we first prove in the Appendix that $n^{-1/2}(U_b'(\beta_0, \gamma_0), U'(\gamma_0))'$ is asymptotically normal with mean zero and a covariance matrix that can be consistently estimated by

$$\hat{\Sigma}_b = \begin{pmatrix} \hat{\Sigma}_{11} & \hat{\Sigma}_{12} \\ \hat{\Sigma}_{12}' & \hat{\Sigma}_{22} \end{pmatrix},$$

where

$$\hat{\Sigma}_{11} = n^{-1}\sum_{i=1}^{n}\left[\int_0^\tau Q(t)\{Z_i - \tilde{Z}(t; \hat{\beta}_b, \hat{\gamma})\}d\tilde{M}_i^*(t) - \int_0^\tau \frac{\tilde{R}(t; \hat{\beta}_b, \hat{\gamma})Y_i(t)d\hat{M}_i^{**}(t)}{\sum_{j=1}^n Y_j(t)\exp(\hat{\gamma}'Z_j)}\right]^{\otimes 2},$$

$$\hat{\Sigma}_{22} = n^{-1}\sum_{i=1}^{n}\left[\int_0^\tau Y_i(t)\{Z_i - \bar{Z}^*(t; \hat{\gamma})\}d\hat{M}_i^{**}(t)\right]^{\otimes 2},$$

and

$$\hat{\Sigma}_{12} = n^{-1}\sum_{i=1}^{n}\left[\left(\int_0^\tau Q(t)\{Z_i - \tilde{Z}(t; \hat{\beta}_b, \hat{\gamma})\}d\tilde{M}_i^*(t) - \int_0^\tau \frac{\tilde{R}(t; \hat{\beta}_b, \hat{\gamma})Y_i(t)d\hat{M}_i^{**}(t)}{\sum_{j=1}^n Y_j(t)\exp(\hat{\gamma}'Z_j)}\right)\right.$$

$$\left.\times\int_0^\tau Y_i(t)\{Z_i - \bar{Z}^*(t; \hat{\gamma})\}' d\hat{M}_i^{**}(t)\right],$$

with

$$\hat{M}_i^{**}(t) = N_i(t) - \exp(\hat{\gamma}'Z_i)\hat{\Lambda}_0(t, \hat{\gamma}),$$

$$\tilde{M}_i^*(t) = X_i^*(t) - \int_0^t Y_i(s)\Phi(\tilde{\mu}_0(s), \hat{\beta}_b, Z_i)\exp(\hat{\gamma}'Z_i)d\hat{\Lambda}_0(s; \hat{\gamma}),$$

$$\tilde{R}(t; \beta, \gamma) = \sum_{i=1}^n Q(t)\{Z_i - \tilde{Z}(t; \beta, \gamma)\}Y_i(t)\Phi(\tilde{\mu}_0(t; \beta, \gamma), \beta, Z_i)\exp(\gamma'Z_i),$$

and

$$\widetilde{Z}(t;\, \beta,\, \gamma) = \frac{\sum_{i=1}^{n} Y_i(t)\phi_1(\widetilde{\mu}_0(t;\, \beta,\, \gamma),\, \beta,\, Z_i)\exp(\gamma' Z_i) Z_i}{\sum_{i=1}^{n} Y_i(t)\phi_1(\widetilde{\mu}_0(t;\, \beta,\, \gamma),\, \beta,\, Z_i)\exp(\gamma' Z_i)}.$$

Then it can be shown that $n^{1/2}(\hat{\beta}_b - \beta_0)$ has an asymptotic normal distribution with mean zero and a covariance matrix that can be consistently estimated by

$$(\hat{B}^{-1},\, -\hat{B}^{-1}\hat{C})\hat{\Sigma}_b(\hat{B}^{-1},\, -\hat{B}^{-1}\hat{C})',$$

where

$$\hat{B} = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Q(t)Y_i(t)\{Z_i - \widetilde{Z}(t;\, \hat{\beta}_b,\, \hat{\gamma})\}\phi_2(\widetilde{\mu}_0(t),\, \hat{\beta}_b,\, Z_i)\exp(\hat{\gamma}' Z_i)\, d\hat{\Lambda}_0(t,\, \hat{\gamma}),$$

and $\hat{C} = \hat{C}_1\hat{C}_2^{-1}$ with

$$\hat{C}_1 = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Q(t)Y_i(t)\Phi(\widetilde{\mu}_0(t),\, \hat{\beta}_b,\, Z_i)\{Z_i - \widetilde{Z}(t;\, \hat{\beta}_b,\, \hat{\gamma})\}$$

$$\cdot\, \{Z_i - \overline{Z}^*(t;\, \hat{\gamma})\}' \exp(\hat{\gamma}' Z_i)\, d\hat{\Lambda}_0(t,\, \hat{\gamma})$$

and

$$\hat{C}_2 = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} Y_i(t)\left\{\frac{\sum_{j=1}^{n} Y_j(t)\exp(\hat{\gamma}' Z_j) Z_j^{\otimes 2}}{\sum_{j=1}^{n} Y_j(t)\exp(\hat{\gamma}' Z_j)} - \overline{Z}^*(t;\, \hat{\gamma})^{\otimes 2}\right\} dN_i(t).$$

4. Simulation Studies

Simulation studies were conducted to assess the performance of the proposed estimation procedures for various practical situations. In the studies, we focused on the situation where observation times depend on covariates through model (3.4) with $\Lambda_0(t) = 1$ and $\gamma_0 = \log 2$, and assumed that Z_i was a Bernoulli random variable with success probability 0.5. The follow-up time C_i was generated from the uniform distribution $U[3, 5]$, which resulted in an average of 6 observations per subject. Among others, we considered the following two models: the proportional mean model

$$X_i(t) = \mu_0(t)\exp(\beta_0 Z_i) + \sigma\varepsilon_i(t), \tag{4.1}$$

and the Box-Cox transformation model

$$X_i(t) = \{1 + \mu_0(t)\exp(\beta_0 Z_i)\}^2 - 1 + \sigma\varepsilon_i(t) \tag{4.2}$$

with $\mu_0(t) = 1$ and independent standard normal variables $\varepsilon_i(t)$'s. Note that the second model is a nonproportional mean model. The results reported here are based on 2000 replicates.

Table 1 below displays the means of point estimates $\hat{\beta}_b$ and their standard error estimates given in Section 3.2 along with the empirical 95% coverage probabilities under the above proportional mean model. In the table, different values of β_0, σ and sample size n were considered and both log-rank and Gehan weight functions were used. For comparison, the sample standard errors of $\hat{\beta}_b$ based on simulated data are also calculated and included in the table. In all cases, τ was set to be the maximum length of follow-up times.

Table 1. Simulation results based on the proportional mean model

		$n = 50$							
		Log-rank				Gehan			
β_0	σ	Mean	SE	SEE	ECP	Mean	SE	SEE	ECP
0.5	0.10	0.4999	0.0735	0.0722	0.9420	0.5006	0.0640	0.0636	0.9430
	0.25	0.5002	0.0770	0.0758	0.9410	0.5009	0.0656	0.0649	0.9440
	0.50	0.4989	0.0797	0.0784	0.9340	0.4993	0.0710	0.0696	0.9380
1.0	0.10	0.9998	0.0737	0.0721	0.9400	1.0003	0.0642	0.0631	0.9420
	0.25	1.0011	0.0748	0.0735	0.9390	1.0004	0.0651	0.0648	0.9410
	0.50	1.0003	0.0779	0.0769	0.9450	0.9992	0.0699	0.0687	0.9400

					n = 100				
0.5	0.10	0.4997	0.0521	0.0517	0.9440	0.5003	0.0456	0.0453	0.9420
	0.25	0.5006	0.0533	0.0526	0.9470	0.5008	0.0460	0.0467	0.9520
	0.50	0.4990	0.0549	0.0547	0.9420	0.4999	0.0492	0.0493	0.9510
1.0	0.10	1.0001	0.0527	0.0518	0.9410	0.9997	0.0457	0.0450	0.9460
	0.25	1.0003	0.0539	0.0530	0.9400	0.9993	0.0462	0.0469	0.9490
	0.50	1.0009	0.0548	0.0543	0.9450	0.9998	0.0494	0.0489	0.9430

Note: Mean represents the mean of the estimates of β_0, SE represents the sampling standard error of the estimates of β_0, SEE represents the mean of the standard error estimators, and ECP represents the empirical coverage probability of the 95% confidence interval.

Table 2 below presents the results for the same situations considered in Table 1, but under the Box-Cox transformation model.

Table 2. Simulation results based on the Box-Cox transformation model

					n = 50				
		Log-rank				Gehan			
β_0	σ	Mean	SE	SEE	ECP	Mean	SE	SEE	ECP
0.5	0.10	0.4991	0.0614	0.0596	0.9470	0.4993	0.0497	0.0495	0.9450
	0.25	0.5004	0.0622	0.0611	0.9430	0.4998	0.0510	0.0501	0.9550
	0.50	0.4979	0.0635	0.0624	0.9390	0.4976	0.0523	0.0514	0.9390
1.0	0.10	0.9937	0.0607	0.0590	0.9420	0.9927	0.0493	0.0490	0.9420
	0.25	0.9942	0.0616	0.0596	0.9380	0.9918	0.0528	0.0515	0.9380
	0.50	0.9925	0.0631	0.0618	0.9350	0.9906	0.0539	0.0521	0.9340
					n = 100				
0.5	0.10	0.4994	0.0422	0.0420	0.9470	0.4999	0.0351	0.0352	0.9510
	0.25	0.5005	0.0434	0.0425	0.9440	0.4986	0.0358	0.0354	0.9430
	0.50	0.4986	0.0446	0.0437	0.9420	0.4988	0.0365	0.0357	0.9400
1.0	0.10	0.9917	0.0428	0.0421	0.9450	0.9909	0.0340	0.0343	0.9440
	0.25	0.9930	0.0439	0.0433	0.9430	0.9920	0.0356	0.0349	0.9420
	0.50	0.9911	0.0448	0.0440	0.9400	0.9903	0.0369	0.0353	0.9390

It can be seen from Tables 1 and 2 that the proposed estimation procedures performed well for all situations considered under both the proportional mean model and the Box-Cox transformation model. It appears that the estimates are unbiased and there is a good agreement between the estimated and empirical standard errors. The confidence intervals have reasonable coverages and the results become better when the sample size increases from 50 to 100. We also simulated other models and the results were similar to those given above.

5. An Example of Application

In this section, we apply the proposed methodology to CD4 cell counts from protocol 116A of the AIDS Clinical Trials Group. This is a three arms randomized double-blind trial comparing three treatments, ZDV, 500mg/day DDI, and 750mg/day DDI (Dolin et al. [4]). The study began in October 1989 and ended in May 1992. Among other variables, CD4 cell counts were supposed to be measured every 8 weeks and as expected, there are a lot of missing measurements. In the analysis here, we focus on the patients for whom at least one CD4 cell count was available at week 2, 8, 12, 16, 24, 32, 40, 48, 56, or 64 after baseline CD4 cell count measurements (at week 0). In total, the data set includes 1302 observations on CD4 cell counts from 156 patients.

In the study, a binary covariate DXAIDS was measured, which equals 1 if a patient was diagnosed with AIDS at entry to the study and 0 otherwise. To investigate the effect of DXAIDS on CD4 cell counts, we applied the proposed method to estimate the regression parameter vector β_0. The proportional mean model (4.1) and the Box-Cox transformation model (4.2) are studied, each with the log-rank and Gehan weight functions. The estimates $\hat{\beta}_a$, together with the estimated standard error (ESE), are summarized in Table 3 below.

Table 3. The estimates $\hat{\beta}_a$ of the regression parameter and their estimated standard errors (in bracket) for the CD4 cell count data

Weight	Model (4.1)	Model (4.2)
Log-rank	–0.5151 (0.1646)	–0.2807 (0.0825)
Gehan	–0.4848 (0.1623)	–0.2637 (0.0805)

From Table 3 we see that, under the proportional mean model (4.1), $\hat{\beta}_a = -0.5151$ with ESE $= 0.1646$ using the log-rank weight function, and $\hat{\beta}_a = -0.4848$ with ESE $= 0.1623$ using the Gehan weight function. Both results suggest that DXAIDS has a significant effect on CD4 cell counts and that the patients with AIDS diagnosed at the beginning tended to have lower CD4 cell counts than the patients free of AIDS. It also shows that estimate $\hat{\beta}_a$ is reasonably robust against the weight function, as the estimates and their estimated standard errors are close between the two weight functions.

Similar conclusions can be drawn from the Box-Cox transformation model (4.2) as well, where $\hat{\beta}_a = -0.2807$ and -0.2637 respectively with the log-rank and Gehan weight functions, and the corresponding estimated standard errors are 0.0825 and 0.0805, respectively. These results again indicate that CD4 cell counts were significantly different between patients diagnosed with AIDS and free of AIDS. Note that, however, we did not consider treatment effect in the above analyses, since the final analysis of the study showed that there were no significant treatment differences (Dolin et al. [4]). Also, the study suggested that it is reasonable to assume that observation times are independent of DXAIDS.

6. Concluding Remarks

We have discussed the regression analysis of longitudinal data under general semiparametric transformation models, which include many existing regression models for panel count data and longitudinal data as special cases. Estimation procedures are proposed for the regression parameters that allow observation times to depend on covariates. Simulation studies show that the methods work well for the situations considered. The methodology was applied to a set of longitudinal data about CD4 cell counts from an AIDS clinical trial. We considered only time-independent covariates in this paper, but it is straightforward to generalize the proposed methods to situations with time-dependent covariates.

There are several open questions for future researches. One is the selection of weight functions, which is usually a complicated problem (Lin and Ying [9]). It would be useful to investigate how to find a weight function that gives the optimal efficiency of the proposed estimate of regression parameters if it exists. Another interesting problem is the development of model-checking procedures for model (2.1). Since model (2.1) allows many choices, it would be helpful to have methods available for selecting a valid or best-fitting model. A third open question is the asymptotic distribution of the proposed estimators of the baseline mean function $\mu_0(t)$, which requires further investigations. Also it would be useful to generalize the proposed methodology to the situation where there exist informative drop-outs (Lin and Ying [10]).

Appendix: Proofs of Asymptotic Properties

In this section we prove the asymptotic properties of $U_a(\beta_0)$, $U_b(\beta_0, \gamma_0)$, $U(\gamma_0)$, $\hat{\beta}_a$ and $\hat{\beta}_b$ given in the previous sections through three theorems.

In the following, we will assume that the Z_i's are bounded and that $Q(t)$ has bounded variation and converges almost surely to $q(t)$ uniformly in $t \in [0, \tau]$. First by using the uniform strong law of large numbers (Pollard [13, p. 41]), it is not difficult to show that $\hat{\mu}_0(t; \beta)$, $\tilde{\mu}_0(t; \beta, \gamma)$, $\overline{Z}(t; \beta)$, $\widetilde{Z}(t; \beta, \gamma)$ and $\overline{Z}^*(t; \gamma)$ converge almost surely to nonrandom functions $\mu_0(t; \beta)$, $\mu_0(t; \beta, \gamma)$, $\bar{z}(t; \beta)$, $\tilde{z}(t; \beta, \gamma)$ and $\bar{z}^*(t; \gamma)$ uniformly in t, β and γ, respectively. Let

$$\mu_0(t) = \mu_0(t;\, \beta_0) = \mu_0(t;\, \beta_0,\, \gamma_0),$$

$$\bar{z}(t) = \bar{z}(t;\, \beta_0),\ \tilde{z}(t) = \tilde{z}(t;\, \beta_0,\, \gamma_0),\ \bar{z}^*(t) = \bar{z}^*(t;\, \gamma_0),$$

$$M_i(t) = M_i(t;\, \beta_0,\, \Lambda),\ \widetilde{M}_i(t) = \widetilde{M}_i(t;\, \beta_0,\, \gamma_0,\, \Lambda_0),\ M_i^*(t) = N_i(t) - \Lambda(t),$$

$$M_i^{**}(t) = N_i(t) - \exp(\gamma_0' Z_i)\Lambda_0(t),\ \pi(t) = EY_1(t),\ \tilde{\pi}(t) = E(Y_1(t)\exp(\gamma_0' Z_1)),$$

$$r(t) = E[Q(t)\{Z_1 - \bar{z}(t)\}Y_1(t)\Phi(\mu_0(t),\, \beta_0,\, Z_1)], \text{ and}$$

$$\tilde{r}(t) = E[Q(t)\{Z_1 - \tilde{z}(t)\}Y_1(t)\Phi(\mu_0(t),\, \beta_0,\, Z_1)\exp(\gamma_0' Z_1)].$$

We first study the following three processes:

$$U_a(t;\, \beta) = \sum_{i=1}^{n}\int_0^t Q(s)Z_i[dX_i^*(s) - Y_i(s)\Phi(\hat{\mu}_0(s;\, \beta),\, \beta,\, Z_i)d\hat{\Lambda}(s)],$$

$$U_b(t;\, \beta,\, \gamma)$$

$$= \sum_{i=1}^{n}\int_0^t Q(s)Z_i[dX_i^*(s) - Y_i(s)\Phi(\tilde{\mu}_0(s;\, \beta,\, \gamma),\, \beta,\, Z_i)\exp(\gamma' Z_i)d\hat{\Lambda}_0(s;\, \gamma)],$$

$$U(t;\, \gamma) = \sum_{i=1}^{n}\int_0^t Y_i(s)\{Z_i - \overline{Z}^*(s;\, \gamma)\}dN_i(s).$$

Their asymptotic properties are shown below:

Theorem 1. *Under the above assumptions,* $n^{-1/2}U_a(t;\, \beta_0)$ *and* $n^{-1/2}(U_b'(t;\, \beta_0,\, \gamma_0),\, U'(t;\, \gamma_0))'$ *converge weakly, as* $n \to \infty$, *to zero-mean Gaussian processes with covariance functions*

$$\Gamma_a(t_1,\, t_2) = E\Bigg[\left(\int_0^{t_1} q(s)\{Z_1 - \bar{z}(s)\}dM_1(s) - \int_0^{t_1}\frac{r(s)Y_1(s)dM_1^*(s)}{\pi(s)}\right)$$

$$\times\left(\int_0^{t_2} q(s)\{Z_1 - \bar{z}(s)\}dM_1(s) - \int_0^{t_1}\frac{r(s)Y_1(s)dM_1^*(s)}{\pi(s)}\right)'\Bigg],$$

and

$$\Gamma_b(t_1, t_2) = \begin{pmatrix} \Gamma_{11}(t_1, t_2) & \Gamma_{12}(t_1, t_2) \\ \Gamma'_{12}(t_2, t_1) & \Gamma_{22}(t_1, t_2) \end{pmatrix}$$

respectively, where

$$\Gamma_{11}(t_1, t_2) = E\left[\left(\int_0^{t_1} q(s)\{Z_1 - \tilde{z}(s)\}d\tilde{M}_1(s) - \int_0^{t_1} \frac{\tilde{r}(s)Y_1(s)dM_1^{**}(s)}{\tilde{\pi}(s)}\right)\right.$$

$$\left.\times\left(\int_0^{t_2} q(s)\{Z_1 - \tilde{z}(s)\}d\tilde{M}_1(s) - \int_0^{t_2} \frac{\tilde{r}(s)Y_1(s)dM_1^{**}(s)}{\tilde{\pi}(s)}\right)'\right],$$

$$\Gamma_{22}(t_1, t_2)$$

$$= E\left[\int_0^{t_1} Y_1(s)\{Z_1 - \bar{z}^*(s)\}dM_1^{**}(s)\int_0^{t_2} Y_1(s)\{Z_1 - \bar{z}^*(s)\}' dM_1^{**}(s)\right],$$

$$\Gamma_{12}(t_1, t_2) = E\left[\left(\int_0^{t_1} q(s)\{Z_1 - \tilde{z}(s)\}d\tilde{M}_1(s) - \int_0^{t_1} \frac{\tilde{r}(s)Y_1(s)dM_1^{**}(s)}{\tilde{\pi}(s)}\right)\right.$$

$$\left.\times\int_0^{t_2} Y_1(s)\{Z_1 - \bar{z}^*(s)\}' dM_1^{**}(s)\right].$$

Proof. Using the linear expansion of $\Phi(u, v, w)$ about u in $U_a(t; \beta_0)$, we have

$$n^{-1/2}U_a(t; \beta_0) = n^{-1/2}\sum_{i=1}^{n}\int_0^t Q(s)Z_i[dM_i(s; \beta_0, \hat{\Lambda}) - Y_i(s)\phi_1(\mu^*(s), \beta_0, Z_i)$$

$$\times\{\hat{\mu}_0(s; \beta_0) - \mu_0(s)\}d\hat{\Lambda}(s)], \tag{A.1}$$

where $\mu^*(t)$ lies between $\hat{\mu}_0(t; \beta_0)$ and $\mu_0(t)$. Note that $\hat{\mu}_0(t; \beta)$ satisfies

$$\sum_{i=1}^{n}\left[X_i^*(t) - \int_0^t Y_i(t)\Phi(\hat{\mu}_0(t; \beta), \beta, Z_i)d\hat{\Lambda}(s)\right] = 0. \tag{A.2}$$

Then the linear expansion yields

$$[\hat{\mu}_0(t;\, \beta_0) - \mu_0(t)]d\hat{\Lambda}(t) = \frac{\sum_{i=1}^{n} dM_i(t;\, \beta_0,\, \hat{\Lambda})}{\sum_{i=1}^{n} Y_i(t)\phi_1(\mu^{**}(t),\, \beta_0,\, Z_i)}, \tag{A.3}$$

where $\mu^{**}(t)$ also lies between $\hat{\mu}_0(t;\, \beta_0)$ and $\mu_0(t)$. It follows from the functional central limit theorem (Pollard [13, p. 53]) and $\sup_{0 \le t \le \tau} |\, \hat{\Lambda}(t) - \Lambda(t)\,| = O_p(n^{-1/2})$ that $\sum_{i=1}^{n} M_i(t;\, \beta_0,\, \hat{\Lambda})$ is $O_p(n^{1/2})$ uniformly in t. Hence, combining (A.1) and (A.3) with the uniform convergence of $\hat{\mu}_0(t;\, \beta_0)$, we obtain

$$
\begin{aligned}
& n^{-1/2} U_a(t;\, \beta_0) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_0^t Q(s) \left[Z_i - \frac{\sum_{j=1}^{n} Y_j(s)\phi_1(\mu^{*}(s),\, \beta_0,\, Z_j) Z_j}{\sum_{j=1}^{n} Y_j(t)\phi_1(\mu^{**}(s),\, \beta_0,\, Z_j)} \right] dM_i(s;\, \beta_0,\, \hat{\Lambda}) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_0^t Q(s)[Z_i - \bar{z}(s)] dM_i(s;\, \beta_0,\, \hat{\Lambda}) + o_p(1) \\
&= U_{a1}(t) - U_{a2}(t) + o_p(1),
\end{aligned}
$$

where

$$U_{a1}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t Q(s)\{Z_i - \bar{z}(s)\} dM_i(s),$$

$$U_{a2}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t Q(s)\{Z_i - \bar{z}(s)\} Y_i(s) \Phi(\hat{\mu}_0(s;\, \beta_0),\, \beta_0,\, Z_i) d[\hat{\Lambda}(s) - \Lambda(s)].$$

Using tools from the modern empirical process theory (e.g., Pollard [13]), we can get

$$U_{a1}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t q(s)\{Z_i - \bar{z}\} dM_i(s) + o_p(1).$$

Note that

$$\hat{\Lambda}(t) - \Lambda(t) = \sum_{i=1}^{n} \int_0^t \frac{Y_i(s) dM_i^*(s)}{\pi(s)} + o_p(n^{-1/2}),$$

which implies

$$U_{a2}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \frac{r(s) Y_i(s) dM_i^*(s)}{\pi(s)} + o_p(1).$$

Thus,

$$n^{-1/2} U_a(t;\, \beta_0) = n^{-1/2} \sum_{i=1}^{n} \Psi(t) + o_p(1), \tag{A.4}$$

where

$$\Psi_i(t) = \int_0^t q(s) \{Z_i - \bar{z}(s)\} dM_i(s) - \int_0^t \frac{r(s) Y_i(s) dM_i^*(s)}{\pi(s)}.$$

Because $\Psi_i(t)$ $(i = 1, \ldots, n)$ are i.i.d. zero-mean random variables for any fixed t, the finite-dimensional normality of (A.4) follows from the multivariate central limit theorem. Using similar arguments to those in the proof of Theorem 1 of Lin et al. [11], we can show that (A.4) is tight and thus converges weakly to a zero-mean Gaussian process with covariance function $E(\Psi_1(t_1)\Psi_1(t_2)) = \Gamma_a(t_1, t_2)$ at (t_1, t_2). In a similar manner and using

$$[\tilde{\mu}_0(t;\, \beta_0,\, \gamma_0) - \mu_0(t)] d\hat{\Lambda}_0(t;\, \gamma_0) = \frac{\sum_{i=1}^{n} d\tilde{M}_i(t;\, \beta_0,\, \gamma_0,\, \hat{\Lambda})}{\sum_{i=1}^{n} Y_i(t) \phi_1(\tilde{\mu}^*(t),\, \beta_0,\, Z_i) \exp(\gamma_0' Z_i)},$$

and

$$\hat{\Lambda}_0(t;\, \gamma_0) - \Lambda_0(t) = \sum_{i=1}^{n} \int_0^t \frac{Y_i(s) dM_i^{**}(s;\, \gamma_0)}{\tilde{\pi}(s)} + o_p(n^{-1/2}),$$

where $\tilde{\mu}^*(t)$ lies between $\tilde{\mu}_0(t;\,\beta_0,\,\gamma_0)$ and $\mu_0(t)$, we have

$$n^{-1/2}U_b(t;\,\beta_0,\,\gamma_0)$$

$$= n^{-1/2}\sum_{i=1}^{n}\left[\int_0^t q(s)\{Z_i - \tilde{z}(s)\}d\widetilde{M}_i(s) - \int_0^t \frac{\tilde{r}(s)Y_i(s)dM_i^{**}(s)}{\tilde{\pi}(s)}\right] + o_p(1),$$

$$n^{-1/2}U(t;\,\gamma_0) = n^{-1/2}\sum_{i=1}^{n}\int_0^t Y_i(s)\{Z_i - \bar{z}^*(s)\}dM_i^{**}(s) + o_p(1).$$

Therefore, $n^{-1/2}(U_b'(t;\,\beta_0,\,\gamma_0),\,U'(t;\,\gamma_0))'$ also converges weakly to zero-mean Gaussian processes with covariance function $\Gamma_b(t_1,\,t_2)$ at $(t_1,\,t_2)$. This completes the proof of Theorem 1.

Next, we study the asymptotic properties of $\hat{\beta}_a$ and $\hat{\beta}_b$. To do so, let $\theta = (\beta',\,\gamma')'$, $\theta_0 = (\beta_0',\,\gamma_0')'$, and $U_b(\beta,\,\gamma) = U_b(\theta)$, and we need to establish the asymptotic linearity of $U_a(\beta)$ and $U_b(\theta)$ in $\mathcal{N}(\beta_0)$ and $\mathcal{N}^*(\theta_0)$, respectively, where $\mathcal{N}^*(\theta_0)$ is a compact neighborhood of θ_0. The results are given in the next theorem.

Theorem 2. *Suppose that the above assumptions hold and* $n \to \infty$. *Then for any sequence* $c_n \to 0$,

$$\sup_{\|\beta-\beta_0\|\le c_n}\frac{\|U_a(\beta) - U_a(\beta_0) + nA(\beta-\beta_0)\|}{n\|\beta-\beta_0\|} = o(1), \tag{A.5}$$

and

$$\sup_{\|\theta-\theta_0\|\le c_n}\frac{\|U_b(\theta) - U_b(\theta_0) - C(U(\gamma) - U(\gamma_0)) + nB(\beta-\beta_0)\|}{n\|\theta-\theta_0\|} = o(1) \tag{A.6}$$

almost surely, *where*

$$A = E\left[\int_0^\tau q(t)Y_1(t)\{Z_1 - \bar{z}(t)\}\phi_2(\mu_0(t),\,\beta_0,\,Z_1)d\Lambda(t)\right],$$

$$B = E\left[\int_0^\tau q(t)Y_1(t)\{Z_1 - \tilde{z}(t)\}\phi_2(\mu_0(t),\,\beta_0,\,Z_1)\exp(\gamma_0'Z_1)d\Lambda_0(t)\right],$$

and $C = C_1 C_2^{-1}$, *with*

$$C_1 = E\bigg[\int_0^\tau q(t) Y_1(t) \Phi(\mu_0(t), \beta_0, Z_1)$$

$$\cdot \{Z_1 - \tilde{z}(t)\}\{Z_1 - \bar{z}^*(t)\}' \exp(\gamma_0' Z_1) d\Lambda_0(t)\bigg],$$

$$C_2 = E\bigg[\int_0^\tau Y_1(t)\{Z_1 - \bar{z}^*(t)\}^{\otimes 2} \exp(\gamma_0' Z_1) d\Lambda_0(t)\bigg].$$

Furthermore, if A has full rank, then $\hat{\beta}_a$ is strongly consistent and $n^{1/2}(\hat{\beta}_a - \beta_0)$ converges in distribution to a zero-mean normal vector with covariance matrix $A^{-1}\Gamma_a(\tau, \tau)A'^{-1}$. If B has full rank, then $\hat{\beta}_b$ is strongly consistent and $n^{1/2}(\hat{\beta}_b - \beta_0)$ converges in distribution to zero-mean normal with covariance matrix $(B^{-1}, -B^{-1}C)\Gamma_b(\tau, \tau)(B^{-1}, -B^{-1}C)'$.

Proof. Let $\hat{A}(\beta) = -n^{-1}\partial U_a(\beta)/\partial\beta'$. Then

$$\hat{A}(\beta) = n^{-1}\sum_{i=1}^n \int_0^\tau Q(t) Y_i(t) Z_i\bigg[\phi_1(\hat{\mu}_0(t;\beta), \beta, Z_i)\frac{\partial\hat{\mu}_0(t;\beta)}{\partial\beta'}$$

$$+ \phi_2(\hat{\mu}_0(t;\beta), \beta, Z_i)\bigg] d\hat{\Lambda}(t).$$

Differentiate (A.2) with respect to β, we get

$$\frac{\partial\hat{\mu}_0(t;\beta)}{\partial\beta'} d\hat{\Lambda}(t) = -\frac{\sum_{i=1}^n Y_i(t)\phi_2(\hat{\mu}_0(t;\beta), \beta, Z_i)}{\sum_{i=1}^n Y_i(t)\phi_1(\hat{\mu}_0(t;\beta), \beta, Z_i)} d\hat{\Lambda}(t).$$

Thus

$$\hat{A}(\beta) = n^{-1}\sum_{i=1}^n \int_0^\tau Q(t) Y_i(t)\{Z_i - \bar{Z}(t;\beta)\}\phi_2(\hat{\mu}_0(t;\beta), \beta, Z_i) d\hat{\Lambda}(t).$$

The uniform convergence of $\hat{\mu}_0(t;\beta)$ and the continuity of $\phi_k(u, v, w)$ $(k = 1, 2)$ together with the uniform strong law of large numbers imply that $\hat{A}(\beta)$ converges almost surely to nonrandom function $A(\beta)$ uniformly in β. It is easily checked that $A(\beta_0) = A$. Hence, (A.5) immediately follows from Taylor's expansion and the continuity of $A(\beta)$. When A is nonsingular, the consistency of $\hat{\beta}_a$ follows from the arguments used in the Appendix of Lin et al. [12, p. 627] (also see Lin et al. [11]). Now it follows from (A.5) and consistency of $\hat{\beta}_a$ that $n^{1/2}(\hat{\beta}_a - \beta_0) = A^{-1}n^{-1/2}U_a(\theta_0) + o_p(1)$, which yields the asymptotic normality of $\hat{\beta}_a$ from Theorem 1. According to the proof of (A.5), we have

$$\sup_{\| \theta-\theta_0 \| \le c_n} \frac{\| U_b(\theta) - U_b(\theta_0) + nB(\beta - \beta_0) + nC_1(\gamma - \gamma_0) \|}{n\| \theta - \theta_0 \|} = o(1). \tag{A.7}$$

Using Taylor's expansion, we get

$$\sup_{\| \gamma-\gamma_0 \| \le c_n} \frac{\| U(\gamma) - U(\gamma_0) + nC_2(\gamma - \gamma_0) \|}{n\| \gamma - \gamma_0 \|} = o(1). \tag{A.8}$$

Thus, (A.6) follows from (A.7) and (A.8). Similar to $\hat{\beta}_a$, we can obtain consistency and the asymptotic normality of $\hat{\beta}_b$.

Finally, we prove the convergence for the estimators of the covariance matrices.

Theorem 3. *Under the same conditions as Theorem* 2, $\hat{\Sigma}_l \to \Gamma_k(\tau, \tau)$ $(l = a, b)$, $\hat{A} \to A$, $\hat{B} \to B$, *and* $\hat{C} \to C$ *almost surely.*

Proof. The uniform convergence of $\hat{\mu}_0(t; \beta)$, $\overline{Z}(t; \beta)$, $R(t, \beta)$ and $\hat{\Lambda}(t)$ and the strong consistency of $\hat{\beta}_a$ entail that $n^{-1}\sum_{i=1}^{n} \| \hat{D}_i - D_i \|^2 \to 0$ almost surely, where

$$\hat{D}_i = \int_0^\tau \{Q(t)\{Z_i - \overline{Z}(t; \hat{\beta}_a)\}d\hat{M}_i(t) - \int_0^\tau \frac{R(t; \hat{\beta}_a)Y_i(t)d\hat{M}_i^*(t)}{\sum_{j=1}^{n} Y_j(t)},$$

$$D_i = \int_0^\tau q(t)\{Z_i - \bar{z}(t)\}dM_i(s) - \int_0^\tau \frac{r(t)Y_i(t)dM_i^*(t)}{\pi(t)}.$$

It follows from the strong law of large numbers that $n^{-1}\sum_{i=1}^{n} D_i^{\otimes 2} \to \Gamma_a(\tau, \tau)$ almost surely. Therefore, $\hat{\Sigma}_a$ converges almost surely to $\Gamma_a(\tau, \tau)$. Similarly, $\hat{\Sigma}_b \to \Gamma_b(\tau, \tau)$, $\hat{A} \to A$, $\hat{B} \to B$, and $\hat{C} \to C$ almost surely. This completes the proof of the theorem.

References

[1] H. Y. Chen and R. J. Little, A profile conditional likelihood approach for the semiparametric transformation regression model with missing covariates, Lifetime Data Anal. 7 (2001), 207-224.

[2] S. C. Cheng and L. J. Wei, Inferences for a semiparametric model with panel data, Biometrika 87 (2000), 89-97.

[3] P. J. Diggle, K. Y. Liang and S. L. Zeger, Analysis of Longitudinal Data, Chapman & Hall, London, 1994.

[4] R. Dolin, D. A. Amato, M. A. Fischl, C. Pettinelli, M. Beltangady, S. H. Liou, M. J. Brown, A. P. Cross, M. S. Hirsch, W. D. Hardy, D. Mildvan, D. C. Blair, W. G. Powderly, M. F. Para, K. H. Fife, R. T. Steigbigel, L. Smaldone and the AIDS Clinical Trials Group, Zidovudine compared with didanosine in patients with advanced HIV type 1 infection and little or no previous experience with zidovudine, Archivex of Internal Medicine 155(9) (1995), 961-974.

[5] N. M. Laird and J. H. Ware, Random effects models for longitudinal data, Biometrics 38 (1982), 963-974.

[6] J. F. Lawless, C. Nadeau and R. J. Cook, Analysis of mean and rate functions for recurrent events, Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis, eds., D. Y. Lin and T. R. Fleming, pp. 37-49, Springer-Verlag, New York, 1997.

[7] K. Y. Liang and S. L. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986), 13-22.

[8] X. Lin and R. J. Carroll, Semiparametric regression for clustered data using generalized estimating equations, J. Amer. Statist. Assoc. 96 (2001), 1045-1056.

[9] D. Y. Lin and Z. Ying, Semiparametric and nonparametric regression analysis of longitudinal data, J. Amer. Statist. Assoc. 96 (2001), 103-113.

[10] D. Y. Lin and Z. Ying, Semiparametric regression analysis of longitudinal data with informative drop-outs, Biostatistics 4 (2003), 385-398.

[11] D. Y. Lin, L. J. Wei and Z. Ying, Accelerated failure time models for counting processes, Biometrika 85 (1998), 605-618.

[12] D. Y. Lin, L. J. Wei and Z. Ying, Semiparametric transformation models for point processes, J. Amer. Statist. Assoc. 96 (2001), 620-628.

[13] D. Pollard, Empirical Processes: Theory and Applications, Institute of Mathematical Statistics, Hayward, CA, 1990.

[14] J. Sun, A nonparametric test for current status data with unequal censoring, J. R. Stat. Soc. B 61 (1999), 243-250.

[15] J. Sun and J. D. Kalbfleisch, Estimation of the mean function of point processes based on panel count data, Statist. Sinica 5 (1995), 279-290.

[16] J. Sun and L. J. Wei, Regression analysis of panel count data with covariate-dependent observation and censoring times, J. R. Stat. Soc. B 62 (2000), 293-302.

[17] P. F. Thall and J. M. Lachin, Analysis of recurrent events: nonparametric methods for random-interval count data, J. Amer. Statist. Assoc. 83 (1988), 339-347.

[18] Y. Zhang, A semiparametric pseudolikelihood estimation method for panel count data, Biometrika 89 (2002), 39-48.

■