Regression Analysis of Longitudinal Data in the Presence of Informative Observation and Censoring Times

Jianguo SUN, Liuquan SUN, and Dandan LIU

Longitudinal data frequently occur in many types of studies, such as medical follow-up studies and observational investigations. Numerous methods for analyzing these data have been proposed (Diggle, Liang, and Zeger 1994; Laird and Ware 1982; Zeger and Diggle 1994). For example, Diggle et al. (1994) gave an excellent review of commonly used methods including estimating equation–based estimators. In this article we consider situations in which longitudinal response variables of interest may be correlated with observation times as well as censoring times. This article considers the analysis of longitudinal data where these correlations may exist and proposes a joint modeling approach that uses some latent variables to characterize the correlations. For inference about regression parameters, estimating equation approaches are developed and both large-sample and final-sample properties of the proposed estimators are established. In addition, some graphical and numerical procedures are presented for model checking. The methodology is applied to a bladder cancer study that motivated this investigation.

KEY WORDS: Estimating equation; Informative observation process; Joint modeling; Latent variables.

1. INTRODUCTION

Longitudinal data frequently occur in many types of studies, such as medical follow-up studies and observational investigations. Various methods for analyzing these data have been proposed (Diggle, Liang, and Zeger 1994; Laird and Ware 1982; Zeger and Diggle 1994). For example, Diggle et al. (1994) gave an excellent review of commonly used methods including estimating equation and random-effect model approaches, and Lin and Ying (2001) and Welsh, Lin, and Carroll (2002) proposed some clever semiparametric methods for the analysis of longitudinal data. A basic assumption behind all of these methods is that both observation and censoring times are noninformative about the longitudinal response variable of interest.

Recently, some authors have considered the situation in which there exists an informative or nonignorable dropout or informative censoring time (Huang and Wang 2004; Little 1995; Roy and Lin 2002; Sun and Song 2001; Wang and Taylor 2001; Wu and Carroll 1988; Wulfsohn and Tsiatis 1997). For example, Huang and Wang (2004) proposed a borrow-strength estimation procedure for the case where the underlying response process is a counting process and the observed data are recurrent-event data. In this case there exists an event time representing dropout or censoring time, such as death, that is related to the underlying longitudinal variable of interest and thus must be modeled together with the longitudinal variable to obtain valid inference. For methods developed for this situation, again observation times are typically assumed to be noninformative. Note that this may not be true in many situations, because the longitudinal response variable of interest may be correlated with observation times as well as censoring time.

A common situation where informative observation and censoring times occur is that these times are subject or response variable–dependent. For example, they may be hospitalization times and dropout time of subjects in the study (Lin, Scharfstein, and Rosenheck 2004; Wang, Qin, and Chiang 2001). Sun and Wei (2000) and Zhang (2002) discussed a set of longitudinal data arising from a bladder cancer follow-up study conducted by the Veterans Administration Cooperative Urological Research Group. The study involves patients with superficial bladder tumors, and one feature of the data set is that some patients in the study had significantly more clinical visits than others, suggesting that observation times may be informative. There exists some limited research that considers the situation where the longitudinal variable and observation times are correlated (Lin et al. 2004; Lipsitz, Fitzmaurice, Ibrahim, Gelber, and Lipshultz 2002; Robins, Rotnitzky, and Zhao 1995; Sun, Park, Sun, and Zhao 2005). For example, Lin et al. (2004) considered a typical marginal regression model and proposed a class of inverse intensity-of-visit process-weighted estimators, and Sun et al. (2005) proposed a joint model and developed some estimating equation–based estimators. In this article we consider situations in which longitudinal response variables of interest may be correlated with both the observation times and the censoring time.

The article is organized as follows. Section 2 discusses joint modeling of longitudinal response and a counting process that characterizes observation times. A semiparametric random-effect model is used for the longitudinal response, and a subject-specific nonstationary Poisson process is used for observation time process. The censoring process can be arbitrary, with all three processes connected through a latent variable. Section 3 presents inference procedures about regression parameters of interest with the focus on the effect of covariates on the longitudinal response. The proposed estimators have a closed form, are consistent, and have asymptotic normal distribution. The ideas behind the joint model and estimation approach are similar to those used by Huang and Wang (2004) for recurrent-event data. In Section 4 we discuss the assessment of the models described in Section 2 and give several residual-based procedures. We report some numerical results from simulation studies for evaluating the proposed methods in Section 5 that indicate that the
methods work well for situations considered. In Section 6 we apply the proposed methodology to the bladder cancer study discussed earlier, and in Section 7 we provide concluding remarks and some discussion.

2. JOINT MODELS

Consider a longitudinal study, and let \( Y(t) \) denote the longitudinal response variable of interest. Also, let \( Z \) be the \( p \times 1 \) vector of covariates, let \( C \) be the follow-up or censoring time, and let \( N(t) \) be the counting process denoting the number of the observation times before or at time \( t \). The longitudinal process \( Y(t) \) is observed only at the time points where \( N(t) \) jumps for \( t \leq C \). Note that for inference about \( Y(t) \), if \( Y(t), N(t), \) and \( C \) are independent completely or conditional on covariates, then a marginal approach is usually taken (Lin and Ying 2001). Here we consider the situation in which they may depend on each other or may be correlated; in other words, the observation times and the censoring time are informative to the response variable. Thus it seems natural and necessary to consider the joint modeling approach through a common subject-specific latent variable (frailty). Our proposed joint models are flexible in that no parametric assumptions on the distributions of latent variables and censoring times are required.

Let \( V \) be a nonnegative valued latent variable with \( E(V|Z) = 1 \). For the analysis, we will assume that given \( Z \) and \( V \), \( Y(t) \) follows the marginal model

\[
E[Y(t)|Z, V] = \mu_0(t) + \beta'Z + V,
\]

where \( \mu_0(t) \) is an unspecified smooth function of \( t \) and \( \beta \) is a vector of unknown regression parameters. The foregoing model without \( V \) has been considered by many authors, including Lin and Ying (2001). For the observation process, we assume that, conditioning on \( Z \) and \( V \), \( N(t) \) is a nonstationary Poisson process with intensity function

\[
\lambda(t|Z, V) = V\lambda_0(t) \exp(\gamma'Z),
\]

where \( \gamma \) is a vector of unknown regression parameters and \( \lambda_0(t) \) is an unknown continuous baseline intensity function. Model (2) has been used by several authors for the analysis of recurrent-event data (e.g., Huang and Wang 2004; Wang et al. 2001); however, it does not seem to exist research on models (1) and (2) together. In what follows, we assume that the censoring time \( C \) may depend on \( Z \) and \( V \) in an arbitrary way, but conditional on \( Z \) and \( V \), \( Y(t), N(t) \) and \( C \) are mutually independent.

In comparison to similar models discussed earlier, a main feature of models (1) and (2) is that they allow the dependence between the response process \( Y(t) \) and the observation process \( N(t) \). In particular, they specify a positive correlation of the two processes, which seems to be the situation for the example discussed in Section 6. We can similarly develop models and a corresponding inference procedure for the case in which the two processes are negatively correlated. We comment on this further later.

Under models (1) and (2), the distribution of \( V \), the baseline functions \( \mu_0(t) \) and \( \lambda_0(t) \), and the distribution of \( C \) serve as nonparametric components. Our main interest here is to assess covariate effects on the longitudinal process or to estimate parameters \( \beta \). Toward this end, we present an estimating equation approach in the next section.

3. ESTIMATION OF REGRESSION PARAMETERS

To estimate regression parameters in models (1) and (2), suppose that we have a random sample of \( n \) subjects. For the \( i \)th subject, let \( Y_i(t), Z_i, V_i, C_i, \) and \( N_i(t) \) be defined as earlier, but with subject \( i, i = 1, \ldots, n \). Also, let \( m_i \) denote the total number of observations and let \( T_{1i}, \ldots, T_{mi} \) denote the observation times on subject \( i \). Let \( \hat{V}_i \) be the (possibly data-dependent) weight function, \( \tau \) is the longest follow-up time, \( \Delta_i(t) = I(C_i \geq t) \), and \( \bar{S}_i(t) \) be the longi-

\[
\hat{V}_i = \frac{m_i}{\Lambda_0(C_i) \exp(\gamma'Z_i)}.
\]

where \( \Lambda_0(t) \) is a possibly data-dependent weight function, \( \tau \) is the longest follow-up time, \( \Delta_i(t) = I(C_i \geq t) \), and \( \bar{S}_i(t) \) be the longi-

\[
\hat{V}_i = \frac{m_i}{\Lambda_0(C_i) \exp(\gamma'Z_i)}.
\]
By replacing $V_i$ with $\hat{V}_i$ in the estimating function (3), we propose estimating $\beta$ using the solution to the equation

$$\hat{U}(\beta; \gamma) = n^{-1} \sum_{i=1}^{n} \int_{0}^{\tau} Q(t) \{ Z_i - \tilde{Z}(t; \hat{\gamma}) \} \times (Y_i(t) - \hat{V}_i - \beta' Z_i) \Delta_i(t) dN_i(t)$$

and

$$\tilde{Z}(t; \gamma) = \frac{\sum_{i=1}^{n} \Delta_i(t) \hat{V}_i \exp(\gamma' Z_i) Z_i}{\sum_{i=1}^{n} \Delta_i(t) \hat{V}_i \exp(\gamma' Z_i)}.$$

Equation (5) about $\beta$ has a closed-form solution given by

$$\hat{\beta} = \left[ \sum_{i=1}^{n} \int_{0}^{\tau} Q(t) \{ Z_i - \tilde{Z}(t; \hat{\gamma}) \} Z_i' \Delta_i(t) dN_i(t) \right]^{-1} \times \sum_{i=1}^{n} \int_{0}^{\tau} Q(t) \{ Z_i - \tilde{Z}(t; \hat{\gamma}) \} [Y_i(t) - \hat{V}_i] \Delta_i(t) dN_i(t).$$

It is easy to show from the law of large numbers and the consistency of $\hat{\Lambda}_0$ and $\hat{\gamma}$ that the foregoing estimator is consistent. Huang and Wang (2004) used a similar idea for the analysis of recurrent-event data in the presence of informative censoring.

To establish the asymptotic normality of $\hat{\beta}$, define

$$\hat{H}(t) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(T_{ij} \leq t),$$

$$\hat{R}(t) = n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} I(T_{ij} \leq t \leq C_i),$$

$$\hat{A}(t) = \sum_{i=1}^{n} \int_{0}^{\tau} \left\{ \frac{(Y_i(t) - \hat{V}_i) \Delta_i(u) dN_i(u)}{\sum_{i=1}^{n} \Delta_i(u) \hat{V}_i \exp(\gamma' Z_i)} \right\},$$

$$\hat{b}_i(t) = \sum_{j=1}^{m_i} \left\{ I(T_{ij} \leq u \leq C_i) d\hat{H}(u) - I(t < T_{ij} \leq \tau) \right\} \sqrt{R(t)^{-1}} \hat{R}(t_i),$$

and

$$\hat{e}_i(\alpha) = -\int w(x) \hat{b}_i(c) dP_{in}(w, x, c, m_i) / \hat{F}(c) + \frac{m_i}{\hat{F}(C_i)} - \exp(\alpha' X_i),$$

where $P_{in}(w, x, c, m_i)$ is the empirical measure of $\{(W_i, X_i, C_i, m_i), i = 1, \ldots, n\}$. Also, define $\hat{h}_i(\alpha)$ to be the vector function $[E(-\partial \hat{e}_i(\alpha) / \partial \alpha)]^{-1} \hat{e}_i(\alpha)$ without its first element, $\phi_0(\alpha)$ to be the first element of $[E(-\partial \hat{e}_i(\alpha) / \partial \alpha)]^{-1} \hat{e}_i(\alpha)$, and $\hat{\xi}_i(t) = \hat{b}_i(t) + \hat{\phi}_0(\alpha)$. Let $\beta_0$ be the true value of $\beta$. Then, as we show in Appendix A, under some regularity conditions, $n^{1/2}(\hat{\beta} - \beta_0)$ converges in distribution to a normal random vector with mean 0 and a covariance matrix that can be consistently estimated by $\hat{D}^{-1} \hat{\Sigma} \hat{D}^{-1}$, where $\hat{\Sigma} = n^{-1} \sum_{i=1}^{n} \hat{\psi}_i' \hat{\psi}_i$, $\hat{\psi}_i = \int_{0}^{\tau} Q(t) \{ Z_i - \tilde{Z}(t; \hat{\gamma}) \}$

$$\times \left\{ \frac{Y_i(t) - \beta' Z_i}{\Lambda_0(C_i)} \right\} \Delta_i(t) dN_i(t)$$

$$+ \int_{0}^{\tau} \sum_{j=1}^{m} Q(t_j) (z - \tilde{Z}(t_j; \hat{\gamma})) \frac{m_i}{\Lambda_0(c)} dP_{2n}(z; c, m_i, t_1, \ldots, t_m)$$

$$\times \{ \hat{e}_i(c) + z' \hat{h}_i(\alpha) \} dP_{3n}(z; c, m_i)$$

$$+ \int_{0}^{\tau} \frac{Q(t)}{\Lambda_0(c)} \left\{ \int I(c \geq t) (z - \tilde{Z}(t; \hat{\gamma})) \frac{m_i}{\Lambda_0(c)} d\hat{A}(t) - \int_{0}^{\tau} \frac{Q(t)}{\Lambda_0(C_i)} d\hat{A}(t) \right\},$$

and

$$\hat{D} = n^{-1} \sum_{i=1}^{n} \left\{ \sum_{j=1}^{n} \Delta_j(t) \hat{V}_j \exp(\gamma' Z_j) \hat{Z}^{\otimes 2}(t; \hat{\gamma}) - \hat{Z}(t; \hat{\gamma}) \hat{Z}^{\otimes 2}(t; \hat{\gamma}) \right\}$$

$$\times \left\{ \sum_{j=1}^{n} \Delta_j(t) \hat{V}_j \exp(\gamma' Z_j) \right\} d\hat{A}(t) - \hat{A}(t),$$

and

$$\hat{D} = n^{-1} \sum_{i=1}^{n} \left\{ \sum_{j=1}^{n} \Delta_j(t) \hat{V}_j \exp(\gamma' Z_j) \right\} d\hat{A}(t) - \hat{A}(t).$$

In the foregoing, $a \otimes^2 = aa'$ for a vector $a$, and $P_{2n}(z, c, m_i, t_1, \ldots, t_m)$ and $P_{3n}(z, c, m_i)$ denote the empirical measures of $\{(Z_i, C_i, m_i, T_{i1}, \ldots, T_{im}), i = 1, \ldots, n\}$ and $\{(Z_i, C_i, m_i), i = 1, \ldots, n\}$.

### 4. PROCEDURES FOR GOODNESS-OF-FIT TESTS

This section considers the assessment of the models described in Section 2, with the focus on model (1). To check model (2), we have complete recurrent-event data and can use some discussion and simple approaches of Huang and Wang (2004). In general, several questions about an assumed regression model may be of interest. One common question is the functional form for a particular component of covariates, and another is the overall fit of the model. Model (1) can be considered a generalized linear model with the identical link function, and thus we may be interested in testing the identical link function. In what follows we discuss these three questions and describe some graphical and numerical procedures for each question.

For each $i$, following Lin, Wei, Yang, and Ying (2000) and Pan and Lin (2005), we define the martingale residual

$$\hat{M}_i(t) = \int_{0}^{\tau} \left\{ \frac{[Y_i(u) - \beta' Z_i - \hat{V}_i] \Delta_i(u) dN_i(u)}{\hat{A}(u)} \right\}$$

$$- \Delta_i(u) \hat{V}_i \exp(\gamma' Z_i) d\hat{A}(u),$$

where $i = 1, \ldots, n$. First, we consider checking the functional form for the $k$th component of $Z_i$; for this, it is apparent that we

$$\hat{M}_i(t) = \int_{0}^{\tau} \left\{ \frac{[Y_i(u) - \beta' Z_i - \hat{V}_i] \Delta_i(u) dN_i(u)}{\hat{A}(u)} \right\}$$

$$- \Delta_i(u) \hat{V}_i \exp(\gamma' Z_i) d\hat{A}(u),$$

where $i = 1, \ldots, n$. First, we consider checking the functional form for the $k$th component of $Z_i$; for this, it is apparent that we
can simply plot \( \hat{M}_i(t) \) against \( Z_{ik} \). To develop a more formal procedure, we define
\[
\mathcal{F}_k(r) = n^{-1/2} \sum_{i=1}^{n} I(Z_{ik} \leq r) \hat{M}_i(t),
\]
the cumulative sum of \( \hat{M}_i(t) \) over the values of \( Z_{ik} \). Also, we define
\[
S_k(t, r) = n^{-1} \sum_{i=1}^{n} I(Z_{ik} \leq r) \Delta_i(t) \hat{V}_i \exp(\hat{g}' Z_i),
\]
\[
S^{(0)}(t) = n^{-1} \sum_{i=1}^{n} \Delta_i(t) \hat{V}_i \exp(\hat{g}' Z_i),
\]
\[
B_k(t, r) = n^{-1} \sum_{i=1}^{n} \int_{0}^{t} \left\{ I(Z_{ik} \leq r) - \frac{S_k(u, r)}{S^{(0)}(u)} \right\} \times Z_i \Delta_i(u) dN_i(u),
\]
and
\[
B_z(t, r) = n^{-1} \sum_{i=1}^{n} \int_{0}^{t} \left\{ I(Z_{ik} \leq r) - \frac{S_k(u, r)}{S^{(0)}(u)} \right\} \times Z_i \Delta_i(u) dN_i(u),
\]
In what follows, we assume that the limits of \( S_k(t, r) \), \( S^{(0)}(t) \), \( B_k(t, r) \), and \( B_z(t, r) \) exist and are denoted by \( s_k(t, r) \), \( s^{(0)}(t) \), \( b_k(t, r) \), and \( b_z(t, r) \).

To apply the statistic \( \mathcal{F}_k(r) \), the Appendix B shows that its null distribution can be approximated by the mean-0 Gaussian process
\[
\tilde{\mathcal{F}}_k(r) = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} \left\{ I(Z_{ik} \leq r) - \frac{s_k(u, r)}{s^{(0)}(u)} \right\} d\hat{M}_i(u)
\]
\[
+ n^{-1/2} \sum_{i=1}^{n} \int \sum_{i=1}^{m} \left\{ I(z_k \leq r) - \frac{s_k(t, r)}{s^{(0)}(t)} \right\} \times v^* \left[ \hat{g}(c; \alpha_0) + z' \eta_i(\alpha_0) \right] dP_2(z, c, m, t_1, \ldots, t_m)
\]
\[
+ n^{-1/2} \sum_{i=1}^{n} \int \sum_{i=1}^{m} \left\{ I(z_k \leq r) - \frac{s_k(u, r)}{s^{(0)}(u)} \right\} \times \exp(g_0(z) dP_3(z, c, m)) d\hat{A}(u)
\]
\[
b_1(\tau, r) D^{-1} n^{-1/2} \sum_{i=1}^{n} \hat{\psi}_i
\]
\[
b_2(\tau, r) n^{-1/2} \sum_{i=1}^{n} \eta_i(\alpha_0).
\]

In this formula, \( z_k \) is the \( k \)th component of \( z \); \( \mathcal{A}(t) = \int_{0}^{t} \mu_0(u) d\Lambda_0(u) \);
\[
M_i(t) = \int_{0}^{t} \left[ Y_i(u) - \beta_0 Z_i - \frac{m_i}{\Lambda_0(C_i) \exp(g_0 Z_i)} \right] \times \Delta_i(u) dN_i(u) - \frac{\Lambda_i(u) m_i}{\Lambda_0(C_i)} d\hat{A}(u).
\]
\[
v^* = m/[\Lambda_0(c) \exp(g_0 z)]; \xi_i(c; \alpha_0), \eta_i(\alpha_0), \text{ and } \Psi_i \text{ are as defined in (A.1), (A.2), and (A.9)} ; \text{ and } P_2 \text{ and } P_3 \text{ denote the limits of } P_{2n} \text{ and } P_{3n} \text{. It is apparent that it is not possible to evaluate the foregoing distribution analytically, because the limiting process of } \tilde{\mathcal{F}}_k(r) \text{ does not have an independent increments structure. For this, we propose using the following simulation approach, which was discussed by Cheng, Wei, and Ying (1997) and Lin et al. (2000).}

Let \( (G_1, \ldots, G_n) \) be independent standard normal variables independent of the data. Then it can be shown, following Cheng et al. (1997) and Lin et al. (2000), that the distribution of the process \( \tilde{\mathcal{F}}_k(r) \) can be approximated by that of the mean-0 Gaussian process
\[
\tilde{\mathcal{F}}_k(r) = n^{-1/2} \sum_{i=1}^{n} \int_{0}^{t} \left\{ I(Z_{ik} \leq r) - \frac{s_k(u, r)}{s^{(0)}(u)} \right\} d\hat{M}_i(u)G_i
\]
\[
+ n^{-1/2} \sum_{i=1}^{n} \int \sum_{i=1}^{m} \left\{ I(z_k \leq r) - \frac{s_k(t, r)}{s^{(0)}(t)} \right\} \times \tilde{v}(\hat{g}_i(c) + z' \eta_i(\hat{\alpha})) dP_{2n}(z, c, m, t_1, \ldots, t_m)G_i
\]
\[
+ n^{-1/2} \sum_{i=1}^{n} \int \int \left\{ I(z_k \leq r) - \frac{s_k(u, r)}{s^{(0)}(u)} \right\} \times \exp(g_0 z) dP_{3n}(z, c, m) d\hat{A}(u)G_i
\]
\[
b_1(\tau, r) D^{-1} n^{-1/2} \sum_{i=1}^{n} \hat{\psi}_i G_i
\]
\[
b_2(\tau, r) n^{-1/2} \sum_{i=1}^{n} \hat{\eta}_i(\hat{\alpha}) G_i.
\]

For the assessment of the identical link function of model (1), we can consider the process
\[
\mathcal{F}_k(r) = n^{-1/2} \sum_{i=1}^{n} I(\hat{\beta}'Z_i \leq r) \hat{M}_i(t),
\]
following Lin et al. (2000) and Pan and Lin (2005). As with \( \tilde{\mathcal{F}}_k(r) \), we can show that the null distribution of \( \mathcal{F}_k(r) \) can be approximated by that of \( \mathcal{F}_k(r) \), the process \( \tilde{\mathcal{F}}_k(r) \) given in (7) with \( I(Z_{ik} \leq r) \) replaced by \( I(\hat{\beta}'Z_i \leq r) \). Thus we can develop similar graphical and numerical procedures as for checking functional form of covariates, for testing the identical link function.
Finally, to construct an omnibus test for the overall fit of model (1), we define
\[ F_0(t, z) = n^{-1/2} \sum_{i=1}^{n} I(Z_i \leq z) \hat{M}_i(t), \]
where the event \( I(Z_i \leq z) \) means that each of the components of \( Z_i \) is no larger than the corresponding component of \( z \). As with \( F_k(r) \) and \( F_s(r) \), we can similarly show that the null distribution of \( F_0(t, z) \) can be approximated by that of the mean-0 Gaussian process \( \hat{F}_k(t, z) \) given by (7), with \( I(Z_{ik} \leq r) \) replaced by \( I(Z_i \leq z) \). An omnibus test statistic is then given by \( \sup_r |F_0(t, z)| \), from which a \( p \) value can be obtained as with \( \sup_r |F_k(r)| \) discussed earlier.

5. NUMERICAL RESULTS

We conducted simulation studies to assess the performance of the estimation procedure proposed in Section 3 with the focus on estimating \( \beta \). Wang et al. (2001) has investigated the performance of the estimators of \( \gamma \) and \( \lambda_0(t) \). In our studies, we considered two situations for the \( Z_i \)'s: (a) the two-sample problem with the \( Z_i \)'s generated from the Bernoulli distribution with success probability .5 and (b) generating the \( Z_i \)'s from the normal distribution N(0, .25). For given \( Z_i \) from the Bernoulli distribution, following Wang et al. (2001), we generated the latent variable \( V_i \) by letting \( V_i = \exp(-\ln(2.75)Z_i)V^*_i \), with \( V^*_i \) generated from the density function
\[ f(v^*|Z_i) = (1 - Z_i)I(.5 \leq v^* \leq 1.5) + \frac{Z_i}{2.5} I(1.5 \leq v^* \leq 4). \]
For given \( Z_i \) from N(0, .25), we used \( V_i = \exp(-\ln(2.75)I(Z_i \geq 0))V^*_i \), with \( V^*_i \) generated from the density function
\[ f(v^*|Z_i) = I(Z_i < 0)I(.5 \leq v^* \leq 1.5) + \frac{I(Z_i \geq 0)}{2.5} I(1.5 \leq v^* \leq 4). \]
It can be easily shown that \( E(V_i|Z_i) = 1 \).

For each given \( Z_i \) and \( V_i \), we generated the observation times \( T_{ij} \)'s from model (2) with \( \gamma = .5 \) and \( \lambda_0(t) = 1 \) (stationary Poisson process) or
\[ \lambda_0(t) = 1 + \frac{t - 2}{8} \]
(nonstationary Poisson process). For the censoring time \( C_i \), we investigated two situations: (a) independent censoring with \( C_i \) generated from the uniform distribution U(1, 6) and (b) dependent censoring with \( C_i \) assumed to follow the truncated exponential distribution over [1, 6] with density function
\[ f(c|Z_i, V_i) = \frac{V_i \exp(-V_i c/6)/6}{\exp(-V_i/6) - \exp(-V_i)} I(1 \leq c \leq 6). \]
For the response variable, we assumed that \( Y_i(t) \) was given by
\[ Y_i(t) = \mu_0(t) + \beta Z_i + V_i + \epsilon_i, \]
where the \( \epsilon_i \)'s are a random sample from the standard normal distribution. For \( \mu_0(t) \), we considered two functions, \( \mu(t) = 1 + t \) and \( \mu_0(t) = 1 + t^{1/2} \). The results, presented later, are based on 1,000 replications and sample size \( n = 200 \).

Table 1. Simulation results from stationary Poisson processes

<table>
<thead>
<tr>
<th>( C_i )</th>
<th>( \mu_0(t) )</th>
<th>BIAS</th>
<th>ESE</th>
<th>SSD</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z_i \sim \text{Bin}(5) )</td>
<td>( 1 + t )</td>
<td>.0270</td>
<td>.2101</td>
<td>.1951</td>
<td>.958</td>
</tr>
<tr>
<td></td>
<td>( 1 + t^{1/2} )</td>
<td>.0376</td>
<td>.1951</td>
<td>.1980</td>
<td>.947</td>
</tr>
<tr>
<td>Uniform</td>
<td>( 1 + t )</td>
<td>.0345</td>
<td>.2079</td>
<td>.2023</td>
<td>.955</td>
</tr>
<tr>
<td></td>
<td>( 1 + t^{1/2} )</td>
<td>.0379</td>
<td>.1920</td>
<td>.1838</td>
<td>.963</td>
</tr>
<tr>
<td>( Z_i \sim N(0, .25) )</td>
<td>( 1 + t )</td>
<td>.0020</td>
<td>.2199</td>
<td>.2003</td>
<td>.952</td>
</tr>
<tr>
<td></td>
<td>( 1 + t^{1/2} )</td>
<td>.0152</td>
<td>.2025</td>
<td>.1885</td>
<td>.959</td>
</tr>
</tbody>
</table>

Table 2. Simulation results from nonstationary Poisson processes

<table>
<thead>
<tr>
<th>( C_i )</th>
<th>( Q(t) )</th>
<th>BIAS</th>
<th>ESE</th>
<th>SSD</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z_i \sim \text{Bin}(5) )</td>
<td>( Q_1(t) )</td>
<td>.0066</td>
<td>.2132</td>
<td>.2013</td>
<td>.961</td>
</tr>
<tr>
<td></td>
<td>( Q_2(t) )</td>
<td>.0258</td>
<td>.1899</td>
<td>.1784</td>
<td>.964</td>
</tr>
<tr>
<td>Uniform</td>
<td>( Q_1(t) )</td>
<td>.0172</td>
<td>.2091</td>
<td>.2011</td>
<td>.952</td>
</tr>
<tr>
<td></td>
<td>( Q_2(t) )</td>
<td>.0348</td>
<td>.1869</td>
<td>.1813</td>
<td>.945</td>
</tr>
<tr>
<td>( Z_i \sim N(0, .25) )</td>
<td>( Q_1(t) )</td>
<td>.0086</td>
<td>.2223</td>
<td>.2145</td>
<td>.953</td>
</tr>
<tr>
<td></td>
<td>( Q_2(t) )</td>
<td>.0139</td>
<td>.1944</td>
<td>.1890</td>
<td>.944</td>
</tr>
<tr>
<td>Uniform</td>
<td>( Q_1(t) )</td>
<td>.0038</td>
<td>.2199</td>
<td>.2056</td>
<td>.961</td>
</tr>
<tr>
<td></td>
<td>( Q_2(t) )</td>
<td>.0112</td>
<td>.1911</td>
<td>.1851</td>
<td>.948</td>
</tr>
</tbody>
</table>

Note: BIAS represents the mean of estimates minus the true value, ESE represents the mean of the estimated standard errors, SSD represents the sample standard deviation of the estimates, and CP represents the empirical 95% coverage probability.
parameter estimator seems to converge slower for the weight function $Q_2(t)$ than for the weight function $Q_1(t)$. But the estimator associated with $Q_2(t)$ seems more efficient than that associated with $Q_1(t)$.

To assess the asymptotic normality of $\hat{\beta}$ given in Section 3, we studied the quantile plots of the standardized $\hat{\beta}$ against the standard normal variable. The plots, which are omitted here, indicate that the asymptotic normality is reasonable for the situations studied here. We also investigated other setups that included larger average number of observations per subject, larger sample sizes, and larger ranges for the censoring times $C_i$’s, and obtained similar results.

6. AN APPLICATION

To illustrate the proposed methodology, in this section we apply it to the longitudinal bladder cancer data discussed earlier and analyzed by Sun and Wei (2000) and Zhang (2002), among others. The data include 85 bladder cancer patients, 47 in the placebo group and 38 in the thiotepa treatment group. For each patient, the observed information includes the clinical visit or observation times (in months) and the number of bladder tumors that occurred between clinical visits. In addition, two baseline covariates were measured, the number of initial tumors that the patients had before entering the study and the size of the largest initial tumor. The longest observation time was 53 months. One of the study’s main objectives was to assess the effect of thiotepa treatment on the bladder tumor recurrence. For our analysis here, we focus on the effects of thiotepa treatment and the number of initial tumors on the recurrence rate of bladder tumor. Note that the size of the largest initial tumor had been shown to have no effect on the recurrence rate (Sun and Wei 2000; Zhang 2002).

For the analysis, we defined $Y_i(t)$ as the natural logarithm of the number of observed tumors at time $t$ plus 1 to avoid 0, $i = 1, \ldots, 85$. We defined the covariates as $Z_{i1} = 0$ if the patient was in the placebo group and 1 if the patient was in the thiotepa group and $Z_{i2}$ as the number of initial tumors, $i = 1, \ldots, 85$. First, we assumed that the data can be described by models (1) and (2) and considered estimation of regression parameters. The application of the proposed method in Section 3 with $W_i = Q(t) = 1$ yielded $\hat{\beta}_1 = -0.4656$ and $\hat{\beta}_2 = 0.0145$ with estimated standard errors of .1137 and .0331. By letting $Q(t) = Q_2(t)$ with $Q_2(t)$ defined as in the simulation study, we obtained $\hat{\beta}_1 = -0.5242$ and $\hat{\beta}_2 = 0.0093$ with estimated standard errors of .1134 and .0322. These results suggest that the thiotepa treatment had a significant effect in reducing the recurrence of bladder tumor, but the initial number of bladder tumors seems to have no significant effect in predicting the recurrence rate of the bladder tumor.

For direct comparison, we also analyzed the data by assuming that the response process and the observation process are independent given covariates, that is, $V_i = 1$ for all $i$. In this case, using the approach of Lin and Ying (2001), who studied models (1) and (2) without the latent variable $V$, we obtained $\hat{\beta}_1 = -1.9496$ and $\hat{\beta}_2 = 0.0492$ with estimated standard errors of .0456 and .0131. It is interesting to note that the estimated effect of the thiotepa treatment was a little larger than but similar to that obtained earlier, but the estimated effect of the initial number of tumors was quite different from that given earlier.

The aforementioned differences indicate that by ignoring the correlation between $Y(t)$ and $N(t)$, we can overestimate the effects of both the treatment and the covariate. This is particularly true for the covariate effect. One possible reason for this is that the observation process carries some information about the tumor recurrence rate, and using the independence assumption could falsely transform the correlation between $Y(t)$ and $N(t)$ to that between $Y(t)$ and the covariate even through a significant relationship between $Y(t)$ and the covariate may not exist.

In summary, the foregoing results suggest a correlation between the tumor occurrence process and the observation process. After adjusting for this relationship, thiotepa treatment still had a significant effect in reducing the recurrence of the bladder tumor, but the initial number of bladder tumors seemed to have no significant effect on overall tumor recurrence. Note that under a conditional model, Sun et al. (2005) concluded that the tumor occurrence rate at any time was related to the number of observations during the previous 6 months.

Now we consider the application of the model-checking procedures given in Section 4 to the data. In particular, we conducted three tests: the functional form for the number of initial tumors, the identical link function, and the overall fit. For the test of the functional form, we used the statistic $F_2(r)$ and obtained $\sup_{r} |\hat{F}_2(r)| = 1.1783$ with a $p$ value of .689 based on 1,000 realizations of the statistic $\sup_{r} |\hat{F}_2(r)|$. This suggests that the linear form for the number of initial tumors in model (1) seems reasonable. To give a graphical comparison, Figure 1 displays the observed process $F_2(r)$ along with 20 realizations of the process $\hat{F}_2(r)$ and reaches a similar conclusion.

To assess the identical link function used in model (1), we calculated the statistic $F_6(r)$ and obtained $\sup_{r} |\hat{F}_6(r)| = 1.9862$ and a $p$ value of .677 based on 1,000 realizations of the statistic $\sup_{r} |\hat{F}_6(r)|$. For the overall fit of model (1), the application of the procedure given in Section 4 gave $\sup_{t,z} |\hat{F}_0(t,z)| = 2.3686$ and a $p$ value close to 1 again using 1,000 realizations of the statistic $\sup_{t,z} |\hat{F}_0(t,z)|$. These results indicate that model (1) fits the observed data well.

7. CONCLUDING REMARKS AND DISCUSSION

We have presented a statistical approach for regression analysis of longitudinal data when the underlying longitudinal
process, the observation process, and the censoring process may be correlated. In the proposed methodology, marginal models are specified with the use of some latent variable for the connection of the three processes. For estimating regression parameters, we have proposed an estimating equation approach that yields consistent and asymptotically normal estimators. The approach can be considered generalizations of both the method given by Lin and Ying (2001) for longitudinal data with noninformative observation and censoring times and the approach of Huang and Wang (2004) for recurrent-event data in the presence of informative censoring. Lin et al. (2004) and Sun et al. (2005) considered the same problem studied here but assumed that the censoring process is independent of both the longitudinal process and the observation process of covariates. Another difference between the methods given here and those of Sun et al. (2005) is that here we adopted the marginal approach, whereas Sun et al. (2005) proposed a conditional approach.

As is well known, model checking is always an important issue in regression analysis, because most regression models have limitations. Huang and Wang (2004) provided some discussion about assessing model (2); in this article we confined the focus to the assessment of model (1) assuming that model (2) holds. In addition to the approach discussed here, another commonly used approach for model checking is to apply the cross-validation procedure to, for example, evaluate the predicted error of the model.

Other work in the literature similar to that presented here includes that of Sun and Wei (2000), which considered the similar inference problem but for interval-censored or panel count data. These authors also applied marginal models for the processes involved and assumed that the processes are independent given covariates. Another situation that is similar to, but not the same as that discussed here and has been discussed by many authors in the literature is the analysis of longitudinal data with informative dropout times as mentioned earlier. In this case, a common approach is to use latent variables to connect the correlated processes as in the method presented here.

As noted earlier, models (1) and (2) allow a positive correlation between a response process $Y(t)$ and the observation process $N(t)$. Although these models fit the example discussed in Section 6 well, as pointed out by a referee, the negative correlation may exist in some situations. For this case, instead of model (1), we could apply the model

$$E[Y(t)|Z,V] = \mu_0(t) + \beta' Z - V,$$

and the could easily develop inference procedures similar to those given in previous sections. A more general approach is to generalize model (1) to

$$E[Y(t)|Z,V] = \mu_0(t) + \beta' Z + \sigma V,$$

where $\sigma$ is an unknown parameter. Estimation of $\sigma$ does not seem to be straightforward in this model.

Another limitation of the approach given here, as in many methods proposed for longitudinal data with informative dropout times, is that the latent variable $V$ is time-independent. Also, we considered only time-independent covariates. In some situations, one may want to consider the model

$$E[Y(t)|Z(t),V(t)] = \mu_0(t) + \beta' Z(t) + V(t)$$

instead of model (1), where $Z(t)$ is a vector of covariates that may depend on time and $V(t)$ is a mean-0 stochastic process. As for future research, another problem that was not discussed here but may be of interest in some cases, is estimation of $\mu_0(t)$. One approach to this is to perform a two-step iterative procedure by estimating $\beta$ and $\mu_0(t)$ alternatively. However, the study of the asymptotic properties of the resulting estimators would not be easy, and in particular, the estimate of $\mu_0(t)$ may not be root-$n$ consistent. For the example discussed in Section 6, as an alternative and suggested by the associate editor, one could define $Y(t)$ as the cumulative counts of bladder tumors (Sun and Wei 2000; Zhang 2002) and obtain the standard weak convergence rate for the estimate of $\mu_0(t)$.

**APPENDIX A: ASYMPTOTIC NORMALITY OF $\hat{\beta}$**

Here we use the same notation defined earlier and take all limits at $n \to \infty$. Define

$$S^{(0)}(t; \gamma) = n^{-1} \sum_{i=1}^{n} \Delta_i(t) \hat{\gamma}_i \exp(\gamma' Z_i),$$

$$S^{(1)}(t; \gamma) = n^{-1} \sum_{i=1}^{n} \Delta_i(t) \hat{\gamma}_i \exp(\gamma' Z_i) Z_i,$$

and

$$S^{(2)}(t; \gamma) = n^{-1} \sum_{i=1}^{n} \Delta_i(t) \hat{\gamma}_i \exp(\gamma' Z_i) Z_i^2.$$

In addition, $S^{(0)}(t; \gamma_0)$, $S^{(1)}(t; \gamma_0)$, and $S^{(2)}(t; \gamma_0)$ denote the limits of $S^{(0)}(t; \gamma_0)$, $S^{(1)}(t; \gamma_0)$, and $S^{(2)}(t; \gamma_0)$.

To study the asymptotic distribution of the proposed estimate $\hat{\beta}$, we need the following regularity conditions:

(R1) $P(C \geq t, V > 0) > 0$ and $E(V^2) < \infty$.

(R2) $Z$ is bounded and $G(t) = E[V \exp(\gamma_0' Z) I(C \geq t)]$ is a continuous function for $t \in [0, \tau]$.

(R3) The weight function $Q(t)$ has bounded variation and converges to a deterministic function $q(t)$ in probability uniformly in $t \in [0, \tau]$.

(R4) $D = E[\int_0^t q(t)|Z_i - \tilde{z}(t)|^2 \Delta_i(t) dN_i(t)]$ is nonsingular, where $\tilde{z}(t) = s^{(1)}(t)/s^{(0)}(t)$.

Define $R(t) = G(t) \Lambda_0(t)$, $H(t) = \int_0^t G(u) d\Lambda_0(u)$,

$$b_i(t) = \sum_{j=1}^{m_i} \left\{ \int_0^t \frac{1}{R(u)} \left( H(u) - I(t < T_{ij} \leq \tau) \right) dH(u) \right\},$$

and

$$e_i(\alpha) = -\int w_i x_i b_i(c) dP_i(w, x, c, m) \frac{F(c)}{F(c_i)} + W_i X_i \left( \frac{m_i}{F(c_i)} - \exp(\alpha' X_i) \right),$$

where $P_i(w, x, c, m)$ is the joint probability measure of $(W_i, X_i, C_i, m_i)$. Let $\eta_i(\alpha)$ denote the vector function $(-\partial e_i(\alpha)/\partial \alpha)^{-1} e_i(\alpha)$ without the first entry, let $\hat{\phi}_i(\alpha)$ denote the first entry of $(-\partial e_i(\alpha)/\partial \alpha)^{-1} e_i(\alpha)$, and let $\hat{\xi}_i(t;\alpha) = b_i(t) + \hat{\phi}_i(\alpha)$. Under conditions (R1) and (R2), it follows from Wang et al. (2001) that

$$n^{1/2} (\hat{\Lambda}_0(t) - \Lambda_0(t)) = n^{-1/2} \Lambda_0(t) \sum_{i=1}^{n} \hat{\xi}_i(t;\alpha_0) + o_p(1) \quad (A.1)$$
and

\[ n^{1/2}(\hat{\gamma} - \gamma_0) = n^{-1/2} \sum_{i=1}^{n} \eta_i(\alpha_0) + o_P(1), \]  

(A.2)

where \( \alpha_0 \) is the true value of \( \alpha \). Using (A.1), (A.2), and the Taylor series expansion, we have

\[
\begin{align*}
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) Z_i \left( Y_i(t) - \frac{m_j}{A_0(C_i)} \exp(\gamma_0 Z_i) - \beta_0' Z_i \right) \\
&\times \Delta_i(t) dN_i(t) \\
&+ n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) Z_i \left( \frac{m_j}{A_0(C_i)} \exp(\gamma_0 Z_i) \right) \\
&\times \Delta_i(t) dN_i(t) \\
&+ n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) Z_i \left( \frac{m_j}{A_0(C_i)} \exp(\gamma_0 Z_i) \right) \\
&\times \Delta_i(t) dN_i(t) + o_p(1) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) Z_i \left( Y_i(t) - \frac{m_j}{A_0(C_i)} \exp(\gamma_0 Z_i) - \beta_0' Z_i \right) \\
&\times \Delta_i(t) dN_i(t) \\
&\times \left( \xi_0(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m) \\
&\times \left( \xi_i(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m, t_1, \ldots, t_m) + o_p(1), \tag{A.3}
\end{align*}
\]

where \( P_3(z, c, m) \) is the joint probability measure of \( (Z_i, C_i, m_i, T_i, \ldots, T_i) \).

Similarly,

\[
\begin{align*}
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) S(1)(t; \gamma_0) dA(t) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) S(1)(t; \gamma_0) dA(t) - n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) \\
&\times \left( \int \frac{I(t \geq t) m_j}{A_0(c)} \left( \xi_0(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m) \right) dA(t) \\
&+ o_p(1), \tag{A.4}
\end{align*}
\]

\[
\begin{align*}
&\int_0^\tau q(t) \bar{z}(t) S(0)(t; \gamma_0) dA(t) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) \bar{z}(t) \frac{A_0(C_i)}{m_j} dA(t) \\
&\times \left( \xi_i(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m) dA(t) + o_p(1), \tag{A.5}
\end{align*}
\]

and

\[
\begin{align*}
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) \bar{z}(t) \left( Y_i(t) - \hat{V}_i - \beta_0' Z_i \right) \Delta_i(t) dN_i(t) \\
&\times \left( \xi_i(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m) dA(t) + o_p(1,1). \tag{A.6}
\end{align*}
\]

where \( A(t) \) is as defined in (6) and \( P_3(z, c, m) \) is the joint probability measure of \( (Z_i, C_i, m_i) \).

Note that the estimating function \( \hat{U}(\beta; \gamma) \) can be rewritten as

\[
\hat{U}(\beta; \gamma) = n^{-1} \sum_{i=1}^{n} \int_0^\tau q(t) Z_i \left( Y_i(t) - \hat{V}_i - \beta_0' Z_i \right) \Delta_i(t) dN_i(t) \\
- n^{-1} \sum_{i=1}^{n} \int_0^\tau q(t) S(1)(t; \gamma_0) dA(t) \\
+ n^{-1} \sum_{i=1}^{n} \int_0^\tau q(t) S(0)(t; \gamma_0) dA(t) \\
+ o_p(1),
\]

and \( \hat{U}(\beta_0; \gamma_0) \to 0 \) in probability by the law of large numbers. Thus, using (A.3)–(A.6), lemma A.1 of Lin and Ying (2001), and the functional version of the Taylor expansion for the mapping, we get

\[
\begin{align*}
&n^{1/2} \hat{U}(\beta_0; \gamma_0) \\
&= n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) Z_i \left( Y_i(t) - \hat{V}_i - \beta_0' Z_i \right) \Delta_i(t) dN_i(t) \\
&- n^{-1} \int_0^\tau q(t) S(1)(t; \gamma_0) dA(t) \\
&+ n^{-1} \int_0^\tau q(t) S(0)(t; \gamma_0) dA(t) \\
&- n^{-1/2} \sum_{i=1}^{n} \int_0^\tau q(t) \bar{z}(t) \left( Y_i(t) - \hat{V}_i - \beta_0' Z_i \right) \Delta_i(t) dN_i(t) \\
&+ o_p(1,1), \tag{A.7}
\end{align*}
\]

which is a sum of \( n \) independent mean-0 random vectors plus an asymptotically negligible term. In (A.7),

\[
\begin{align*}
\Phi_i = \int_0^\tau q(t) Z_i \left( Y_i(t) - \frac{m_j}{A_0(C_i)} \exp(\gamma_0 Z_i) - \beta_0' Z_i \right) \\
&\times \Delta_i(t) dN_i(t) \\
&+ \int \sum_{j=1}^{m} q(j) [z - \bar{z}(j)] \frac{m_j}{A_0(c) \exp(\gamma_0 z)} \\
&\times \left( \xi_i(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m, t_1, \ldots, t_m) \\
&+ \int_0^\tau q(t) \left( \int I(c \geq t) m_j \frac{I(c \geq t) m_j}{A_0(c)} \right) \\
&\times \left( \xi_i(c; \alpha_0) + \gamma_0' \eta_i(\alpha_0) \right) dP_3(z, c, m) dA(t) \\
&- \int_0^\tau q(t) \left( \Delta_i(t) m_j \frac{A_0(C_i)}{m_j} \right) [Z_i - \bar{z}(t)] dA(t).
\end{align*}
\]
We can check that $-\partial \hat{U}(\beta_0; \gamma) / \partial \gamma|_{\gamma=\gamma_0}$ converges in probability to
\[
P = \mathbb{E} \left[ \int_0^1 q(t) (s(2)(t)/s(0)(t) - \bar{z}(t)^2) \times \{ V_i(t) - V_i - \beta_0 Z_i \} \Delta_i(t) dN_i(t) \right].
\]
The Taylor series expansion of $\hat{U}(\beta_0; \hat{\gamma})$ at $\hat{U}(\beta_0; \gamma_0)$, together with the consistency of $\hat{\gamma}$ and (A.2), yields
\[
n^{1/2} \hat{U}(\beta_0; \hat{\gamma}) = n^{1/2} \hat{U}(\beta_0; \gamma_0) - Pn^{-1/2} \sum_{i=1}^n \eta_i(a_0) + o_p(1).
\]
(A.8)

It then follows from the multivariate central limit theorem and (A.7) and (A.8) that $n^{1/2} \hat{U}(\beta_0; \hat{\gamma})$ converges in distribution to a mean-0 normal random vector with covariance matrix $\Sigma = E[\Psi^2]$, where
\[
\Psi = \Phi_i - P \eta_i(a_0).
\]
(A.9)

Note that $-\partial \hat{U}(\beta; \gamma)/\partial \beta$ converges in probability to $D_\beta$, which is defined in condition (R4). The Taylor expansion of $\hat{U}(\beta; \gamma)$ at $\hat{U}(\beta_0; \gamma_0)$ gives
\[
n^{1/2}(\hat{\beta} - \beta_0) = D^{-1} n^{1/2} \hat{U}(\beta_0; \gamma_0) + o_p(1).
\]
(A.10)

Thus it follows from (A.8) that $n^{1/2}(\hat{\beta} - \beta_0)$ is asymptotically mean-0 normal with covariance matrix $D^{-1} \Sigma D^{-1}$. Using the method in appendix A.3 of Lin et al. (2000), we can show that $D^{-1} \Sigma D^{-1}$ can be consistently estimated by $\hat{D}^{-1} \hat{\Sigma} \hat{D}^{-1}$.

APPENDIX B: WEAK CONVERGENCE OF $F_k(r)$, $F_g(t)$, AND $F_0(t, z)$

Here we sketch the proof only for the weak convergence of $F_k(r)$ under models (1) and (2). The weak convergence of $F_g(t)$ and $F_0(t, x)$ can be derived similarly. To prove the weak convergence of $F_k(r)$, first, using lemma A.1 of Lin and Ying (2001) and the functional version of the Taylor expansion, we have
\[
F_k(r) = n^{-1/2} \sum_{i=1}^n \int_0^1 \left\{ I(Z_{ik} \leq r) - \frac{s(q_i, u)}{s(0)(u)} \right\} dM_i(u)
\]
\[
+ n^{-1/2} \sum_{i=1}^n \int_0^r \sum_{l=1}^m \left\{ I(z_{ik} \leq r) - \frac{s(q_i, r_i)}{s(0)(r_i)} \right\}
\times v^*(\xi_l(c; a_0) + z^* \eta_l(a_0)) \exp(y(0)z) dP_3(z, c, m, t_1, \ldots, t_m)
\]
\[
+ n^{-1/2} \sum_{i=1}^n \int_0^r \left\{ I(z_{ik} \leq r) - \frac{s(q_i, r_i)}{s(0)(r_i)} \right\} I(c \geq u)
\times v^*(\xi_l(c; a_0) + z^* \eta_l(a_0)) \exp(y(0)z) dP_3(z, c, m)
\]
\[
+ b_k(r, r)r^{1/2}(\hat{\beta} - \beta_0) - b_k(r, r)r^{1/2}(\hat{\gamma} - \gamma_0) + o_p(1).
\]

In this formula, the tightness of the first three terms directly follows the arguments given in appendix A.5 of Lin et al. (2000). The last two terms are also tight because $n^{1/2}(\hat{\beta} - \beta_0)$ and $n^{1/2}(\hat{\gamma} - \gamma_0)$ converge in distribution and $b_k(r, r)$ and $b_k(r, r)$ are some deterministic functions. Therefore, it follows that $F_k(r)$ is tight.

Based on (A.2), (A.8), (A.9), and (A.10), we can further write $F_k(r)$ as
\[
F_k(r) = n^{-1/2} \int_0^r \left\{ I(Z_{ik} \leq r) - \frac{s(q_i, r_i)}{s(0)(r_i)} \right\} dM_i(u)
\]
\[
+ n^{-1/2} \sum_{i=1}^n \int_0^r \sum_{l=1}^m \left\{ I(z_{ik} \leq r) - \frac{s(q_i, r_i)}{s(0)(r_i)} \right\}
\times v^*(\xi_l(c; a_0) + z^* \eta_l(a_0)) \exp(y(0)z) dP_3(z, c, m, t_1, \ldots, t_m)
\]
\[
+ n^{-1/2} \sum_{i=1}^n \int_0^r \left\{ I(z_{ik} \leq r) - \frac{s(q_i, r_i)}{s(0)(r_i)} \right\} I(c \geq u)
\times v^*(\xi_l(c; a_0) + z^* \eta_l(a_0)) \exp(y(0)z) dP_3(z, c, m)
\]
\[
+ b_k(r, r)r^{1/2}(\hat{\beta} - \beta_0) - b_k(r, r)r^{1/2}(\hat{\gamma} - \gamma_0) + o_p(1).
\]

It follows from the multivariate central limit theorem and the tightness of $F_k(r)$ that $F_k(r)$ converges weakly to the mean-0 Gaussian process given in (6).

[Received February 2002. Revised June 2007.]

REFERENCES


