数理解析研究所講究録 1273 巻 2002 年 100-114

Model checking for semiparemetric linear transformation model

北里大学大学院臨床統計専攻　服部　聡 (Satoshi Hattori)
Kitasato University Graduate School, Div. of Biostatistics

1 Introduction

In clinical trials, regression models are frequently used and play important roles. For example, in cancer clinical trials which are desined to evaluate the survival benefit of anti-cancer drugs, regression models for time-to-event data are used to evaluate the treatment effect adjusting the effect of covariates and to estimate(predict) survival functions of various kind of patients. The Cox proportional hazard model is frequenty and routinely used in the recent clinical trials with time-to-event data as endpoints. Of course the Cox proportional hazard model is quite useful and important because the Cox proportional hazard model is a semiparametric model, which seems to be more robust than the parametric models, and can be fitted easily by many commercial software packages. On the other hand, the Cox proportional hazard model may not fit well since the proportional hazard assumption, the Cox proportional hazard model requires, isn't neccesarily a weak one. Recently the practical inference procedure of some kind of semiparametric models have been established with the theoretical justifications. For example, Wei, Ying and Lin(1990), Tsiatis(1990), Ying(1993) and other authers have developed the inference procedure for the semiparametric accelerated failure time model and Lin and Ying(1994) has established the simple inference procedure for the semiparametric additive hazard model. The linear transformation model is another attractive alternative semiparametric model. The linear transformation model is defined as

$$S_Z(t) = g\{h(t) + Z^T\beta\}, \tag{1}$$

where $S_Z(t)$ is the survival function of the patient with the covariate Z, $h(t)$ is the unknown non-decreasing function and g is the known continuous and strictly increasing function. Here we assume that Z is bounded and, without loss of generality, $\mid Z \mid \leq 1$ is assumed. When $g^{-1}(t) = log(-log(t))$, (1) reduces to

$$log(-log(S_Z(t))) = h(t) + Z^T\beta.$$

This is the Cox proportional hazard model. And when $g^{-1}(t) = -logit(t)$, (1) reduces to

$$-logit(S_Z(t)) = h(t) + Z^T\beta.$$

This is the proportional odds model which is the important alternative to the Cox proportinal hazard model (Bennett (1982)). So changing the link function, g, the linear transformation model provides a large class of semiparametric models containing the Cox proportional hazard model and the proportional odds regression model as special cases. Recently Cheng, Wei and Ying(1995), Cheng, Wei, Ying(1997) and Fine, Wei and Ying(1998) proposed the inference procedure for the linear transformation model with the univariate, possibly right-censored data. And the extensions to more complicated data were done (Cai, Wei and Wilcox(2000), Cheng and Wang(2001), Lin, Wei and Ying(2001)). For the univariate data, Cheng, Wei and Ying(1997) proposed an graphical model checking procedure based on p-p plot. Though their model checking procedure is very useful to check and select the model, it is desirable to establish the formal model checking procedure since the graphical procedure may be subjective. For the Cox proportional hazard model, Lin, Wei and Ying(1993) proposed an quite useful model checking procedure based on the cumulative martingale-based residuals. Their method provides the formal omnibus test and , in addition, the graphical model checking technique which is very useful to investigate what kind of misspecification occure. In this article, we develop the model checking technique for the linear transformation model with the univariate, possible right-censored data based on the cumlative sum of martingale-type residuals.

2 The inference procedure for the linear transformation model

In this section, we summarize the inference procedure for the linear transformation model proposed by Cheng, Wei and Ying(1995, 1997). The linear transformation model is defined equivalently to (1) as

$$h(T) = -Z^T\beta + \varepsilon,$$

where T is the failure time, ε is the random variable whose distribution function $F = 1 - g$ is completely known. The specification of the distribution of ε corresponds to that of the link function g. The distribution of ε is the standard extreme value distribution and the standart logistic distribution for the Cox proportional hazard model and the proportional odds model respectively. To estimate the unknown parameter β and h, the survival function G of the censoring time C is assumed not to depend on the covariate Z. To estimate the regression coefficient β, Cheng, Wei and Ying(1995) proposed the unbiased estimating equation,

$$U(\beta) = \sum_{i=1}^{n}\sum_{j=1}^{n} w(Z_{ij}^T\beta)Z_{ij}\Big\{\frac{\Delta_j I(X_i \geq X_j)}{\hat{G}^2(t)} - \xi(Z_{ij}^T\beta)\Big\}, \tag{2}$$

where $\Delta_i = I(T_j \geq C_i)$, $\xi(s) = \int_{-\infty}^{\infty}\{1-F(t+s)\}dF(t)$, $w(.)$ is a weight function, $Z_{ij} = Z_i - Z_j$ and $\hat{G}$ is the Kaplan-Meier estimator for G. Cheng, Wei and Ying(1995) showed that $\hat{\beta}$, which is the solution of the equation $U(\beta) = 0$, converges to the true value β_0 almost surely and the distribution of $n^{\frac{1}{2}}(\hat{\beta} - \beta_0)$ converges asymptotically to normal distribution with mean zero whose asymptotic variance-covariance matrix can be consistently estimated by the sandwich-type variance estimator. In addition, Cheng, Wei and Ying(1997) proposed the inference procedure for the survival function

$S_z(t)$. To this end, they proposed the unbiased estimating equation for $h(t)$,

$$V(h(t)) = \sum_{i=1}^{n} \Big\{ \frac{I(X_i \geq t)}{\hat{G}(t)} - g(Z_i^T \hat{\beta}) \Big\}, \quad t \in [0, \tau] \tag{3}$$

where τ is a constant satisfying $P(X > \tau) > 0$ and $\hat{\beta}$ is the estimator derived by (2), and proved that $\hat{h}(t)$, the solution of $V(h(t)) = 0$, converges to $h_0(t)$ uniformly in $[0, \tau]$ almost surely.

2.1 Model cheching method based on the cumulative sum of martingale residuals

At first, we define the martingale-type residual for the linear transformation model in the similar matter for the Cox proportional hazard model. For the Cox proportional hazard model, the martingale residual is defined as

$$\tilde{M}_i(t) = N_i(t) - \int_0^t Y_i(u) e^{Z_i^T \tilde{\beta}} d\tilde{\Lambda}(u)$$

where $N_i(t) = I(X_i \leq t, \Delta_i = 1)$, $Y_i(t) = I(X_i \geq t)$, $\tilde{\beta}$ is the maximum partial likelihood estimator and $\tilde{\Lambda}(t)$ is the Breslow estimator of the baseline hazard function (Barlow and Prentice 1988, Thernea, Grambsch and Fleming 1990). Similarly we define the martingale residual for the linear transformation model, using the relation $\Lambda_{Z_i}(t) = -log(S_{Z_i}(t))$ and (1) as

$$\hat{M}_i(t) = N_i(t) + \int_0^t Y_i(u) dlog(g\{\hat{h}(t+) + Z_i^T \hat{\beta}\}).$$

Using this martingale residual, we define the goodness-of-fit statistics similar to that proposed by Lin, Wei and Ying(1993) for the Cox proportional hazard model. Define the multi-parameter stochastic process,

$$H(t, z) = \sum_{i=1}^{n} I(Z_i \leq z) \hat{M}_i(t).$$

If the model is correctly specified, by the Taylor series expansion, $n^{-\frac{1}{2}} H(t, z)$ is asymptotically equivalent to

$$\begin{aligned} n^{-\frac{1}{2}} \tilde{H}(t, z) &= n^{-\frac{1}{2}} \sum_{l=1}^{n} \int_0^t I(Z_l \leq z) dM_l(u) \\ &+ n^{-1} \sum_{l=1}^{n} \int_0^t I(Z_l \leq z) Y_l(u) \\ &\times d\Big[\tilde{g}\{h_0(u) + \beta_0^T Z_l\} \\ &\quad n^{\frac{1}{2}}\{\hat{h}(u) - h_0(u) + (\hat{\beta} - \beta_0)^T Z_l\Big] \end{aligned}$$

where $\tilde{g}(x) = \frac{dg(x)}{dx} / g(x)$.
From Cheng, Wei and Ying(1997), $n^{\frac{1}{2}}\{\hat{h}(t) - h_0(t) + (\hat{\beta} - \beta_0)^T Z_l]$ is asymptotically equivalent to

$$W_{Z_l}(t) = \frac{1}{a(t)} \Big[\{b(t) + a(t) Z_l\}^T D$$

$$\begin{aligned}
& \times \quad \Big\{n^{-\frac{3}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n} w(Z_{ij}^T\beta_0)Z_{ij}e_{ij}(\beta_0) \\
& + \quad 2n^{\frac{-1}{2}}\sum_{i=1}^{n}\int_0^\infty \frac{q(u)}{\pi(u)}dM_i^c(u)\Big\} \\
& + \quad n^{\frac{-1}{2}}\sum_{i=1}^{n} r_i(t) \\
& + \quad n^{\frac{-1}{2}}\sum_{i=1}^{n}\frac{\pi(t)}{G(t)}\int_0^\infty \frac{1}{\pi(u)}dM_i^c(u)\Big], \qquad (4)
\end{aligned}$$

where varioius kind of quantities in (4) are as follows,

$$\begin{aligned}
a(t) &= \lim_{n\to\infty}\hat{a}(t) \\
&= \lim_{n\to\infty}\frac{-1}{n}\sum_{i=1}^{n} f(\hat{h_0}(t)+Z_i^T\hat{\beta}), \\
b(t) &= \lim_{n\to\infty}\hat{b}(t) \\
&= \lim_{n\to\infty}\frac{-1}{n}\sum_{i=1}^{n} f(\hat{h_0}(t)+Z_i^T\hat{\beta})Z_i, \\
D^{-1} &= \lim_{n\to\infty}\hat{D}^{-1} \\
&= \lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} w(Z_{ij}^T\beta_0)\xi'(Z_{ij}^T\beta_0)Z_{ij}Z_{ij}^T, \\
\pi(t) &= \lim_{n\to\infty}\hat{\pi}(t) \\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} I(X_i \geq t), \\
q(t) &= \lim_{n\to\infty}\hat{q}(t) \\
&= \lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} w(Z_{ij}^T\beta_0)Z_{ij}\frac{\delta_j I(X_i\geq X_j)}{\hat{G}^2(t)}I(X_j\geq t), \\
e_{ij}(\beta_0) &= \frac{\delta_j I(X_i\geq X_j)}{\hat{G}^2(t)} - \xi(Z_{ij}^T\beta_0), \\
r_i(t) &= \frac{I(X_i\geq t)}{G(t)} - S_{Z_i}(t), \\
M_i^c(t) &= I(X_i\leq t, \Delta_i = 0) - \int_0^\infty I(X_i\geq u)d\Lambda_G(u),
\end{aligned}$$

f is the density function of F, the distribution function of ε and Λ_G is the cumulative hazard function of the common censoring time. Note that $M_i^c(t)$ is the martingale corresponding to the counting process for the censoring time C. Then

$$\begin{aligned}
n^{-\frac{1}{2}}H(t,z) \quad &\simeq \quad n^{-\frac{1}{2}}\sum_{l=1}^{n}\int_0^t I(Z_l\leq z)dM_l(u) \\
&+ \quad n^{-1}\sum_{l=1}^{n}\int_0^t I(Z_l\leq z)Y_l(u)
\end{aligned}$$

$$\times \quad d\Big[\tilde{g}\{h_0(u) + Z_l^T\beta_0\}W_{Z_l}(t)\Big]. \tag{5}$$

In appendix, the first and second term of the r.h.s of (5) is proved to be asmptotically equivalent to the zero-mean Gaussian process. To evaluate the goodness-of-fit objectively, the null distribution of the goodness-of-fit multi-parameter stochastic process is needed. Although, under the correct model specification, the goodness-of-fit stochastic process is asymptotically equivalent to the zero-mean Gaussian process, it is difficult to know the covariance structure of the Gaussian process analytically. So we use the simulation techniques to approximate the null Gaussian process following to the idea originally proposed by Lin, Wei and Ying(1993).
To approximate the null process, we define another stochastic process,

$$\begin{aligned}
n^{-\frac{1}{2}}\hat{H}(t,z) &= n^{-\frac{1}{2}}\sum_{i=1}^{n} I(Z_i \le z)\hat{M}_i(t)\mathcal{L}_i \\
&+ n^{-\frac{5}{2}}\sum_{l=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\int_0^t I(Z_l \le z)Y_l(u)Z_{ij}\hat{e_{ij}}(\hat{\beta}) \\
&\quad \times d\Big[\tilde{g}\{\hat{h}(u+) + Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\{\hat{b}(u+) + \hat{a}(u+)Z_l\}^T\hat{D}\Big] \\
&\quad \times \mathcal{L}_l \\
&+ 2n^{-\frac{3}{2}}\sum_{l=1}^{n}\sum_{i=1}^{n}\int_0^t I(Z_l \le z)Y_l(u)\int_0^\infty \frac{\hat{q}(u)}{\hat{\pi}(u)}d\hat{M}_i^c(u) \\
&\quad \times d\Big[\tilde{g}\{\hat{h}(u+) + Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\{\hat{b}(u+) + \hat{a}(u+)Z_l\}^T\hat{D}\Big] \\
&\quad \times \mathcal{L}_l \\
&+ n^{-\frac{3}{2}}\sum_{l=1}^{n}\sum_{i=1}^{n}\int_0^t I(Z_l \le z)Y_l(u) \\
&\quad \times d\Big[\tilde{g}\{\hat{h}(u+) + Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\hat{r}_i(u+)\Big] \\
&\quad \times \mathcal{L}_l \\
&+ n^{-\frac{3}{2}}\sum_{l=1}^{n}\sum_{i=1}^{n}\int_0^t I(Z_l \le z)Y_l(u) \\
&\quad \times d\Big[\tilde{g}\{\hat{h}(u) + Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\frac{\hat{\pi}(u+)}{\hat{G}(u+)}\int_0^u \frac{1}{\hat{\pi}(x)}d\hat{M}_i^c(x)\Big] \\
&\quad \times \mathcal{L}_l
\end{aligned} \tag{6}$$

where $\mathcal{L}_i, i = 1, 2, ..n$ are the sequence of the standard normal random variable independent of the data, $\{(X_i, \Delta_i, Z_i)\}$. In appendix, it is shown that conditional on the data, $n^{-\frac{1}{2}}\hat{H}(z,t)$ is the zero-mean Gaussian process and converges to the limiting process of $n^{-\frac{1}{2}}H(z,t)$ asymptotically. Note that conditional on the data, $\mathcal{L}_i, i = 1, 2, ..n$ are the only random components in $\hat{H}(z,t)$. In practice, arbitary numbers of realizations of the null process $\hat{H}(z,t)$ can be easily simulated by generating $\{\mathcal{L}_i, i = 1, 2, ..n\}$s in computer. Comparing the observed cumlative martingale residuals to the simulated null process, we can evaluate the goodness-of-fit.

2.2 An omnibus test

Similar to Lin, Wei and Ying(1993), an omnibus goodness-of-fit test is defined as

$$\begin{aligned} H_{omn} &= \sup_{z\in[-1,1]^p, 0\le t\le\tau} |H(t,z)| \\ &= \sup_{z\in[-1,1]^p, 0\le t\le\tau} |\sum_{i=1}^{n} I(Z_i \le z)\hat{M}_i(t)| \end{aligned}$$

Since the null distribution of $H(t,z)$ is zero-mean Gaussian process, an remarkably large value of H_{omn} suggests model misspecification. As the distribution of $n^{-\frac{1}{2}}H(t,z)$ can be approximated by $n^{-\frac{1}{2}}\hat{H}(t,z)$, H_{omn} can be approximated by $\hat{H}_{omn} = \sup_{z,t}|\hat{H}(t,z)|$ where the supremum is taken over $z \in [-1,1]^p, 0 \le t \le \tau$. So the p-value, $P\{H_{omn} \ge h_{omn}\}$ where h_{omn} is a realization of H_{omn}, can be approximated by $P\{\hat{H}_{omn} \ge h_{omn}\}$. The realizations of $\hat{H}_{omn}$ can be easily generated by simulation.

2.3 Checking the misspecification of functional form of covariates

An omnibus goodness-of-fit test may be a powerful guide to judge whether the fitted model is appropriate or not. When the fitted model, however, seem to be not appropreate, it is desirable to know what kind of model misspecification was made. Similar to Lin, Wei and Ying(1993), we define one-parameter stochastic process as

$$H^{(k)}(\tau,z) = \sum_{i=1}^{n} I(Z_i^{(k)} \le z^{(k)})\hat{M}_i(\tau)$$

where $z \in [-1,1]$ and $Z_i^{(k)}$ is the k-th element of Z_i. This stochastic process is a special case of $H(t,z)$. So the sample path of $H^{(k)}(\tau,z)$ under null hypothesis(i.e. the fitted model is correct) can be easily obtained though simulation and be displayed graphically because the sample path is one-dimensional function. Plotting the realization of $H^{(k)}(\tau,z)$ with some simulated realizations(say 20 realizations) of null distribution, it may be evaluated graphicall how strange the obtained realization is. Furthermore p-value based on $sup|H^{(k)}(\tau,z)|$ can be evaluated in the same way as the omnibus test.

2.4 Checking the misspecification of link function

To test the link misspecification, we can use the special case of $H(t,z)$ setting z=1.

$$H(t,1) = \sum_{i=1}^{n} \hat{M}_i(t).$$

Similar to $H_f^{(k)}(z)$,we can evaluate the goodness-of-fit of the fitted model graphically by ploting the obtained realization and the some(say 20) simulated realizations of the null process and subjectively by evaluating the p-value using the simulation technique.

3 Appendix

3.1 Weak convergence of $n^{-\frac{1}{2}}H(t,z)$

To show the weak convergence of $n^{-\frac{1}{2}}H(t,z)$, it is sufficient to prove the weak convergence of the 1st and 2nd terms of $n^{-\frac{1}{2}}\tilde{H}(t,z)$ converge to the zero-mean Gaussian process, where

$$\begin{aligned} n^{-\frac{1}{2}}\tilde{H}(t,z) &\simeq n^{-\frac{1}{2}}\sum_{l=1}^{n}\int_0^t I(Z_l \le z)dM_l(u) \\ &+ n^{-1}\sum_{l=1}^{n}\int_0^t I(Z_l \le z)Y_l(u) \\ &\times d\Big[\tilde{g}\{h_0(u)+Z_l^T\beta_0\}W_{Z_l}(t)\Big], \end{aligned} \quad (7)$$

since $n^{-\frac{1}{2}}H(t,z)$ and $n^{-\frac{1}{2}}\tilde{H}(t,z)$ are asymptotically equivalent.

Weak convergence of the 1st term of (7)

From the multivariate central limit theorem, the arbitary finite dimensional projection converge to the Gaussian distribution. So to show the weak convergence to the zero-mean Gaussian process of $n^{-\frac{1}{2}}\sum_{i=1}^{n}\int_0^t I(Z_i \le z)dM_i(u)$, it is sufficient to show its tightness in $D([-1,1]^p \times [0,\tau])$. Using the same argument as appendix 1 of Spiekerman and Lin(1996), the covariates can be restricted to the continuous ones without loss of generality. And for simplicity, the dimension of the covariates is assumed unity. To show the tightness, it is sufficient to show the moment inequalities,

$$E[\psi^2(z_1,z_2;t_1,t)\psi^2(z_1,z_2;t,t_2)] \le K(t-t_1)(t_2-t)Pr^2\{Z\in[z_1,z_2)\} \quad (8)$$

$$\begin{aligned} E[\psi^2(z_1,z;t_1,t_2)\psi^2(z,z_2;t_1,t_2)] &\le K(t_2-t_1)^2 \\ &\times Pr\{Z\in[z_1,z)\}Pr\{Z\in[z,z_2)\} \end{aligned} \quad (9)$$

for $\forall\ z_1<z<z_2, t_1<t<t_2$, where K is some constant and

$$\psi(z_1,z_2;t_1,t_2) = n^{-\frac{1}{2}}\sum_{i=1}^{n}\int_{t_1}^{t_2} dM_i(u)I(Z_i\in[z_1,z_2)).$$

It is sufficient to the moment inequarities (8) and (9) under the assumption that $Pr\{Z\in[z_1,z)\}\ge\frac{1}{n}$ and $Pr\{Z\in[z,z_2)\}\ge\frac{1}{n}$ (Bickel and Wichura(1971), Lin, Wei and Ying(1993)).

$$\begin{aligned} &l.s.h\ \ of\ \ (8) \\ &= E\Big[\psi^2(z_1,z_2;t_1,t)E\Big[\psi^2(z_1,z_2;t,t_2)|Z_i,i=1,2...n,\psi^2(z_1,z_2;t_1,t)\Big]\Big] \\ &= E\Big[\psi^2(z_1,z_2;t_1,t)\times \\ &\quad \frac{1}{n}\sum_{i=1}^{n}E\Big[\int_t^{t_2} Y_i(u)dlog(g\{h_0(u)+Z_i^T\beta_0\}I(Z_i\in[z_1,z))\Big|Z_i,\psi^2(z_1,z_2;t_1,t)\Big]\Big] \\ &= E[\psi^2(z_1,z_2;t_1,t)\frac{1}{n}\sum_{i=1}^{n}\int_t^{t_2} Y_i(u)dlog(g\{h_0(u)+Z_i^T\beta_0\}I(Z_i\in[z_1,z))] \end{aligned}$$

The 2nd equality is obtained by the independent increament property of the martingale and the standard moment caluculation of the counting process martingale. Furthermore

$$
\begin{aligned}
&\leq KE[\psi^2(z_1,z_2;t_1,t)\frac{1}{n}\sum_{i=1}^{n} I(Z_i \in [z_1,z))(t_2-t)] \\
&= KE[\frac{1}{n}\sum_{i=1}^{n} I(Z_i \in [z_1,z))\frac{1}{n}\sum_{j=1}^{n}\int_{t_1}^{t} Y_i(u)dlog(g\{h_0(u)+Z_j^T\beta_0\}I(Z_j \in [z_1,z)](t_2-t) \\
&\leq K(t_2-t)(t-t_1)\Big\{\frac{n(n-1)}{n^2}Pr^2(Z \in [z_1,z)) + \frac{1}{n}Pr^2(Z \in [z_1,z))\Big\} \\
&\leq K(t_2-t)(t-t_1)\Big\{\frac{n(n-1)}{n^2}Pr^2(Z \in [z_1,z)) + Pr^2(Z \in [z_1,z))\Big\} \\
&\leq 2K(t_2-t)(t-t_1)Pr^2(Z \in [z_1,z))
\end{aligned}
$$

So the moment inequility (8) is obtained. Next we show the 2nd moment inequility (9).

$$
\begin{aligned}
l.s.h \;\; of \;\; (9)& \\
&= E\Big[E[\psi^2(z_1,z;t_1,t_2)\psi^2(z,z_1;t_1,t_2)|Z_i, i=1,2,...n]\Big] \\
&= \frac{1}{n^2}\sum_{i\neq j} E\Big[\{\int_{t_1}^{t_2} dM_i(u)\}^2\{\int_{t_1}^{t_2} dM_j(u)\}^2 I(Z_i \in [z_1,z), Z_j \in [z,z_2))\Big] \\
&\leq K\frac{1}{n^2}n(n-1)(t_2-t_1)^2 Pr\{Z_i \in [z_1,z), Z_j \in [z,z_2)\} \\
&\leq K(t_2-t_1)^2 Pr\{Z_i \in [z_1,z)\}Pr\{Z_j \in [z,z_2)\}
\end{aligned}
$$

So the moment inequility (9) is also obtained.

Weak convergence of the 2nd term of (7)

Using the integration by part,

$$
\begin{aligned}
the\ 2nd\ term \;\; of \;\; (7)& \\
&= \frac{1}{n}\sum_{i=1}^{n} I(Z_i \leq z)\int_0^t Y_i(u)\tilde{g}\{h_0(u)+Z_i^T\beta\}dW_{Z_i}(u) \qquad (10) \\
&+ \frac{1}{n}\sum_{i=1}^{n} I(Z_i \leq z)\int_0^t Y_i(u)W_{Z_i}(u)d\tilde{g}\{h_0(u)+Z_i^T\beta\} \qquad (11)
\end{aligned}
$$

Here $W_{Z_i}(t)$ is re-expressed as,

$$
W_{Z_i}(t) = \frac{1}{a(t)}\Big[(b(t)+a(t)Z_i)^T D(W_1+W_2) + W_3(t) + W_4(t)\Big]
$$

where

$$
\begin{aligned}
W_1 &= n^{-\frac{3}{2}}\sum_{i=1}^{n}\sum_{j=1}^{n} w(Z_{ij}^T\beta_0)Z_{ij}e_{ij}(\beta_0) \\
W_2 &= 2n^{-\frac{1}{2}}\sum_{i=1}^{n}\int_0^{\infty}\frac{q(u)}{\pi(u)}dM_i^c(u)
\end{aligned}
$$

$$
\begin{aligned}
W_3(t) &= n^{-\frac{1}{2}}\sum_{i=1}^{n} r_i(t) \\
W_4(t) &= n^{-\frac{1}{2}}\sum_{i=1}^{n} \frac{\pi(t)}{G(t)}\int_0^t \frac{1}{\pi(u)} dM_i^c(u)
\end{aligned}
$$

Cheng, Wei and Ying(1997) proved that W_1, W_2, $W_3(t)$ and $W_4(t)$ are asymptotically equivalent to the zero-mean Gaussian process(appendix B of Cheng, Wei and Ying(1997)).
To show the weak convergence of (10) to the zero-mean Gaussian process,we prepare a lemma.

Lemma 3.1
Let $\{f_n(t)\}$ be a sequences of bounded variation functions on $[0,\tau]$ and $\{g_n(z,t)\}$ be a sequence of bounded variateion with respect to t in $[0,\tau]$ for each z in [-1,1] such that

1. $\sup_{0\le t\le\tau} |f_n(t) - f_\infty(t)| \to 0$, *where $f_\infty(t)$ is continuous on $[0,\tau]$,*
2. $\sup_{0\le t\le\tau,-1\le z\le 1} |g_n(z,t) - g_\infty(z,t)| \to 0$ *,where $g_\infty(z,t)$ is bounded on $[-1,1]\times[0,\tau]$ with bounded variation with respect to t for each z.*

Then

$$
\sup_{0\le t\le\tau,-1\le z\le 1} \left| \int_0^t f_n(u) dg_n(z,u) - \int_0^t f_\infty(u) dg_\infty(z,u) \right| \to 0. \tag{12}
$$

Proof
(12) is a simple extension of lemma A.3 of Bilias, Gu and Ying (1997). (Q.E.D)

Here we show the weak convergence of (10). Using the integration by part,

$$
\begin{aligned}
(10) &= \frac{1}{n}\sum_{i=1}^{n} I(Z_i\le z)\int_0^t Y_i(u)\tilde{g}\{h_0(u)+Z_i^T\beta\}Z_i^T D(W_1+W_2) d\frac{a(u)}{a(u)} \\
&+ \frac{1}{n}\sum_{i=1}^{n} I(Z_i\le z)\int_0^t Y_i(u)\tilde{g}\{h_0(u)+Z_i^T\beta\} \\
&\times d\Big[\frac{b^T(u)}{a(u)}D(W_1+W_2)+\frac{1}{a(u)}(W_3(u)+W_4(u))\Big] \\
&= \frac{1}{n}\sum_{i=1}^{n} I(Z_i\le z)\int_0^t Y_i(u)\tilde{g}\{h_0(u)+Z_i^T\beta\}Z_i^T D(W_1+W_2) d\frac{a(u)}{a(u)} \\
&+ \int_0^t d\Big[\frac{1}{n}\sum_{i=1}^{n} I(Z_i\le z)Y_i(u)\tilde{g}\{h_0(u)+Z_i^T\beta\}\frac{b^T(u)}{a(u)}D(W_1+W_2)+\frac{1}{a(u)}(W_3(u)+W_4(u))\Big] \\
&- \int_0^t \frac{b^T(u)}{a(u)}D(W_1+W_2)+\frac{1}{a(u)}(W_3(u)+W_4(u)) \\
&\times d\Big[\frac{1}{n}\sum_{i=1}^{n} I(Z_i\le z)Y_i(u)\tilde{g}\{h_0(u)+Z_i^T\beta\}\Big]
\end{aligned} \tag{13}
$$

Note that the 1st term of (13) is zero. Let $\mathcal{W}_1$, $\mathcal{W}_2$, $\mathcal{W}_3(t)$ and $\mathcal{W}_4(t)$ be the limiting zero-mean Gaussian process of W_1, W_2, $W_3(t)$ and $W_4(t)$ respectively whose weak convergences were proved in the appendix B of Cheng, Wei and Ying (1997). Note that there exist nondecreasing functions,

$\tilde{g}^+(u)$ and $\tilde{g}^-(u)$ such that $\tilde{g}\{h_0(u)+Z_i^T\beta\} = \tilde{g}^+(u)-\tilde{g}^-(u)$. Since both $-Y_i(u)\tilde{g}^+(u)$ and $I(Z_i \le z)$ are monotone functions on $\mathbf{R}$, their psedo dimensions are unity. From lemma 5.3 and theorem 4.8 of Pollard(1990), $I(Z_i \le z)Y_i(u)\tilde{g}^+(u)$ is manageable (Pollard 1990, p. 38). Similarly, $I(Z_i \le z)Y_i(u)\tilde{g}^-(u)$ is also manageable. Then $I(Z_i \le z)Y_i(u)\tilde{g}\{h_0(u) + Z_i^T\beta\}$ is also manageable(the lemma A2 of Bilias, Gu and Ying(1997)). So by the uniform law of large numbers(Pollard(1990), p.41), $\frac{1}{n}\sum_{i=1}^n I(Z_i \le z)Y_i(u)\tilde{g}\{h_0(u) + Z_i^T\beta\}$ converges uniformly to some nonrandom function almost surely. Since this fact and the continuity on $[0,\tau]$ of $\mathcal{W}_3(t)$ and $\mathcal{W}_4(t)$, which is proved in the later part of this appendix, ensure the conditions of the lemma 3.1, the 3rd term of (13) is asmptotically equivalent to the zero-mean Gaussian process $\int_0^t [\frac{b^T(u)}{a(u)}D(\mathcal{W}_1 + \mathcal{W}_2) + \frac{1}{a(u)}(\mathcal{W}_3(u) + \mathcal{W}_4(u))]d\lim\frac{1}{n}\sum_{i=1}^n I(Z_i \le z)Y_i(u)\tilde{g}\{h_0(X_i) + Z_i^T\beta\}$. The 2nd term of (13) converges to a zero-mean Gaussian process since $\frac{1}{n}\sum_{i=1}^n I(Z_i \le z)Y_i(u)\tilde{g}\{h_0(u) + Z_i^T\beta\}$ converges uniformly to some nonrandom function almost surely.

Next we show the weak convergence of (11).

$$
\begin{aligned}
(11) \;=\; & \frac{1}{n}\sum_{i=1}^n I(Z_i \le z)\int_0^t Y_i(u)\frac{1}{a(u)}\{b(u)+a(u)Z_i\}^T h_0'(u)\tilde{g}'\{h_0(u)+Z_i^T\beta\}du \\
\times\; & D(W_1+W_2) \qquad (14)\\
+\; & \frac{1}{n}\sum_{i=1}^n I(Z_i \le z)\int_0^t Y_i(u)\frac{1}{a(u)}\{W_3(u)+W_4(u)\} \\
\times\; & h_0'(u)\tilde{g}'\{h_0(u)+Z_i^T\beta\}du \qquad (15)
\end{aligned}
$$

Using the similar argument to that for the 1st term of (10), by the uniform law of large number(Pollard(1990), p. 41), $\frac{1}{n}\sum_{i=1}^n I(Z_i \le z)\int_0^t Y_i(u)\frac{1}{a(u)}\{b(u)+a(u)Z_i\}^T h_0'(u)\tilde{g}'\{h_0(u)+Z_i^T\beta\}du$ converges uniformly to some non-random function almost surely. And $D(W_1 + W_2)$ converge to Gaussian random variable. So (14) is tight and converges weakly to some Gauusian process. Furthermore, it can be shown by the similar argument to that for the 1st term (10) that $\frac{1}{n}\sum_{i=1}^n I(Z_i \le z)Y_i(u)\frac{1}{a(u)}h_0'(u)\tilde{g}'\{h_0(u) + Z_i^T\beta\}$ converges uniformaly to some non-random function almost surely by the uniform law of large number. Because of the weak convergence of $W_3(u) + W_4(u)$ to $\mathcal{W}_3(t) + \mathcal{W}_4(t)$ and the almost sure representation theorem(Pollard(1990), p. 45), $W_3(u) + W_4(u)$ can converges to $\mathcal{W}_3(u) + \mathcal{W}_4(u)$ uniformly and almost surely in some probability space. So

$$
\begin{aligned}
& \sup_{0\le t\le\tau,-1\le z\le 1}\left|\int_0^t f_n(z,u)(W_3(u)+W_4(u))du - \int_0^t f_\infty(z,u)(\mathcal{W}_3(u)+\mathcal{W}_4(u))du\right| \\
\le\; & \sup_{0\le t\le\tau,-1\le z\le 1}\left|f_n(z,u)(W_3(u)+W_4(u)) - f_\infty(z,u)(\mathcal{W}_3(u)+\mathcal{W}_4(u))\right|\tau,
\end{aligned}
$$

where $f_n(z,u) = \frac{1}{n}\sum_{i=1}^n I(Z_i \le z)Y_i(u)\frac{1}{a(u)}h_0'(u)\tilde{g}'\{h_0(u) + Z_i^T\beta\}$ and $f_\infty = \lim f_n(u)$. Again applying the almost sure presentation theorem, this means that (13) converges weakly to the Gaussian process $\int f_\infty(z,u)(\mathcal{W}_3(u) + \mathcal{W}_4(u))du$ in the original probability space.

Continuity of $\mathcal{W}_3(t)$ and $\mathcal{W}_4(t)$

It is sufficient to show $E|\mathcal{W}_3(t) - \mathcal{W}_3(s)|^4 \leq K|t-s|^2$ by the Kolmogorov-Centsov theorem.By simple algebra,

$$\begin{aligned} E|W_3(t) - W_3(s)|^4 &= E\Big[\frac{1}{n^2}|\sum_{i=1}^n (r_i(t) - r_i(s))|^4\Big] \\ &= \frac{1}{n}E|r_1(t) - r_1(s)|^4 \\ &+ \frac{n^2-n}{n^2}E|r_1(t) - r_1(s)|^2 E|r_2(t) - r_2(s)|^2 \end{aligned} \quad (16)$$

Because of the uniform boundedness of $r_1(t)$, the 1st term of (16) is dominated by $\frac{K}{n}$ where K is some constant. From the simple algebra, one can show $E|r_1(t) - r_1(s)|^2 \leq K^{\frac{1}{2}}|t-s|$. So $E|W_3(t) - W_3(s)|^4 \leq \frac{K}{n} + (1-\frac{1}{n})K|t-s|^2$ is obtained.Using Fatou's lemma, the aimed inequality, $E|\mathcal{W}_3(t) - \mathcal{W}_3(s)|^4 \leq K|t-s|^2$, is obtained.

Similar to $\mathcal{W}_3(t)$, it is sufficient to prove $E|\mathcal{W}_4(t) - \mathcal{W}_3(4)|^4 \leq K|t-s|^2$ to show the continuity of $\mathcal{W}_4(t)$.

$$\begin{aligned} E|W_4(t) - W_4(s)|^4 &= E\Big[\frac{1}{n^2}\Big|\sum_{i=1}^n \Big(\frac{\pi(t)}{G(t)}\int_0^t \frac{1}{\pi(u)}dM_i^c(u) - \frac{\pi(s)}{G(s)}\int_0^s \frac{1}{\pi(u)}dM_i^c(u)\Big)\Big|^4 \\ &= E\Big[\frac{1}{n^2}\sum_{i=1}^n \big(\frac{\pi(t)}{G(t)}\big)^4\Big|\int_s^t \frac{1}{\pi(u)}dM_i^c(u)\Big|^4 \\ &+ \frac{2}{n^2}\sum_{i\neq j}\Big(\frac{\pi(t)}{G(t)}\Big)^2\Big(\frac{\pi(t)}{G(t)} - \frac{\pi(s)}{G(s)}\Big)^2 \\ &\times \Big|\int_s^t \frac{1}{\pi(u)}dM_i(u)\Big|^2\Big|\int_0^s \frac{1}{\pi(u)}dM_i(u)\Big|^2\Big] \end{aligned} \quad (17)$$

The 1st term and 2nd terms are bounded by $\frac{K}{n}$ and $(1-\frac{1}{n})K|t-s|^2$ respectively. From the Fatou's lemma, $E|\mathcal{W}_4(t) - \mathcal{W}_4(s)|^4 \leq K|t-s|^2$ is obtained.

3.2 Weak convergence of $n^{-\frac{1}{2}}\hat{H}(t,z)$ conditional on the data

Here we show that conditional on the data $\{X_i, \Delta_i, Z_i\}$, $n^{-\frac{1}{2}}\hat{H}(t,z)$ converges weakly to the unconditional limiting Gaussian process of $n^{-\frac{1}{2}}\tilde{H}(t,z)$ when the fitted model is correctly specified. It is straightforward to show the covariance function of $n^{-\frac{1}{2}}\hat{H}(t,z)$ converges almost surely to that of the limiting Gaussian process of $n^{-\frac{1}{2}}\tilde{H}(t,z)$. So the finite dimensional conditional distribution of the $n^{-\frac{1}{2}}\tilde{H}(t,z)$ converges to that of he limiting Gaussian process of $n^{-\frac{1}{2}}\hat{H}(t,z)$. To show the weak convergence, it is sufficient to show the tightness of the each term of (6). To this end, it is sufficient to show the moment inequalities(Bickel and Wichura(1971)). Define

$$\begin{aligned} \Psi_l^{(1)}(z_1, z_2 : t_1, t_2) &= \{\hat{M}_l(t_2) - \hat{M}_l(t_1)\}I(z_1 \leq Z_l \leq z_2) \\ \Psi_{+,l}^{(2)}(z_1, z_2 : t_1, t_2) &= \sum_{i=1}^n\sum_{j=1}^n \int_{t_1}^{t_2} Y_l(u) Z_{ij}\hat{e}_{ij}(\hat{\beta}) \end{aligned}$$

$$\times d\Big[\tilde{g}^{+}\{\hat{h}(u+)+Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\{\hat{b}(u+)+\hat{a}(u+)Z_l\}^T\hat{D}\Big]$$
$$\times I(z_1 \le Z_l \le z_2)$$

$$\Psi^{(3)}_{+,l}(z_1,z_2:t_1,t_2) = 2n^{-\frac{3}{2}}\sum_{i=1}^{n}\int_0^t I(Z_l\le z)Y_l(u)\int_0^\infty \frac{\hat{q}(u)}{\hat{\pi}(u)}d\hat{M}_i^c(u)$$
$$\times d\Big[\tilde{g}^{+}\{\hat{h}(u+)+Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\{\hat{b}(u+)+\hat{a}(u+)Z_l\}^T\hat{D}\Big]$$
$$\times I(z_1 \le Z_l \le z_2)$$

$$\Psi^{(4)}_{+,l}(z_1,z_2:t_1,t_2) = n^{-\frac{3}{2}}\sum_{i=1}^{n}\int_0^t I(Z_l\le z)Y_l(u)$$
$$\times d\Big[\tilde{g}^{+}\{\hat{h}(u+)+Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\hat{r}_i(u+)\Big]$$
$$\times I(z_1 \le Z_l \le z_2)$$

$$\Psi^{(5)}_{+,l}(z_1,z_2:t_1,t_2) = n^{-\frac{3}{2}}\sum_{i=1}^{n}\int_0^t I(Z_l\le z)Y_l(u)$$
$$\times d\Big[\tilde{g}^{+}\{\hat{h}(u)+Z_l^T\hat{\beta}\}\frac{1}{\hat{a}(u+)}\frac{\hat{\pi}(u+)}{\hat{G}(u+)}\int_0^u \frac{1}{\hat{\pi}(x)}d\hat{M}_i^c(x)\Big]$$
$$\times I(z_1 \le Z_l \le z_2) \tag{18}$$

where $\tilde{g}^{+}\{\hat{h}(u+)+Z_l^T\hat{\beta}\}$ and $\tilde{g}^{-}\{\hat{h}(u+)+Z_l^T\hat{\beta}\}$ are nondecreasing functions such that $\tilde{g}=\tilde{g}^{+}-\tilde{g}^{-}$. $\Psi^{(k)}_{-,l}(z_1,z_2:t_1,t_2)$ are defined in the similar matter. By the boundedness of $M_l(t)$ and $g\{\hat{h}(u+)+Z_l^T\hat{\beta}\}$ on $[0,\tau]$,

$$\begin{aligned}\Big\{\Psi_i^{(1)}(z_1,z_2:t_1,t)\Big\}^2\Big\{\Psi_j^{(1)}(z_1,z_2:t,t_2)\Big\}^2 &\le (M_i(t_2)-M_i(t))^2(M_j(t)-M_j(t_1))^2\\ &\times I(z_1\le Z_i\le z_2)I(z_1\le Z_j\le z_2)\\ &+ L_{11}|t_2-t||t-t_1|I(z_1\le Z_i\le z_2)I(z_1\le Z_j\le z_2).\end{aligned} \tag{19}$$

where L_{11} is some constant. Noting that conditional on the data, $\{\mathcal{L}_l\}$ is the only random elements of each terms of (6), for $\forall\ z_1<z<z_2, t_1<t<t_2$,

$$\begin{aligned}&E\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_l^{(1)}(z_1,z_2:t_1,t)\mathcal{L}_l\Big|^2\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_l^{(1)}(z_1,z_2:t,t_2)\mathcal{L}_l\Big|^2\\
=\ &\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Psi_i^{(1)}(z_1,z_2:t_1,t)^2\Psi_j^{(1)}(z_1,z_2:t,t_2)^2E[\mathcal{L}_i^2\mathcal{L}_j^2]\\
+\ &\frac{4}{n^2}\sum_{i_1<i_2,i_3<i_4}\Psi_{i_1}^{(1)}(z_1,z_2:t_1,t)\Psi_{i_2}^{(1)}(z_1,z_2:t_1,t)\\
\times\ &\Psi_{i_3}^{(1)}(z_1,z_2:t,t_2)\Psi_{i_4}^{(1)}(z_1,z_2:t,t_2)E[\mathcal{L}_{i_1}\mathcal{L}_{i_2}\mathcal{L}_{j_1}\mathcal{L}_{j_2}]\\
\le\ &\frac{3}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(M_i(t_2)-M_i(t))^2(M_j(t)-M_j(t_1))^2I(z_1\le Z_i\le z_2)I(z_1\le Z_j\le z_2)\end{aligned}$$

$$+ \quad 3L_{11}|t_2 - t||t - t_1|\frac{1}{n}\sum_{i=1}^{n} I(z_1 \le Z_i \le z_2)\frac{1}{n}\sum_{j=1}^{n} I(z_1 \le Z_j \le z_2), \tag{20}$$

where the last inequality holds because of (19) and $E[\mathcal{L}_{i_1}\mathcal{L}_{i_2}\mathcal{L}_{j_1}\mathcal{L}_{j_2}] = 0$ if $i_1 < i_2, j_1 < j_2$. Since the 1st term vanishes as n tends to ∞ by the independent incremant property of martingale, the moment inequality was obtained (theorem 3 and p.1666 of Bickel and Wichura (1971)). Similarly

$$\begin{aligned} & E\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_{+,l}^{(1)}(z_1, z : t_1, t_2)\mathcal{L}_l\Big|^2\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_{+,l}^{(1)}(z, z_2 : t, t_2)\mathcal{L}_l\Big|^2 \\ \le \quad & \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(M_i(t_2) - M_i(t_1))^2(M_j(t_2) - M_j(t_1))^2 I(z_1 \le Z_i \le z)I(z \le Z_j \le z_2) \\ + \quad & 3L_{11}|t_2 - t_1|^2\frac{1}{n}\sum_{i=1}^{n} I(z_1 \le Z_i \le z)\frac{1}{n}\sum_{i=1}^{n} I(z \le Z_j \le z_2), \end{aligned} \tag{21}$$

can be obtained. Since the 1st term converges to

$$\begin{aligned} & \Big\{E\Big[\Big(\int_{t_1}^{t_2} dM_1(u)\Big)^2 I(z_1 \le Z_1 \le z)\Big]\Big\}^2 \\ & = \Big\{E\Big[I(z_1 \le Z_1 \le z)E\Big[\Big(\int_{t_1}^{t_2} dM_1(u)\Big)^2\Big|Z_1\Big]\Big]\Big\}^2 \\ & \le -\Big\{E\Big[I(z_1 \le Z_1 \le z)\int_{t_1}^{t_2} Y_1(u)d\log(g\{h_0(u) + Z_1^T\beta_0\})\Big]\Big\}^2. \end{aligned}$$

Then the 1st term of (21) converges to smoe measure whose marginal is continuous and thereby tightness is obtained (Bickel and Wichura 1971 theorem3 and p.1666).
Since $Y_l(u), Z_{ij}, \hat{e}_{ij}(\hat{\beta}), \frac{1}{\hat{a}(u+)}\{\hat{b}(u+) + \hat{a}(u+)Z_l\}^T\hat{D}, \hat{q}(u), \hat{\pi}(u), \hat{M}_i^c(u)$ and $\frac{1}{\hat{G}(u+)}$ are bounded, for k=2,3,4,5,

$$\begin{aligned} \Big|\Psi_{+,l}^{(k)}(z_1, z_2 : t_1, t_2)\Big| \quad & \le \quad K_k\Big|\tilde{g}^+\{\hat{h}(t_2+) + Z_l^T\hat{\beta}\} - \tilde{g}^+\{\hat{h}(t_1+) + Z_l^T\hat{\beta}\}\Big| \\ & \times \quad I(z_1 \le Z_l \le z_2) \\ & \le \quad \tilde{K}_k|\hat{h}(t_2) - \hat{h}(t_1)|I(z_1 \le Z_l \le z_2), \end{aligned} \tag{22}$$

where K_k and $\tilde{K}_k$ are some constants. Noting that conditional on the data, $\{\mathcal{L}_l\}$ is the only random elements of each terms of 6, for $\forall\ z_1 < z < z_2, t_1 < t < t_2$,

$$\begin{aligned} & E\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_{+,l}^{(k)}(z_1, z_2 : t_1, t)\mathcal{L}_l\Big|^2\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_{+,l}^{(k)}(z_1, z_2 : t, t_2)\mathcal{L}_l\Big|^2 \\ \le \quad & \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Psi_{+,i}^{(k)}(z_1, z_2 : t_1, t)^2\Psi_{+,j}^{(k)}(z_1, z_2 : t, t_2)^2 E[\mathcal{L}_i^2\mathcal{L}_j^2] \\ + \quad & \frac{1}{n^2}\frac{4}{n^2}\sum_{i_1 < i_2, i_3 < i_4}\Psi_{+,i_1}^{(k)}(z_1, z_2 : t_1, t)\Psi_{+,i_2}^{(k)}(z_1, z_2 : t_1, t) \qquad (23) \\ \times \quad & \Psi_{+,i_3}^{(k)}(z_1, z_2 : t, t_2)\Psi_{+,i_4}^{(k)}(z_1, z_2 : t, t_2)E[\mathcal{L}_{i_1}\mathcal{L}_{i_2}\mathcal{L}_{j_1}\mathcal{L}_{j_2}] \\ \le \quad & L_{k1}\Big|\hat{h}(t) - \hat{h}(t_1)\Big|^2\Big|\hat{h}(t_2) - \hat{h}(t)\Big|^2\frac{1}{n}\sum_{i=1}^{n} I(z_1 \le Z_i \le z_2), \qquad (24) \end{aligned}$$

where L_{k1} is some constant and the last inequality holds because of (22) and $E[\mathcal{L}_{i_1}\mathcal{L}_{i_2}\mathcal{L}_{j_1}\mathcal{L}_{j_2}] = 0$ if $i_1 \neq i_2$ and $j_1 \neq j_2$. Similarly,

$$
\begin{aligned}
& E\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_{+,l}^{(k)}(z_1, z : t_1, t_2)\mathcal{L}_l\Big|^2\Big|n^{-\frac{1}{2}}\sum_{l=1}^{n}\Psi_{+,l}^{(k)}(z, z_2 : t, t_2)\mathcal{L}_l\Big|^2 \\
\leq \;& L_{k2}\Big|\hat{h}(t_2) - \hat{h}(t_1)\Big|^2\frac{1}{n}\sum_{i=1}^{n} I(z_1 \leq Z_i \leq z)\frac{1}{n}\sum_{i=1}^{n} I(z \leq Z_i \leq z_2).
\end{aligned} \tag{25}
$$

From (24) and (25), the tightness of the 2nd-5th term of (6) is obtained (theorem3 and p.1666 of Bickel and Wichura(1971)). The moment inequalities for $\Psi_{-,l}^{(k)}(z_1, z_2 : t_1, t_2)$ are obtained in the same matter.

参考文献

[1] S. Bennett. Analysis of survival data by the proportional odds model. *Statistics in Medicine*, Vol. 2, pp. 273–277, 1982.

[2] S.C. Cheng and N. Wang. Linear transformation models for faillure time data with covariate measurement error. *Journal of the American Statistical Association.*, Vol. 96, No. 454, pp. 706–16, 2001.

[3] S.C. Cheng, L.J. Wei, and Z. Ying. Analysis of transformation models with censored data. *Biometrika.*, Vol. 82, No. 4, pp. 835–45, 1995.

[4] S.C. Cheng, L.J. Wei, and Z. Ying. Predicting survival probabilities with semiparametric transformation models. *Journal of the American Statistical Association.*, Vol. 92, No. 437, pp. 227–35, 1997.

[5] J.P. Fine, L.J. Wei, and Z. Ying. On the lilnear transformation model for censored data. *Biometrika.*, Vol. 85, No. 4, pp. 980–986, 1998.

[6] T.R. Fleming and D.P. Harrington. *Counting Processes and Survival Analysis.* John Wiley and Sons., 1991.

[7] D.Y. Lin, L.J Wei, and Z. Ying. Regression diagnostics based on cumulative residuals. preprint.

[8] D.Y. Lin, L.J. Wei, and Z. Ying. Checking the cox model with cumulative sums of martingale-based residuals. *Biometrika.*, Vol. 80, No. 3, pp. 557–72, 1993.

[9] D.Y. Lin, L.J. Wei, and Z. Ying. Semiparametric transformation models for point processes. *Journal of the American Statistical Association.*, Vol. 96, No. 454, pp. 620–28, 2001.

[10] D.Y. Lin and Z. Ying. Semiparametric analysis of the additive risk model. *Biometrika.*, Vol. 81, No. 1, pp. 66–71, 1994.

[11] D. Pollard. *Empirical Processes:Theory and Applications.* Institute of Mathematical Statistics., Hayward,California., 1998. NSF-CBMS Regional Conference Series in Probability and Statistics Volume2.

[12] C.F. Spiekerman and D.Y. Lin. Checking the marginal cox model for correlated failure time data. *Biometrika.*, Vol. 83, No. 1, pp. 143–56, 1996.

[13] A.A. Tsiatis. Estimating regression parameters using linear rank tests for censored data. *The Annals of Statistics.*, Vol. 18, No. 1, pp. 354–372, 1990.

[14] L.J. Wei, Z. Ying, and D.Y. Lin. Linear regression analysis of censored survival data based on rank test. *Biometrika.*, Vol. 77, No. 4, 1990.

[15] Z. Ying. A large sample study of rank estimation for censored regression data. *The Annals of Statistics.*, Vol. 21, No. 1, pp. 76–99, 1993.