SMOOTH COMPOSITE LIKELIHOOD ANALYSIS OF LENGTH-BIASED AND RIGHT-CENSORED DATA WITH AFT MODEL

Xuerong Chen[1], Na Hu[2] and Jianguo Sun[3,4]

[1] *Center of Statistical Research, School of Statistics, Southwestern University of Finance and Economics, Chengdu, 611130, China*
[2] *Boehringer Ingelheim (China) Investment Co., Ltd., Shanghai, 200040, China*
[3] *Department of Statistics, University of Missouri, Columbia, Missouri, 65201, U.S.A*
[4] *Institute of Mathematics, Jilin University, Changchun, 130012, China*

Supplementary Material

Appendix I: Proof of the Asymptotic Properties.

In this supplementary material, we will sketch the proofs for the asymptotic properties of the proposed estimator of regression parameters. For this, we first give Lemma 1 below, which is needed below.

Lemma 1. *Suppose that $f(Z)$ and $E\{\phi(Z,u)|Z=z\}$ are continuous and twice differentiable at z and $E\{|\phi(Z,U)|^2<\infty\}$. Then as $n\to\infty$, we have*

$$\sup_{Z\in\mathcal{Z}}\left|\frac{1}{nb_n}\sum_{i=1}^{n}K\left(\frac{Z_i-z}{b_n}\right)\phi(Z_i,U_i)-E\left[\frac{1}{b_n}K\left(\frac{Z-z}{b_n}\right)\phi(Z,U)\right]\right|=O(\bar{\delta}_n)\quad a.s.,$$

where $\mathcal{Z}$ is the support of Z, $\bar{\delta}_n=b_n^2+(\frac{\log b_n^{-1}}{nb_n})^{1/2}$.

Lemma 1 could be proved along the lines of Lemma A.2 in Xia and Li (1999).

Proof of Theorem 1: To prove Theorem 1, by Theorem 5.7 in Van der Vaart (1998), it is suffices to prove that:

(1). $\sup_{\boldsymbol{\beta}\in\mathcal{B}}\|\ell_n^s(\boldsymbol{\beta})-\ell(\boldsymbol{\beta})\|=o_p(1)$

(2). $\boldsymbol{\beta}_0$ is the unique maximizer of $\ell(\boldsymbol{\beta})$.

The proof of statement (1) can be obtained by the following results,

$$\begin{aligned}
J_1 &= \sup_{\boldsymbol{\beta}\in\mathcal{B},s}\left|\frac{1}{nb_n}\sum_{i=1}^{n}\delta_iK(\frac{R_i(\boldsymbol{\beta})-s}{b_n})-\frac{dP(\delta=1,R(\boldsymbol{\beta})\le s)}{ds}\right|=o_p(1),\\
J_2 &= \sup_{\boldsymbol{\beta}\in\mathcal{B},s}\left|\frac{1}{nb_n}\sum_{i=1}^{n}\int_{\frac{R_i(\beta)-s}{b_n}}^{\infty}K(u)du-P(R(\boldsymbol{\beta})\ge s)\right|=o_p(1),
\end{aligned}$$

$$
\begin{aligned}
J_3 &= \sup_{\boldsymbol{\beta}\in\mathcal{B},s} |\frac{1}{nb_n}\sum_{i=1}^{n}\int_{\frac{H_i(\beta)-s}{b_n}}^{\infty} K(u)du - P(H(\boldsymbol{\beta})\geq s)| = o_p(1),\\
J_4 &= \sup_{\boldsymbol{\beta}\in\mathcal{B},s} |\frac{1}{nb_n}\sum_{i=1}^{n}\int_{\frac{R_i(\beta)-s}{b_n}}^{\infty} \delta_i K(u)du - P(\delta=1, R(\boldsymbol{\beta})\geq s)| = o_p(1),\\
J_5 &= \sup_{\boldsymbol{\beta}\in\mathcal{B},s} |\frac{1}{nb_n}\sum_{i=1}^{n}\int_{\frac{I_i(\beta)-s}{b_n}}^{\infty} \delta_i K(u)du - P(\delta=1, I(\boldsymbol{\beta})\geq s)| = o_p(1).
\end{aligned}
$$

We just need to verify $J_1 = o_p(1)$. The remaining terms can be obtained in similar way. By Lemma 2.8 of Pakes and Pollard (1989), it is sufficient to verify that $\epsilon_1 = \{\delta K(\frac{R(\boldsymbol{\beta})-s}{b_n}), \boldsymbol{\beta}\in\mathcal{B}, s\}$ is an Euclidean class with an integrable envelope function. The Euclidean property of ϵ_1 can be obtained by the Euclidean properties of $\{\delta\}$, $\{K(\frac{R(\boldsymbol{\beta})-s}{b_n}, \boldsymbol{\beta}, s\}$ with constant envelope $F_1 = 1$, $F_2 = \sup|K(.)|$. By example (2.10) in Pakes and Pollard (1989), and condition (C5), the latter class has Euclidean property. By Lemma 22(ii) in Nolan and Pollard (1987), the former class is Euclidean class as the indicator function is bounded variation function. Hence, $J_1 = o_p(1)$. By similar argument, we have $J_i = o_p(1), i = 2,3,4,5$. Therefore, statement (1) holds.

For statement (2), note that $\ell(\boldsymbol{\beta})$ has unique maximizer is equivalent to that the score function has unique root. We just need to verify that:

$$
\sup_{\boldsymbol{\beta}\in\{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|\geq\delta\}} \left\|\frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right\| > 0.
$$

Note that by condition (C4),

$$
\begin{aligned}
\inf_{\boldsymbol{\beta}\in\{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|\geq\delta\}} \left\|\frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right\| &= \inf_{\boldsymbol{\beta}\in\{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|\geq\delta\}} \left\|\frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} - \frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}\right\|\\
&= \inf_{\boldsymbol{\beta}\in\{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|\geq\delta\}} \|\ell^{(2)}(\boldsymbol{\beta}^*)(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\|\\
&= \inf_{\boldsymbol{\beta}\in\{\|\boldsymbol{\beta}-\boldsymbol{\beta}_0\|\geq\delta\}} \left\|\nabla^2_{\boldsymbol{\beta}} E\left\{\delta\log\left(\frac{dP(\delta=1, Ye^{\boldsymbol{\beta}^T\mathbf{X}}\leq t)/dt}{N_1 - N2 + N_3 - N_4}\right)\right\}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\right\|\\
&> 0,
\end{aligned}
$$

where $\boldsymbol{\beta}^*$ lies between $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_0$. This completes the proof.

Proof of Theorem 2: For simplicity, for any functions f_1, f_2 and f, denote $f_1(Y,X;\boldsymbol{\beta}) + f_2(Y,X;\boldsymbol{\beta}) = \{f_1+f_2\}(Y,X;\boldsymbol{\beta})$ and define $f^{(1)}$ to be the first derivative of f. Since $\frac{\partial\ell_n^s(\widehat{\boldsymbol{\beta}})}{\partial\boldsymbol{\beta}} = 0$, we have

$$
\frac{1}{n}\sum_{i=1}^{n}\left\{\delta_i X - \delta_i R_i^{(1)}(\widehat{\boldsymbol{\beta}}) + \delta_i\frac{g_{1n}(Y_i,X_i;\widehat{\boldsymbol{\beta}})}{g_{2n}(Y_i,X_i;\widehat{\boldsymbol{\beta}})} - \delta_i\frac{\{g_{3n}-g_{4n}+g_{5n}-g_{6n}\}(Y_i,X_i;\widehat{\boldsymbol{\beta}})}{\{g_{7n}-g_{8n}+g_{9n}-g_{10n}\}(Y_i,X_i;\widehat{\boldsymbol{\beta}})}\right\} = 0,
$$

where

$$\begin{aligned}
\psi(y,x;\boldsymbol{\beta}) &= \log(y\exp(\boldsymbol{\beta}^T x)), \quad \psi^{(1)} = \frac{\partial\phi(y,x,\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}, \\
g_{1n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\delta_j K^{(1)}(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n})\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n^2}, \\
g_{2n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\frac{\delta_j}{b_n}K(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}), \\
g_{3n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n}K(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}), \\
g_{4n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\frac{H_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n}K(\frac{H_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}), \\
g_{5n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\delta_j\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n}K(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}), \\
g_{6n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\delta_j\frac{I_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n}K(\frac{I_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}), \\
g_{7n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\int_{-\infty}^{\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}}K(s)ds, \\
g_{8n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\int_{-\infty}^{\frac{H_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}}K(s)ds, \\
g_{9n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\delta_j\int_{-\infty}^{\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}}K(s)ds, \\
g_{10n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\delta_j\int_{-\infty}^{\frac{I_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}}K(s)ds,
\end{aligned}$$

Furthermore, we denote the expectations of $g_{k,n}(y,x;\boldsymbol{\beta}), k=1,...,10$ as $g_{k,0}(y,x;\boldsymbol{\beta})$.

Note that

$$\begin{aligned}
g_{1n}(y,x;\boldsymbol{\beta}) &= \frac{1}{n}\sum_{j=1}^{n}\delta_j\lim_{h\to 0}\frac{K(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}+\frac{h}{b_n})-K(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n})}{h/b_n}\times\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n^2} \\
&= \lim_{h\to 0}\frac{b_n}{h}\left[\frac{1}{n}\sum_{j=1}^{n}\delta_j K\left(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})+h}{b_n}\right)\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n^2}\right. \\
&\quad\left.-\frac{1}{n}\sum_{j=1}^{n}\delta_j K\left(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}\right)\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n^2}\right] \\
&=: \lim_{h\to 0}\frac{1}{h}\{J_1+J_2\},
\end{aligned}$$

and

$$\begin{aligned} EJ_1 &= E\left\{\delta_j K\left(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}+\frac{h}{b_n}\right)\times\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n}\right\} \\ &= E\left\{E\left[\delta_j K\left(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}+\frac{h}{b_n}\right)\times\frac{R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta})}{b_n}\Big|_{\delta_j,R_j^{(1)}(\boldsymbol{\beta})}\right]\right\} \\ &= E\left\{\delta_j(R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta}))E\left[K\left(\frac{R_j(\boldsymbol{\beta})-\psi(y,x;\boldsymbol{\beta})}{b_n}+\frac{h}{b_n}\right)\frac{1}{b_n}\Big|_{\delta_j,R_j^{(1)}(\boldsymbol{\beta})}\right]\right\} \\ &= E\{\delta_j(R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta}))\{f_{R_j(\boldsymbol{\beta})|R_j^{(1)}(\boldsymbol{\beta}),\delta_j}(\psi(y,x;\boldsymbol{\beta})-h)+O(b_n^2)\}\}, \end{aligned}$$

where $f_{R(\boldsymbol{\beta})|R^{(1)}(\boldsymbol{\beta}),\delta}$ is the conditional density function of $R(\boldsymbol{\beta})$ given $R^{(1)}(\boldsymbol{\beta})$ and δ.

Similarly, we have

$$EJ_2 = E\{\delta_j(R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta}))\{f_{R_j(\boldsymbol{\beta})|R_j^{(1)}(\boldsymbol{\beta}),\delta_j}(\psi(y,x;\boldsymbol{\beta}))+O(b_n^2)\}\}.$$

By lemma 1, we have $\sup_{R_j(\boldsymbol{\beta})\in\chi}|J_1-EJ_1| = O(\bar{\delta}_n)$ and $\sup_{R_j(\boldsymbol{\beta})\in\chi}|J_2-EJ_2| = O(\bar{\delta}_n)$, where χ is the support of $R(\boldsymbol{\beta})$ and $\bar{\delta}_n = b_n^2+\frac{\log b_n^{-1}}{nb_n}$. Hence $J_1 = EJ_1+O(\bar{\delta}_n)$ and $J_2 = EJ_2+O(\bar{\delta}_n)$ and

$$\begin{aligned} g_{1n}(y,x;\boldsymbol{\beta}) &= \lim_{h\to 0}\frac{1}{h}\{J_1-J_2\} \\ &= \lim_{h\to 0}\frac{1}{h}\left\{E\left[\delta_j(R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta}))[f_{R_j(\boldsymbol{\beta})|R_j^{(1)}(\boldsymbol{\beta}),\delta_j}(\psi(y,x;\boldsymbol{\beta})-h)\right.\right. \\ &\qquad \left.\left.-f_{R_j(\boldsymbol{\beta})|R_j^{(1)}(\boldsymbol{\beta}),\delta_j}(\psi(y,x;\boldsymbol{\beta}))+O(b_n^2)]\right]\right\} \\ &\to -E\{\delta_j(R_j^{(1)}(\boldsymbol{\beta})-\psi^{(1)}(y,x;\boldsymbol{\beta}))f^{(1)}_{R_j(\boldsymbol{\beta})|R_j^{(1)}(\boldsymbol{\beta}),\delta}(\psi(y,x;\boldsymbol{\beta}))\}. \end{aligned}$$

Therefore,

$$g_{1n}(y,x;\boldsymbol{\beta}_0) \to -E\{\delta_j(R_{j0}^{(1)}-\psi^{(1)}(y,x;\boldsymbol{\beta}_0))f^{(1)}_{R_{j0}|R_{j0}^{(1)},\delta_j}(\psi(y,x;\boldsymbol{\beta}_0))\},$$

and similarly we can obtain that

$$\begin{aligned} g_{2n}(y,x;\boldsymbol{\beta}_0) &\to E[\delta_j f_{R_{j0}|\delta_j}(\psi(y,x;\boldsymbol{\beta}_0))], \\ g_{3n}(y,x;\boldsymbol{\beta}_0) &\to E[(R_{j0}^{(1)}-\psi^{(1)}(y,x;\boldsymbol{\beta}_0))f_{R_{j0}|R_{j0}^{(1)},\delta_j}(\psi(y,x;\boldsymbol{\beta}_0))], \\ g_{4n}(y,x;\boldsymbol{\beta}_0) &\to E[(H_{j0}^{(1)}-\psi^{(1)}(y,x;\boldsymbol{\beta}_0))f_{H_{j0}|H_{j0}^{(1)},\delta_j}(\psi(y,x;\boldsymbol{\beta}_0))], \\ g_{5n}(y,x;\boldsymbol{\beta}_0) &\to E[\delta_j(R_{j0}^{(1)}-\psi^{(1)}(y,x;\boldsymbol{\beta}_0))f_{R_{j0}|R_{j0}^{(1)},\delta_j}(\psi(y,x;\boldsymbol{\beta}_0))], \\ g_{6n}(y,x;\boldsymbol{\beta}_0) &\to E[\delta_j(I_{j0}^{(1)}-\psi^{(1)}(y,x;\boldsymbol{\beta}_0))f_{I_{j0}|I_{j0}^{(1)},\delta_j}(\psi(y,x;\boldsymbol{\beta}_0))], \\ g_{7n}(y,x;\boldsymbol{\beta}_0) &\to P(R_{j0}>\psi(y,x;\boldsymbol{\beta}_0)), \\ g_{8n}(y,x;\boldsymbol{\beta}_0) &\to P(H_{j0}>\psi(y,x;\boldsymbol{\beta}_0)), \\ g_{9n}(y,x;\boldsymbol{\beta}_0) &\to P(R_{j0}>\psi(y,x;\boldsymbol{\beta}_0),\delta_j=1), \\ g_{10n}(y,x;\boldsymbol{\beta}_0) &\to P(I_{j0}>\psi(y,x;\boldsymbol{\beta}_0),\delta_j=1), \end{aligned}$$

where $R_{j0} = R_j(\boldsymbol{\beta}_0), R_{j0}^{(1)} = R_j^{(1)}(\boldsymbol{\beta}_0)$, $H_{j0}, H_{j0}^{(1)}$ and $I_{j0}, I_{j0}^{(1)}$ have same definition. $f_{H_{j0}|H_{j0}^{(1)},\delta_j}, f_{I_{j0}|I_{j0}^{(1)},\delta_j}$ have same definition as $f_{R_j(\boldsymbol{\beta}_0)|R_j^{(1)}(\boldsymbol{\beta}_0),\delta_j}$.

By Taylor expansion, we can obtain

$$\begin{aligned}
0 &= \sqrt{n}\frac{\partial \ell_n^s(\widehat{\boldsymbol{\beta}})}{\partial \boldsymbol{\beta}} \\
&= \sqrt{n}\frac{\partial \ell_n^s(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0} + \frac{\partial \ell_n^{s2}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0) + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{\delta_i X - \delta_i R_i^{(1)}(\boldsymbol{\beta}_0) + \delta_i\frac{g_{10}(Y_i,X_i;\boldsymbol{\beta}_0)}{g_{20}(Y_i,X_i;\boldsymbol{\beta}_0)} - \delta_i\frac{\{g_{30}-g_{40}+g_{50}-g_{60}\}(Y_i,X_i;\boldsymbol{\beta}_0)}{\{g_{70}-g_{80}+g_{90}-g_{100}\}(Y_i,X_i;\boldsymbol{\beta}_0)}\right\} \\
&\quad + \frac{\partial \ell_n^{s2}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0) + o_p(1).
\end{aligned}$$

Hence it is easy to obtain that

$$\begin{aligned}
\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0) &= -\left(\frac{\partial \ell_n^{s2}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{\delta_i X - \delta_i R_i^{(1)}(\boldsymbol{\beta}_0) + \delta_i\frac{g_{10}(Y_i,X_i;\boldsymbol{\beta}_0)}{g_{20}(Y_i,X_i;\boldsymbol{\beta}_0)}\right. \\
&\quad \left. -\delta_i\frac{\{g_{30}-g_{40}+g_{50}-g_{60}\}(Y_i,X_i;\boldsymbol{\beta}_0)}{\{g_{70}-g_{80}+g_{90}-g_{100}\}(Y_i,X_i;\boldsymbol{\beta}_0)}\right\} + o_p(1), \\
&= -\left(\frac{\partial \ell^{2}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}\right)^{-1}\frac{1}{\sqrt{n}}\sum_{i=1}^n\left\{\delta_i X - \delta_i R_i^{(1)}(\boldsymbol{\beta}_0) + \delta_i\frac{g_{10}(Y_i,X_i;\boldsymbol{\beta}_0)}{g_{20}(Y_i,X_i;\boldsymbol{\beta}_0)}\right. \\
&\quad \left. -\delta_i\frac{\{g_{30}-g_{40}+g_{50}-g_{60}\}(Y_i,X_i;\boldsymbol{\beta}_0)}{\{g_{70}-g_{80}+g_{90}-g_{100}\}(Y_i,X_i;\boldsymbol{\beta}_0)}\right\} + o_p(1).
\end{aligned}$$

Thus it follows that $\sqrt{n}(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)$ converges to the normal distribution with mean 0 and the variance-covariance matrix that can be consistently estimated by $\widehat{\mathbf{A}}^{-1}\widehat{\mathbf{V}}\widehat{\mathbf{A}}^{-1}$. The proof of this theorem is completed.

Proof of Theorem 3: By the Taylor expansion and Theorem 1, we have

$$\begin{aligned}
\widehat{\Lambda}_n(t) &= \int_{-\infty}^{\log t}\frac{2(nb_nt)^{-1}\sum_{i=1}^n\delta_i K(\frac{R_i(\boldsymbol{\beta}_0)-u}{b_n})}{\frac{1}{n}\sum_{i=1}^n\int_{(H_i(\boldsymbol{\beta}_0)-u)/b_n}^{(R_i(\boldsymbol{\beta}_0)-u)/b_n}K(s)ds + \delta_i\int_{(I_i(\boldsymbol{\beta}_0)-u)/b_n}^{(R_i(\boldsymbol{\beta}_0)-u)/b_n}K(s)ds}du \\
&\quad + \frac{\partial\widehat{\Lambda}_n(t)}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}(\widehat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}_0) + o_p(|\widehat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}_0|) \\
&= \int_{-\infty}^{\log t}2\frac{dP(\delta=1,R(\boldsymbol{\beta})\le u)/du}{P(H(\boldsymbol{\beta})\le u\le R(\boldsymbol{\beta}))+P(\delta=1,I(\boldsymbol{\beta})\le u\le R(\boldsymbol{\beta}))}du + o_p(1).
\end{aligned}$$

Note that $u = \log t$ in the equation above. By the third equation on page 1194 of Shen, Ning, and Qin (2009) and the third equation on page 955 of Huang and Qin (2012), we have

$$\widehat{\Lambda}_n(t) = \int_{-\infty}^{t}\frac{f_{e^\epsilon|\mathbf{X}}(s|\mathbf{X})}{S_{e^\epsilon|\mathbf{X}}(s|\mathbf{X})}du + o_p(1) = \Lambda(t) + o_p(1),$$

where $f_{e^\epsilon|\mathbf{X}}(u|\mathbf{X})$ and $S_{e^\epsilon|\mathbf{X}}(u|\mathbf{X})$ are the density function and survival function of e^ϵ given $\mathbf{X}$, respectively. Hence, the proof is completed.

Reference

Huang, C. Y. and Qin, J. (2012). Composite partial likelihood estimation under length-biased sampling, with application to a prevalent cohort study of dementia. *J. Amer. Statist. Assoc.*, **107**, 946-957.

Nolan, D. and Pollard, D. (1987). U-processes: rates of convergence. *Ann. Statist.*, **15**, 780-799.

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, **57**, 1027-1057.

Shen, Y., Ning, J., and Qin, J. (2009). Analyzing length-biased data with semiparametric transformation and accelerated failure time models. *J. Amer. Statist. Assoc.*,, **104**, 1192-1202.

Van der Vaart, A. W.(1998). *Asymptotic Statistics*, New York, USA: Cambridge University Press.

Xia, Y. C. and Li, W. K. (1999). On single-index coeffcient regression models. *J. Amer. Statist. Assoc.*,, **94**, 1275-1285.