0.4pt=0pt

ASYMPTOTIC NORMALITY OF NONPARAMETRIC M-ESTIMATORS WITH APPLICATIONS TO HYPOTHESIS TESTING FOR PANEL COUNT DATA

The Hong Kong Polytechnic University and Indiana University

Supplementary Material

PROOFS OF THEOREMS

S1 Proof of Theorem 2.1

By assumptions A4 and A5 with A2, we have

$$-\sqrt{n}\dot{G}_{\Lambda_0}(\hat{\Lambda}_n - \Lambda_0)[h] = -\sqrt{n}G(\hat{\Lambda}_n)[h] + o_p(1). \tag{S1.1}$$

By assumptions A1 and A2, we have

$$-\sqrt{n}G(\hat{\Lambda}_n)[h] = \sqrt{n}(G_n - G)(\Lambda_0)[h] + o_p(1). \tag{S1.2}$$

Thus, it follows from (S1.1) and (S1.2) that

$$-\sqrt{n}\dot{G}_{\Lambda_0}(\hat{\Lambda}_n - \Lambda_0)[h] = \sqrt{n}(G_n - G)(\Lambda_0)[h] + o_p(1),$$

which completes the proof of the theorem.

S2 Proof of Theorem 3.1

To prove Theorem 3.1, we need to show the following lemma first.

Lemma 1. *Define* $\psi_{ps}(\Lambda; \mathbf{X})[h] = \sum_{j=1}^{K} \left\{ \frac{N(T_{K,j})}{\Lambda(T_{K,j})} - 1 \right\} h(T_{K,j})$ *and*

$$\mathcal{G}_n(\delta)[h] \\ = \left\{ \psi_{ps}(\Lambda; \mathbf{X})[h] - \psi_{ps}(\Lambda_0; \mathbf{X})[h] : \begin{array}{c} d_1(\Lambda, \Lambda_0) < \delta, \\ \sup_{\tau_0 \le t \le \tau} |\Lambda(t) - \Lambda_0(t)| < \delta_0, \end{array} \Lambda \in \Psi_n \right\}.$$

Let $\|\cdot\|_{P,B}$ *be the "Bernstein norm" defined as* $\|f\|_{P,B} = \{2P(e^{|f|} - 1 - |f|)\}^{1/2}$ *by van der Vaart and Wellner (1996). Then the* ε*-bracketing number associated with* $\|\cdot\|_{P,B}$ *for* $\mathcal{G}_n(\delta)[h]$*, denoted by* $N_{[]}(\varepsilon, \mathcal{G}_n(\delta)[h], \|\cdot\|_{P,B})$*, is bounded by* $(\delta/\varepsilon)^{cq_n}$*, that is,*

$$N_{[]}(\varepsilon, \mathcal{G}_n(\delta)[h], \|\cdot\|_{P,B}) \lesssim (\delta/\varepsilon)^{cq_n}$$

for a constant c *independent of* h*, where the symbol* $\lesssim$ *denotes that the left-hand side is bounded above by a constant times the right-hand side.*

Proof. For Λ with $\sup_{\tau_0 \le t \le \tau} |\Lambda(t) - \Lambda_0(t)| < \delta_0$, we obtain that $M_1 \le \Lambda(t) \le M_2$ over $t \in [\tau_0, \tau]$ where M_1 and M_2 are positive constants. Denote the ceiling of x by $\lceil x \rceil$. By the calculation in Shen and Wong (1994, page 597), for any $\varepsilon < \delta$, there exists a set of brackets $\left\{ [\Lambda_i^L, \Lambda_i^U] \colon i = 1, \ldots, (\delta/\varepsilon)^{cq_n} \right\}$ such that for any $\Lambda \in \Psi_n$, $\Lambda_i^L(t) \le \Lambda(t) \le \Lambda_i^U(t)$ over $t \in [\tau_0, \tau]$ for some

$1 \le i \le (\delta/\varepsilon)^{c_1 q_n}$, where $\|\Lambda_i^U - \Lambda_i^L\|_\infty \le \varepsilon$, and c is a constant. Define

$$m_i^L(\mathbf{X})[h] = \sum_{j=1}^{K} N(T_{K,j}) \left[\left\{ \frac{I(h(T_{K,j}) \ge 0)}{\max\left(\Lambda_i^U(T_{K,j}), M_2\right)} + \frac{I(h(T_{K,j}) < 0)}{\max\left(\Lambda_i^L(T_{K,j}), M_1\right)} \right\} - \frac{1}{\Lambda_0(T_{K,j})} \right] h(T_{K,j})$$

and

$$m_i^U(\mathbf{X})[h] = \sum_{j=1}^{K} N(T_{K,j}) \left[\left\{ \frac{I(h(T_{K,j}) \ge 0)}{\max\left(\Lambda_i^L(T_{K,j}), M_1\right)} + \frac{I(h(T_{K,j}) < 0)}{\max\left(\Lambda_i^U(T_{K,j}), M_2\right)} \right\} - \frac{1}{\Lambda_0(T_{K,j})} \right] h(T_{K,j}).$$

After some calculations, we have $||m_i^U(\mathbf{X})[h] - m_i^L(\mathbf{X})[h]||_{P,B}^2 \lesssim \varepsilon^2$ and for any $m(\Lambda; \mathbf{X})[h] \in \mathcal{G}_n(\delta)[h]$, there exist some i such that $m(\Lambda, \mathbf{X})[h] \in [m_i^L(\mathbf{X})[h], m_i^U(\mathbf{X})[h]]$. Therefore, we have

$$N_{[]}(\varepsilon, \mathcal{G}_n(\delta)[h], \|\cdot\|_{P,B}) \lesssim (\delta/\varepsilon)^{cq_n}$$

for a universal constant c, which completes the proof of the lemma.

Proof of Theorem 3.1. To derive the asymptotic normality of the estimators, we need to verify conditions A1-A5 stated in Theorem 2.1.

To prove part (i), we define a sequence of maps S_n^{ps} mapping a neighborhood of Λ_0, denoted by $\mathcal{U}$, in the parameter space for Λ into $l^\infty(\mathcal{H}_r)$

as

$$\begin{aligned} S_n^{ps}(\Lambda)[h] &= n^{-1}\frac{d}{d\varepsilon} l_n^{ps}(\Lambda+\varepsilon h)\Big|_{\varepsilon=0} \\ &= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{K_i}\left\{\frac{N_i(T_{K_i,j})}{\Lambda(T_{K_i,j})}-1\right\} h(T_{K_i,j}) \\ &= \mathbb{P}_n\psi_{ps}(\Lambda;\mathbf{X})[h]. \end{aligned}$$

Correspondingly, we define the limit map $S^{ps} : \mathcal{U} \longrightarrow l^{\infty}(\mathcal{H}_r)$ as

$$S^{ps}(\Lambda)[h] = P\left[\sum_{j=1}^{K}\left\{\frac{N(T_{K,j})}{\Lambda(T_{K,j})}-1\right\} h(T_{K,j})\right].$$

We will show (A1) by applying Lemma 1. For $h \in \mathcal{H}_r$, let the class $\mathcal{G}_n(\delta)[h]$ be as defined in Lemma 1 for some $\delta > 0$. Then by Lemma 1, we have

$$N_{[\,]}(\varepsilon, \mathcal{G}_n(\delta)[h], \|\cdot\|_{P,B}) \lesssim (\delta/\varepsilon)^{cq_n}$$

uniformly in h, and

$$J_{[\,]}(\varepsilon, \mathcal{G}_n(\delta)[h], \|\cdot\|_{P,B}) = \int_0^{\delta}\sqrt{1+logN_{[\,]}(\varepsilon, \mathcal{G}_n(\delta)[h], \|\cdot\|_{P,B})}\,d\varepsilon \lesssim q_n^{1/2}\delta.$$

Lu, Zhang and Huang (2007) showed that $d_1(\hat{\Lambda}_n^{ps}, \Lambda_0) \to 0$ almost surely and hence that the uniform consistency of $\hat{\Lambda}^{ps}$ can be shown by using arguments similar to Proposition 5 of Schick and Yu (2000) under conditions C2-C6; that is,

$$\sup_{\tau_0 \le t \le \tau} |\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)| \to 0 \quad \text{almost surely.}$$

By Theorem 2 of Lu, Zhang and Huang (2007), $n^{r/(1+2r)} d_1(\hat{\Lambda}_n^{ps}, \Lambda_0) = O_p(1)$ with $r > 1$. Thus we have $\psi_{ps}(\hat{\Lambda}_n^{ps}; \mathbf{X})[h] - \psi_{ps}(\Lambda_0; \mathbf{X})[h] \in \mathcal{G}_n(\delta)[h]$ with $\delta = \delta_n = O(n^{-r/(1+2r)})$. Furthermore, for any $\psi_{ps}(\Lambda; \mathbf{X})[h] - \psi_{ps}(\Lambda_0; \mathbf{X})[h] \in \mathcal{G}_n(\delta_n)[h]$, we have

$$\sup_{h \in \mathcal{H}_r} ||\psi_{ps}(\Lambda; \mathbf{X})[h] - \psi_{ps}(\Lambda_0; \mathbf{X})[h]||_{P,B}^2 \lesssim d_1^2(\Lambda, \Lambda_0).$$

Hence, using the maximal inequality in Lemma 3.4.3 of van der Vaart and Wellner (1996), we obtain that

$$\begin{aligned} E_P \| n^{1/2}(P_n - P) \|_{\mathcal{G}_n(\delta_n)[h]} &\lesssim J_{[]}(\delta_n, \mathcal{G}_n(\delta_n)[h], \| \cdot \|_{P,B}) \\ &\quad \times \left\{ 1 + c \frac{J_{[]}(\delta_n, \mathcal{G}_n(\delta_n)[h], \| \cdot \|_{P,B})}{\delta_n^2 \sqrt{n}} \right\} \\ &\lesssim q_n^{1/2} \delta_n + q_n n^{-1/2} \\ &= O(n^{1/(2(1+2r)) - r/(1+2r)}) + O(n^{1/(1+2r) - 1/2}) \\ &= O(n^{(1-2r)/(2(1+2r))}) + O(n^{(1-2r)/(2(1+2r))}) \\ &= o(1), \end{aligned}$$

where c is a positive constant. Therefore, employing the Markov inequality, we have

$$\sqrt{n}(P_n - P)(\psi_{ps}(\hat{\Lambda}_n^{ps}; \mathbf{X})[h] - \psi_{ps}(\Lambda_0; \mathbf{X})[h]) = o_p(1)$$

uniformly in h. Thus, (A1) holds.

For (A2), clearly $S^{ps}(\Lambda_0)[h] = 0$ for $h \in \mathcal{H}_r$, and we need to show that

$S_n^{ps}(\hat{\Lambda}_n^{ps})[h] = o(n^{-1/2})$ for $h \in \mathcal{H}_r$. Note that $\hat{\Lambda}_n^{ps} = \sum_{\ell=1}^{q_n} \hat{\alpha}_{\ell n}^{ps} B_\ell$ satisfies the following score equation

$$n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \left\{ \frac{N_i(T_{K_i,j})}{\hat{\Lambda}_n^{ps}(T_{K_i,j})} - 1 \right\} B_\ell(T_{K_i,j}) = 0, \quad \ell = 1, \ldots, q_n.$$

Thus, for any $h = \sum_{\ell=1}^{q_n} \alpha_\ell B_\ell \in \Phi_n$, we have

$$n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \left\{ \frac{N_i(T_{K_i,j})}{\hat{\Lambda}_n^{ps}(T_{K_i,j})} - 1 \right\} h(T_{K_i,j}) = 0,$$

that is, $S_n^{ps}(\hat{\Lambda}_n^{ps})[h] = 0$ for any $h \in \Phi_n$.

For any $h \in \mathcal{H}_r$, there exists $h_n \in \Phi_n$ such that $||h_n - h||_\infty = O(n^{-rv})$. Next we need to show that

$$S_n^{ps}(\hat{\Lambda}_n^{ps})[h - h_n] = o(n^{-1/2}).$$

For this, we write

$$\begin{aligned} S_n^{ps}(\hat{\Lambda}_n^{ps})[h - h_n] &= \{S_n^{ps}(\hat{\Lambda}_n^{ps})[h - h_n] - S_n^{ps}(\Lambda_0)[h - h_n]\} + S_n^{ps}(\Lambda_0)[h - h_n] \\ &\equiv I_{1n} + I_{2n}. \end{aligned}$$

Since

$$\begin{aligned} P|I_{1n}| &= n^{-1} \Bigg| \sum_{i=1}^{n} \sum_{j=1}^{K_i} N_i(T_{K_i,j}) \\ &\quad \times \left\{ \frac{1}{\hat{\Lambda}_n^{ps}(T_{K_i,j})} - \frac{1}{\Lambda_0(T_{K_i,j})} \right\} \{h(T_{K_i,j}) - h_n(T_{K_i,j})\} \Bigg| \\ &\lesssim d_1(\hat{\Lambda}_n^{ps}, \Lambda_0) ||h - h_n||_\infty \end{aligned}$$

and

$$
\begin{aligned}
PI_{2n}^2 &= n^{-1}P\left[\sum_{j=1}^{K}\left\{\frac{N(T_{K,j})}{\Lambda_0(T_{K,j})}-1\right\}\{h(T_{K,j})-h_n(T_{K,j})\}\right]^2 \\
&\lesssim n^{-1}||h-h_n||_\infty^2,
\end{aligned}
$$

then it follows that $I_{1n} = o_p(n^{-1/2})$ and $I_{2n} = o_p(n^{-1/2})$, which implies (A2).

Condition (A3) holds because $\mathcal{H}_r$ is a Donsker class and the functional S_n^{ps} is a bounded Lipschitz function with respect to $\mathcal{H}_r$.

For (A4), by the smoothness of $S^{ps}(\Lambda)$, the Fr$\acute{e}$chet differentiability holds and the derivative of $S^{ps}(\Lambda)$ at Λ_0, denoted by $\dot{S}^{ps}_{\Lambda_0}$, is a map from the space $\{(\Lambda-\Lambda_0) : \Lambda \in \mathcal{U}\}$ to $l^\infty(\mathcal{H}_r)$ and

$$
\begin{aligned}
&\dot{S}^{ps}_{\Lambda_0}(\Lambda-\Lambda_0)[h] \\
&= \frac{d}{d\varepsilon}\left\{S^{ps}(\Lambda_0+\varepsilon(\Lambda-\Lambda_0))[h]\right\}\Big|_{\varepsilon=0} \qquad \text{(S2.1)} \\
&= -P\left[\sum_{j=1}^{K} h(T_{K,j})\left\{\frac{\Lambda(T_{K,j})-\Lambda_0(T_{K,j})}{\Lambda_0(T_{K,j})}\right\}\right].
\end{aligned}
$$

Thus, by condition C8, we have

$$
-\dot{S}^{ps}_{\Lambda_0}(\Lambda-\Lambda_0)[h] = \int(\Lambda(t)-\Lambda_0(t))dQ^{ps}(h)(t) \qquad \text{(S2.2)}
$$

where

$$
Q^{ps}(h)(t) = P\left[\sum_{j=1}^{K} I(T_{K,j}\le t)\frac{h(T_{K,j})}{\Lambda_0(T_{K,j})}\right].
$$

Next we show that (A5) holds. Note that

$$
\begin{aligned}
& S^{ps}(\hat{\Lambda}_n^{ps})[h] - S_{\Lambda_0}^{ps}[h] - \dot{S}_{\Lambda_0}^{ps}(\hat{\Lambda}_n^{ps} - \Lambda_0)[h] \\
&= P\left[\sum_{j=1}^{K}\left\{\frac{N(T_{K,j})}{\hat{\Lambda}_n^{ps}(T_{K,j})} - 1\right\} h(T_{K,j})\right] \\
&\quad + P\left[\sum_{j=1}^{K} h(T_{K,j})\left\{\frac{\hat{\Lambda}_n^{ps}(T_{K,j}) - \Lambda_0(T_{K,j})}{\Lambda_0(T_{K,j})}\right\}\right] \\
&= P\left[\sum_{j=1}^{K}\frac{h(T_{K,j})}{\Lambda_0(T_{K,j})\hat{\Lambda}_n^{ps}(T_{K,j})}\left\{\hat{\Lambda}_n^{ps}(T_{K,j}) - \Lambda_0(T_{K,j})\right\}^2\right] \\
&= O_p(d_1^2(\hat{\Lambda}_n^{ps}, \Lambda_0)).
\end{aligned}
$$

By Theorem 2 of Lu, Zhang and Huang (2007),

$$
d_1^2(\hat{\Lambda}_n^{ps}, \Lambda_0) = O_p(n^{-2r/(1+2r)}) = o_p(n^{-1/2}),
$$

and thus (A5) holds.

It follows from Theorem 2.1 that

$$
\sqrt{n}\int\{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)\}dQ^{ps}(h)(t) = \sqrt{n}(S_n^{ps} - S^{ps})(\Lambda_0)[h] + o_p(1). \quad \text{(S2.3)}
$$

Next, we show that Q^{ps} is one-to-one, that is, for $h \in \mathcal{H}_r$, if $Q^{ps}(h) = 0$, then $h = 0$.

Suppose that $Q^{ps}(h) = 0$. Then $\dot{S}_{\Lambda_0}^{ps}(\Lambda - \Lambda_0)[h] = 0$ for any Λ in the neighborhood $\mathcal{U}$. In particular, we take $\Lambda = \Lambda_0 + \epsilon h$ for a small constant ϵ.

Thus we have

$$\begin{aligned} 0 &= \dot{S}^{ps}_{\Lambda_0}(\Lambda - \Lambda_0)[h] \\ &= -\epsilon P\left[\sum_{j=1}^{K}\Lambda_0(T_{K,j})\left\{\frac{h(T_{K,j})}{\Lambda_0(T_{K,j})}\right\}^2\right], \end{aligned}$$

which yields

$$h(T_{K,j}) = 0,\ j = 1, \ldots, K,\ \ a.s.$$

and so $h = 0$ by condition C10.

For each $h \in \mathcal{H}_r$, since Q^{ps} is invertible, there exists $h^{ps} \in \mathcal{H}_r$ such that $Q^{ps}(h^{ps}) = h$. Therefore, we have

$$\sqrt{n}\int\{\hat{\Lambda}^{ps}_n(t) - \Lambda_0(t)\}dh(t) = \sqrt{n}(S^{ps}_n - S^{ps})(\Lambda_0)[h^{ps}] + o_p(1) \to_d N(0, \sigma^2_{ps}),$$

where

$$\sigma^2_{ps} = E\{\psi^2_{ps}(\Lambda_0; \mathbf{X})[h^{ps}]\}. \qquad \text{(S2.4)}$$

To prove part (ii), we define a sequence of maps S_n mapping a neighborhood of Λ_0, $\mathcal{U}$, in the parameter space for Λ into $l^\infty(\mathcal{H}_r)$ as:

$$S_n(\Lambda)[h] = n^{-1}\frac{d}{d\varepsilon}l_n(\Lambda + \varepsilon h)\Big|_{\varepsilon=0}.$$

Write $\Delta N_i(T_{K_i,j}) = N_i(T_{K_i,j}) - N_i(T_{K_i,j-1})$, $\Delta\Lambda(T_{K_i,j}) = \Lambda(T_{K_i,j}) - \Lambda(T_{K_i,j-1})$, and $\Delta h(T_{K_i,j}) = h(T_{K_i,j}) - h(T_{K_i,j-1})$. Then, we have

$$\begin{aligned}
&S_n(\Lambda)[h] \\
&= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{K_i}\left[\left\{\frac{\Delta N_i(T_{K_i,j})}{\Delta\Lambda(T_{K_i,j})} - 1\right\}\Delta h(T_{K_i,j})\right] \\
&\equiv \ \mathbb{P}_n\psi(\Lambda;\mathbf{X})[h].
\end{aligned}$$

Correspondingly, we define the limit map $S : \mathcal{U} \longrightarrow l^\infty(\mathcal{H}_r)$ as

$$S(\Lambda)[h] = P\left[\sum_{j=1}^{K}\left\{\frac{\Delta N(T_{K,j})}{\Delta\Lambda(T_{K,j})} - 1\right\}\Delta h(T_{K,j})\right].$$

Furthermore, the derivative of $S(\Lambda)$ at Λ_0, denoted by $\dot{S}_{\Lambda_0}$, is a map from the space $\{(\Lambda - \Lambda_0) : \Lambda \in \mathcal{U}\}$ to $l^\infty(\mathcal{H}_r)$ and

$$\begin{aligned}
&\dot{S}_{\Lambda_0}(\Lambda - \Lambda_0)[h] \\
&= -P\sum_{j=1}^{K}\Delta h(T_{K,j})\left\{\frac{\Delta\Lambda(T_{K,j}) - \Delta\Lambda_0(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right\} \qquad \text{(S2.5)} \\
&= -\int\{\Lambda(t) - \Lambda_0(t)\}dQ(h)(t)
\end{aligned}$$

where

$$Q(h)(t) = P\left[\sum_{j=1}^{K}\{I(T_{K,j} \le t) - I(T_{K,j-1} \le t)\}\frac{\Delta h(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right].$$

Similarly, we can show that $\sqrt{n}(S_n - S)(\hat{\Lambda}_n)[h] - \sqrt{n}(S_n - S)(\Lambda_0)[h] = o_p(1)$, $S(\Lambda_0)[h] = 0$, $S_n(\hat{\Lambda}_n)[h] = o_p(n^{-1/2})$, and

$$S(\hat{\Lambda}_n)[h] - S(\Lambda_0)[h] - \dot{S}_{\Lambda_0}(\hat{\Lambda}_n - \Lambda_0)[h] = O_p(d_2^2(\hat{\Lambda}_n, \Lambda_0)).$$

By Theorem 2 of Lu, Zhang and Huang (2007), we have $d_2(\hat{\Lambda}_n, \Lambda_0) = O_p(n^{-r/(1+2r)})$, and so

$$S(\hat{\Lambda}_n)[h] - S(\Lambda_0)[h] - \dot{S}_{\Lambda_0}(\hat{\Lambda}_n - \Lambda_0)[h] = o_p(n^{-1/2}).$$

Thus it follows from Theorem 2.1 that

$$\sqrt{n}\int \{\hat{\Lambda}_n(t) - \Lambda_0(t)\} dQ(h)(t) = \sqrt{n}(S_n - S)(\Lambda_0)[h] + o_p(1). \qquad \text{(S2.6)}$$

Next, we show that Q is one-to-one, that is, for $h \in \mathcal{H}_r$, if $Q(h) = 0$, then $h = 0$

Suppose that $Q(h) = 0$. Then $\dot{S}_{\Lambda_0}(\Lambda - \Lambda_0)[h] = 0$ for any Λ in the neighborhood $\mathcal{U}$. In particular, we take $\Lambda = \Lambda_0 + \epsilon h$ for a small constant ϵ. Thus we have

$$\begin{aligned} 0 &= \dot{S}_{\Lambda_0}(\Lambda - \Lambda_0)[h] \\ &= -\epsilon P\left[\sum_{j=1}^{K} \Delta\Lambda_0(T_{K,j}) \left\{\frac{\Delta h(T_{K,j})}{\Delta\Lambda_0(T_{K,j})}\right\}^2\right], \end{aligned}$$

which yields

$$\Delta h(T_{K,j}) = 0,\ j = 1, \ldots, K,\ \ a.s.$$

Thus,

$$h(T_{K,j}) = 0,\ j = 1, \ldots, K,\ \ a.s.$$

and so $h = 0$ by condition C10.

For each $h \in \mathcal{H}_r$, since Q is invertible, there exists unique $h^* \in \mathcal{H}_r$ such that $Q(h^*) = h$. Thus, we have

$$\sqrt{n} \int \{\hat{\Lambda}_n(t) - \Lambda_0(t)\} dh(t) = \sqrt{n}(S_n - S)(\Lambda_0)[h^*] + o_p(1) \to_d N(0, \sigma^2),$$

where

$$\sigma^2 = E\{\psi^2(\Lambda_0; \mathbf{X})[h^*]\}. \tag{S2.7}$$

Proof of Corollary 3.1. (i) Note that

$$P\left[\sum_{j=1}^{K} h(T_{K,j}) \left\{ \frac{\hat{\Lambda}_n^{ps}(T_{K,j}) - \Lambda_0(T_{K,j})}{\Lambda_0(T_{K,j})} \right\}\right] = \int h(t) \frac{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)}{\Lambda_0(t)} d\mu_1(t).$$

Thus it follows from (S2.1)-(S2.3) that

$$\sqrt{n} \int h(t) \frac{\hat{\Lambda}_n^{ps}(t) - \Lambda_0(t)}{\Lambda_0(t)} d\mu_1(t) = \sqrt{n}(S_n^{ps} - S^{ps})(\Lambda_0)[h] + o_p(1),$$

which completes the proof of (i).

Similarly, the result in part (ii) follows from (S2.5) and (S2.6).

S3 Proof of Theorem 4.1

(i) Note that

$$\begin{aligned} U_n^{(ps)} &= \sqrt{n}\mathbb{P}_n\left\{\sum_{j=1}^{K} h_n(T_{K,j})\frac{\hat{\Lambda}_1^{ps}(T_{K,j})-\hat{\Lambda}_2^{ps}(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})}\right\} \\ &= \sqrt{n}\mathbb{P}_n\left\{\sum_{j=1}^{K} h_n(T_{K,j})\frac{\hat{\Lambda}_1^{ps}(T_{K,j})-\Lambda_0(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})}\right\} \\ &\quad -\sqrt{n}\mathbb{P}_n\left\{\sum_{j=1}^{K} h_n(T_{K,j})\frac{\hat{\Lambda}_2^{ps}(T_{K,j})-\Lambda_0(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})}\right\}, \end{aligned}$$

and

$$\begin{aligned} &\sqrt{n}\mathbb{P}_n\left\{\sum_{j=1}^{K} h_n(T_{K,j})\frac{\hat{\Lambda}_1^{ps}(T_{K,j})-\Lambda_0(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})}\right\} \\ &= U_{1n}^{ps}+U_{2n}^{ps}+U_{3n}^{ps}+U_{4n}^{ps} \end{aligned}$$

where

$$U_{1n}^{ps} = \sqrt{n}(\mathbb{P}_n - P)\left\{\sum_{j=1}^{K} h_n(T_{K,j})\frac{\hat{\Lambda}_1^{ps}(T_{K,j})-\Lambda_0(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})}\right\},$$

$$U_{2n}^{ps} = \sqrt{n}P\left[\sum_{j=1}^{K}\{h_n(T_{K,j})-h(T_{K,j})\}\frac{\hat{\Lambda}_1^{ps}(T_{K,j})-\Lambda_0(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})}\right],$$

$$U_{3n}^{ps} = \sqrt{n}P\left[\sum_{j=1}^{K} h(T_{K,j})\{\hat{\Lambda}_1^{ps}(T_{K,j})-\Lambda_0(T_{K,j})\}\left\{\frac{1}{\hat{\Lambda}_0^{ps}(T_{K,j})}-\frac{1}{\Lambda_0(T_{K,j})}\right\}\right],$$

$$U_{4n}^{ps} = \sqrt{n}P\left[\sum_{j=1}^{K} h(T_{K,j})\frac{\hat{\Lambda}_1^{ps}(T_{K,j})-\Lambda_0(T_{K,j})}{\Lambda_0(T_{K,j})}\right].$$

By the arguments similar to those used in the proof of Theorem 3.1 of Balakrishnan and Zhao (2009), we can show that $U_{1n}^{ps} = o_p(1)$, $U_{2n}^{ps} = o_p(1)$, and $U_{3n}^{ps} = o_p(1)$.

From (S2.1)-(S2.3), we have

$$U_{4n}^{ps} = \sqrt{n}(\mathbb{P}_{n_1} - P)\psi_{ps}(\Lambda_0; \mathbf{X})[h] + o_p(1),$$

where $\mathbb{P}_{n_1}$ is the empirical measure based on group 1. Similarly, we have

$$\begin{aligned}
&\sqrt{n}\mathbb{P}_n \left\{ \sum_{j=1}^{K} h_n(T_{K,j}) \frac{\hat{\Lambda}_2^{ps}(T_{K,j}) - \Lambda_0(T_{K,j})}{\hat{\Lambda}_0^{ps}(T_{K,j})} \right\} \\
&= \sqrt{n}(\mathbb{P}_{n_2} - P)\psi_{ps}(\Lambda_0; \mathbf{X})[h] + o_p(1),
\end{aligned}$$

where $\mathbb{P}_{n_2}$ is the empirical measure based on group 2. Hence, we have

$$\begin{aligned}
U_n^{ps} &= \sqrt{\frac{n}{n_1}}\sqrt{n_1}(\mathbb{P}_{n_1} - P)\psi_{ps}(\Lambda_0; \mathbf{X})[h] \\
&\quad - \sqrt{\frac{n}{n_2}}\sqrt{n_2}(\mathbb{P}_{n_2} - P)\psi_{ps}(\Lambda_0; \mathbf{X})[h] + o_p(1).
\end{aligned}$$

Here $\mathbb{P}_{n_1}$ and $\mathbb{P}_{n_2}$ are independent. Thus it follows that U_n^{ps} converges in distribution to $N(0, \sigma_{ps}^2)$.

(ii) Using the arguments similar to the proof of (i), we can obtain

$$\begin{aligned}
U_n &= \sqrt{\frac{n}{n_1}}\sqrt{n_1}(\mathbb{P}_{n_1} - P)\psi(\Lambda_0; \mathbf{X})[h] \\
&\quad - \sqrt{\frac{n}{n_2}}\sqrt{n_2}(\mathbb{P}_{n_2} - P)\psi(\Lambda_0; \mathbf{X})[h] + o_p(1),
\end{aligned}$$

which yields the asymptotic normal distribution $N(0, \sigma^2)$.

(iii) The proof of this part is omitted since it is similar to those used in the proof of Theorem 3.1 (iii) of Balakrishnan and Zhao (2009).

References

Balakrishnan, N. and Zhao, X. (2009). New multi-sample nonparametric tests for panel count data. *Ann. Statist.* **37**, 1112–1149.

Lu, M., Zhang, Y. and Huang, J. (2007). Estimation of the mean function with panel count data using monotone polynomial splines. *Biometrika* **94**, 705–718.

Schick, A. and Yu, Q. (2000). Consistency of the GMLE with mixed case interval-censored data. *Scand. J. Statist.* **27**, 45–55.

Shen, X. and Wong, W. H. (1994). Convergence rate of sieve estimates. *Ann. Statist.* **18**, 580–615.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes.* New York: Springer-Verlag.