NONPARAMETRIC ESTIMATION AND TESTING
FOR PANEL COUNT DATA
WITH INFORMATIVE TERMINAL EVENT

Xiangbin Hu\(^1\), Li Liu\(^2\), Ying Zhang\(^3\), and Xingqiu Zhao\(^1\)

\(^1\) Hong Kong Polytechnic University, \(^2\) Wuhan University
and \(^3\) University of Nebraska Medical Center

Supplementary Material

This Supplementary Material contains calculations of the loss function, some preliminary lemmas, proofs of Theorems 1–5, and some additional works on simulation studies and the real data analysis.
S1 Calculation of Loss Function

For the first part of $\ell_n(\Lambda, \hat{F}_n; X)$, replacing $\Lambda(s)$ by $I^T(s)\alpha$, we have

$$
\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \Delta_i \left\{ \Delta N_{i,j} - \Delta I_j(Y_i)^T \alpha \right\}^2
= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \Delta_i \left\{ \alpha^T \Delta I_j(Y_i) \Delta I_j(Y_i)^T \alpha - 2\Delta N_{i,j} \Delta I_j(Y_i)^T \alpha + \Delta N_{i,j}^2 \right\}
= \alpha^T \left\{ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \Delta_i \Delta I_j(Y_i) \Delta I_j(Y_i)^T \right\} \alpha
- \left\{ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \Delta_i \Delta N_{i,j} \Delta I_j(Y_i) \right\}^T \alpha
+ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \Delta_i \Delta N_{i,j}^2
= : \alpha^T A_1 \alpha - 2B_1^T \alpha + C_1.
$$

For the second part, the KM estimator is a nondecreasing step function

$$
\hat{F}_n(u) = \sum_{l=1}^{L} f_l \mathbb{1}_{[t_l, t_{l+1})}(u),
$$

where $t_1 = 0$, $t_{L+1} = \tau$, and $\{[t_l, t_{l+1}) : l = 1, \cdots, L \}$ is a partition of $[0, \tau)$.

Furthermore, for each subject, we assume $Y_i$ to be in the interval $[t_{l_i}, t_{l_{i+1}})$,
and take \( f_{L+1} = f_L \). Then we have

\[
\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} (1 - \Delta_i) \int_{Y_i}^{\infty} \left\{ \Delta N_{i,j} - \Delta I_{j}(u)^T \alpha \right\}^2 d\hat{F}_n(u) \\
= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_i} \frac{(f_{i+1} - f_i) \{ \Delta N_{i,j} - \Delta I_{j}(t_{i+1}^{(1)})^T \alpha \}^2}{1 - \hat{F}_n(Y_i)} \\
= \sum_{i=1}^{n} \frac{1 - \Delta_i}{n(1 - \hat{F}_n(Y_i))} \sum_{j=1}^{L} \sum_{l=1}^{K_i} \left\{ (f_{i+1} - f_i) \{ \alpha^T \Delta I_j(t_{i+1}^{(1)}) \Delta I_j(t_{i+1}^{(1)})^T \alpha \\
- 2 \Delta N_{i,j} \Delta I_j(t_{i+1}^{(1)})^T \alpha + \Delta N_{i,j}^2 \} \right\} \\
= \alpha^T \left[ \sum_{i=1}^{n} \frac{1 - \Delta_i}{n(1 - \hat{F}_n(Y_i))} \sum_{j=1}^{K_i} \sum_{l=1}^{L} \left\{ (f_{i+1} - f_i) \Delta I_j(t_{i+1}^{(1)}) \Delta I_j(t_{i+1}^{(1)})^T \right\} \right] \alpha \\
- 2 \left[ \sum_{i=1}^{n} \frac{1 - \Delta_i}{n(1 - \hat{F}_n(Y_i))} \sum_{j=1}^{K_i} \sum_{l=1}^{L} \left\{ \Delta N_{i,j} \sum_{l=1}^{L} (f_{i+1} - f_i) \Delta I_j(t_{i+1}^{(1)})^T \right\} \right] ^T \alpha \\
+ \sum_{i=1}^{n} \frac{(1 - \Delta_i)(f_{i+1} - f_i)}{n(1 - \hat{F}_n(Y_i))} \sum_{j=1}^{K_i} \Delta N_{i,j}^2 \\
= : \alpha^T A_2 \alpha - 2B_2^T \alpha + C_2.
\]

Thus, to obtain the estimator \( \hat{\Lambda}_n \), we should minimize \( \alpha^T A \alpha - 2B^T \alpha + C \) under the constraints that \( \alpha_l \geq 0 \) for \( l = 1, \cdots, q_n \), where \( A = A_1 + A_2 \),

\( B = B_1 + B_2 \) and \( C = C_1 + C_2 \).
S2 Lemmas

Lemma 1. (i) Suppose that Condition (C2) holds. For sufficiently small δ, any \( F \in \mathcal{F}_\delta \) and any differentiable function \( g \), we have

\[
\mathcal{P} \left[ (1 - \Delta) \sum_{j=1}^{K} \left| \int_{Y} g(u - T_j) dF(u) - \int_{Y} g(u - T_j) dF_0(u) \right| \right]
\leq \left( E \left[ \sum_{j=1}^{K} |g'(U - T_j)| \right] + E \left[ \sum_{j=1}^{K} |g(U - T_j)| \right] \right) ||F - F_0||_\infty.
\]

(S2.1)

(ii) In addition, suppose that Conditions (C1) and (C3) hold. It follows that for all \( \Lambda \in \Phi \),

\[
|\mathcal{P} m(\Lambda, F; X) - \mathcal{P} m(\Lambda, F_0; X) | \lesssim d_2(F, F_0).
\]

Thus, we have \( \mathcal{P} m(\Lambda, \hat{F}_n; X) - \mathcal{P} m(\Lambda, F_0; X) = o_p(1) \) for all \( \Lambda \in \Phi \).

Proof. (i) By direct calculations, we have

\[
\mathcal{P} \left[ (1 - \Delta) \sum_{j=1}^{K} \left| \int_{Y} g(u - T_j) dF(u) - \int_{Y} g(u - T_j) dF_0(u) \right| \right] \\
\leq \mathcal{P} \left[ \frac{1 - \Delta}{(1 - F_0(Y))(1 - F(Y))} \sum_{j=1}^{K} \left| (1 - F_0(Y)) \int_{Y} g(u - T_j) d(F_0(u) - F(u)) \\ + (F_0(Y) - F(Y)) \int_{Y} g(u - T_j) dF_0(u) \right| \right] \\
\leq \mathcal{P} \left[ (1 - \Delta) \sum_{j=1}^{K} \left| \int_{Y} g(u - T_j) d(F_0(u) - F(u)) \right| \right] \\
+ \mathcal{P} \left[ \frac{1 - \Delta}{(1 - F_0(Y))(1 - F(Y))} \sum_{j=1}^{K} \left| \int_{Y} g(u - T_j) dF_0(u) \right| \right] ||F - F_0||_\infty.
\]
Using integration by parts, we have

\[
P \left[ (1 - \Delta) \sum_{j=1}^{K} \left| \int_{Y}^{\infty} g(u - T_j) d(F_0(u) - F(u)) \right| \right]
= P \left[ (1 - \Delta) \sum_{j=1}^{K} \left| -\left( F_0(Y) - F(Y) \right) g(Y - T_j) - \int_{Y}^{\infty} g'(u - T_j)(F_0(u) - F(u)) du \right| \right]
\leq P \left[ (1 - \Delta) \sum_{j=1}^{K} \left| g(Y - T_j) \right| \right] ||F - F_0||_{\infty} + P \left[ (1 - \Delta) \sum_{j=1}^{K} \int_{Y}^{\infty} \left| g'(u - T_j) \right| du \right] ||F - F_0||_{\infty}.
\]

Condition (C2) implies that \( 1 - F_0(Y) \) is larger than a positive constant. Hence, for sufficiently small \( \delta \) and any \( F \in \mathcal{F}_\delta \), \( 1 - F(Y) \) is also larger than a positive constant. It follows that

\[
P \left[ (1 - \Delta) \sum_{j=1}^{K} \left| g(Y - T_j) \right| \right] \leq P \left[ \Delta \sum_{j=1}^{K} \left| g(Y - T_j) \right| \right] \lessapprox P \left[ \sum_{j=1}^{K} \left| g(U - T_j) \right| \right]
\]
and

\[
P \left[ (1 - \Delta) \sum_{j=1}^{K} \int_{Y}^{\infty} \frac{g(u - T_j) dF_0(u)}{1 - F(Y)} \right] \lessapprox P \left[ (1 - \Delta) \sum_{j=1}^{K} \int_{Y}^{\infty} \frac{g'(u - T_j) dF_0(u)}{1 - F_0(Y)} \right] \leq P \left[ \sum_{j=1}^{K} \left| g(U - T_j) \right| \right].
\]

Moreover, by the second part of Condition (C2), we obtain

\[
P \left[ (1 - \Delta) \sum_{j=1}^{K} \int_{Y}^{\infty} \frac{g'(u - T_j) dF_0(u)}{1 - F(Y)} \right] \lessapprox P \left[ \frac{1 - \Delta}{1 - F_0(Y)} \sum_{j=1}^{K} \int_{Y}^{\infty} \frac{g'(u - T_j) dF_0(u)}{f_0(u)} \right]
\lessapprox P \left[ \sum_{j=1}^{K} \left| g'(U - T_j) \right| \right].
\]

Thus, \( \text{(S2.1)} \) holds.
(ii) By the first part of the Lemma, Conditions (C1) and (C3), we have

\[
\left| \mathcal{P}[m(\Lambda, F; X) - m(\Lambda, F_0; X)] \right|
\leq \mathcal{P} \left[ \sum_{j=1}^{K} (1 - \Delta) \left| \int_{1}^\infty \frac{(\Delta N_j - \Delta \Lambda_j(u))^2 dF(u)}{1 - F(Y)} - \int_{1}^\infty \frac{(\Delta N_j - \Delta \Lambda_j(u))^2 dF_0(u)}{1 - F_0(Y)} \right| \right]
\lesssim \left( E \left[ \sum_{j=1}^{K} \left\{ |\Lambda_j'(U)| (\Delta N_j - \Delta \Lambda_j(U)) | + (\Delta N_j - \Delta \Lambda_j(U))^2 \right\} \right] \right) ||F - F_0||_\infty
\lesssim d_2(F, F_0).
\]

\[\square\]

**Lemma 2.** Suppose that Conditions (C2), (C4) and (C6) hold. Then for sufficiently small $\delta$, \{m(\Lambda, F; X) : \Lambda \in \Phi, F \in \mathcal{F}_\delta, \Lambda \text{ is uniformly bounded}\} is Donsker, where

\[
m(\Lambda, F; X) = \sum_{j=1}^{K} \Delta \{\Delta N_j - \Delta \Lambda_j(Y)\}^2 + \sum_{j=1}^{K} (1 - \Delta) \int_{1}^\infty \frac{\{\Delta N_j - \Delta \Lambda_j(u)\}^2 dF(u)}{1 - F(Y)}.
\]

**Proof.** Note that functions in $\mathcal{F}_\delta$ and $\Phi$ are monotone and uniformly bounded.

It follows from Section 3 of [van der Vaart (1996)] that \{\Lambda \in \Phi : \Lambda \text{ is uniformly bounded}\} and $\mathcal{F}_\delta$ are Donsker. Since $\{\Delta N_j - \Delta \Lambda_j(Y)\}^2$ is Lipschitz for $\Lambda$, by Theorem 2.10.6 of [van der Vaart and Wellner (1996)],

\{[[\Delta N_j - \Delta \Lambda_j(Y)]^2 : \Lambda \in \Phi \text{ is uniformly bounded}\}

is Donsker. Note that

\[
\left\{ \int_{1}^\infty \left[ \Delta N_j - \Delta \Lambda_j(u) \right]^2 dF(u) : \Lambda \in \Phi \text{ is uniformly bounded} \right\}
\]
is a subset of the convex combinations of functions in
\[ \{ [\triangle N_j - \triangle \Lambda_j(Y)]^2 : \Lambda \in \Phi \text{ is uniformly bounded} \}. \]

By Theorem 2.10.1 and Theorem 2.10.3 of \cite{vanDerVaartWellner1996},
\[ \left\{ \int_Y [\triangle N_j - \triangle \Lambda_j(u)]^2 dF(u) : \Lambda \in \Phi \text{ is uniformly bounded} \right\} \]
is Donsker. Since \( 1 - F(Y) \) and \( K \) are bounded from Conditions (C2), (C4) and (C6),
\[ \{ m(\Lambda, F; X) : \Lambda \in \Phi, F \in F_\delta, \text{ \Lambda \ is uniformly bounded} \} \]
is Donsker by Theorem 2.10.6 of \cite{vanDerVaartWellner1996}. \qed

**Lemma 3** (Rate of Convergence of M-estimator with Nuisance Parameter).

Suppose that for every \( \Lambda \in \Phi_n \), sufficiently large \( n \) and sufficiently small \( \eta \),
\[ \mathcal{P}(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \lesssim -d_1(\Lambda, \Lambda_0)^2 + d_1(\Lambda, \Lambda_0) d_2(\hat{F}_n, F_0) + d_2(\hat{F}_n, F_0)^2 \]
and
\[ E \sup_{\{\Lambda \in \Phi_n : d_1(\Lambda, \Lambda_0) < \eta\}} |(\mathbb{P}_n - \mathcal{P})(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X))| \lesssim \frac{\phi_n(\eta)}{\sqrt{n}} \]
hold, where \( \phi_n(\eta) \) satisfies that \( \eta \mapsto \phi_n(\eta)/\eta^\alpha \) is decreasing for some \( \alpha < 2 \).

Let \( r_n > 0 \) satisfy \( \phi_n(r_n) \lesssim \sqrt{nr_n^2} \). If the sequence \( \hat{\Lambda}_n \) satisfies
\[ \mathbb{P}_n m(\Lambda_0, \hat{F}_n; X) \geq \mathbb{P}_n m(\hat{\Lambda}_n, \hat{F}_n; X) - O_p(r_n^2) \]
and converges in probability to \( \Lambda_0 \), then \( d_1(\hat{\Lambda}_n, \Lambda_0) = O_p(r_n + d_2(\hat{F}_n, F_0)) \).
Proof. This Lemma is similar to Theorem 5.55 of van der Vaart (1998). In order to verify 
\[ d_1(\hat{\Lambda}_n, \Lambda_0) = O_p(r_n + d_2(\hat{F}_n, F_0)), \]
we need to prove that for sufficiently large \( n \),
\[
\lim_{M \to \infty} P(\hat{\Lambda}_n \in \Phi_n : d_1(\hat{\Lambda}_n, \Lambda_0) \geq 2^M (d_2(\hat{F}_n, F_0) + r_n)) = 0.
\]
Then we divide \( \Phi_n \) into shells
\[
S_{n,j,M} = \{ \Lambda \in \Phi_n : 2^j r_n \leq d_1(\Lambda, \Lambda_0) < 2^{j+1} r_n, 2^M d_2(\hat{F}_n, F_0) \leq d_1(\Lambda, \Lambda_0) \}.
\]
For any \( \Lambda \in S_{n,j,M} \), we have
\[
2 d_1(\hat{\Lambda}_n, \Lambda_0) \geq 2^M (d_2(\hat{F}_n, F_0) + r_n).
\]
Hence,
\[
\left\{ \hat{\Lambda}_n \in \Phi_n : 2 d_1(\hat{\Lambda}_n, \Lambda_0) \geq 2^M (d_2(\hat{F}_n, F_0) + r_n) \right\} \subseteq \bigcup_{j \geq M} \{ \hat{\Lambda}_n \in S_{n,j,M} \}.
\]
It follows that
\[
P \left( \hat{\Lambda}_n \in \Phi_n : d_1(\hat{\Lambda}_n, \Lambda_0) \geq 2^{M-1} (d_2(\hat{F}_n, F_0) + r_n) \right) \leq P \left( \hat{\Lambda}_n \in \bigcup_{j \geq M} S_{n,j,M} \right).
\]
(S2.2)

Furthermore, since \( \hat{\Lambda}_n \) satisfies that
\[
P_n(m(\Lambda_0, \hat{F}_n; X) \geq P_n(m(\hat{\Lambda}_n, \hat{F}_n; X) - O_p(r_n^2)), \]
for \( \hat{\Lambda}_n \in S_{n,j,M} \) we can find a variable \( R_n = O_p(r_n^2) \) such that
\[
\sup_{\Lambda \in S_{n,j,M}} P_n(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \geq P_n(m(\Lambda_0, \hat{F}_n; X) - m(\hat{\Lambda}_n, \hat{F}_n; X)) \geq -R_n.
\]
Then for any constant \( \kappa \), we have
\[
P \left( \hat{\Lambda}_n \in \bigcup_{j \geq M} S_{n,j,M} \right)
\leq P \left( \sup_{\Lambda \in \bigcup_{j \geq M} S_{n,j,M}} P_n(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \geq -\kappa r_n^2 \right) + P(R_n \geq \kappa r_n^2).
\]
(S2.3)
By \([S2.2]\) and \([S2.3]\), we obtain

\[
P \left( \hat{\Lambda}_n \in \Phi_n : d_1(\hat{\Lambda}_n, \Lambda_0) \geq 2M (d_2(\hat{F}_n, F_0) + r_n) \right)
\]

\[
\leq P(d_1(\hat{\Lambda}_n, \Lambda_0) \geq \eta) + P(R_n \geq \kappa r_n^2)
\]

\[+ \sum_{j \geq M+1, 2j+1 \leq \eta/r_n} P \left( \sup_{\Lambda \in S_{n,j,M}} \mathbb{P}_n(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \geq -\kappa r_n^2 \right). \]

Since \(\hat{\Lambda}_n\) is consistent and \(R_n = O_p(r_n^2)\), \(P(d_1(\hat{\Lambda}_n, \Lambda_0) \geq \eta)\) and \(P(R_n \geq \kappa r_n^2)\) can be arbitrarily small for sufficiently large \(n\) by the choice of \(\eta\) and \(\kappa\). Thus, we need to prove the limitation of the summation on the right hand side is 0 as \(M\) goes to infinity.

Note that for any positive integer \(M\), \(1/4 \leq 1 - 2^{-M} - 2^{-2M} < 1\). Then for all \(\Lambda \in S_{n,j,M}\),

\[
\mathcal{P}(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X))
\]

\[
\lesssim -d_1(\Lambda, \Lambda_0)^2 + d_1(\Lambda, \Lambda_0)d_2(\hat{F}_n, F_0) + d_2(\hat{F}_n, F_0)^2
\]

\[
\leq -(1 - 2^{-M} - 2^{-2M})d_1(\Lambda, \Lambda_0)^2 \leq -2^{2j}r_n^2.
\]

Hence, \(\mathcal{P}(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \leq -c_1 2^{2j}r_n^2\) for some constant \(c_1\).

Taking \(M\) with \(M \geq \frac{1}{2} \log_2(2\kappa/c_1)\), then by the Markov’s inequality, for
$j \geq M + 1$ and sufficiently large $n$ with $r_n \leq 2^{-(j+1)} \eta$, we have

\[
P\left( \sup_{\Lambda \in S_{n,j,M}} \mathbb{P}_n(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \geq -\kappa r_n^2 \right)
\leq P\left( \sup_{\Lambda \in S_{n,j,M}} (\mathbb{P}_n - \mathcal{P})(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \geq \frac{c_1}{2} 2^{2j} r_n^2 \right)
\leq \frac{2}{c_1 2^{2j} r_n^2} E \sup_{\Lambda \in S_{n,j,M}} \left| (\mathbb{P}_n - \mathcal{P})(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \right| \lesssim \frac{\phi_n(2^{(j+1)} r_n)}{2^{2j} r_n^2 \sqrt{n}}.
\]

Since $\phi_n(\eta)/\eta^\alpha$ is decreasing for some $\alpha < 2$, we have $\phi_n(c \eta) \leq c^\alpha \phi_n(\eta)$ for any $c > 1$. Then $\phi_n(r_n) \lesssim \sqrt{n} r_n^2$ ensures that

\[
\frac{\phi_n(2^{(j+1)} r_n)}{2^{2j} r_n^2 \sqrt{n}} \lesssim \frac{2^{\alpha(j+1)} \sqrt{n} r_n^2}{2^{2j} r_n^2 \sqrt{n}} = \frac{1}{2^{(2-\alpha)j-\alpha}}.
\]

Thus,

\[
\sum_{j \geq M+1, 2^{j+1} \leq \eta/r_n} P\left( \sup_{\Lambda \in S_{n,j,M}} \mathbb{P}_n(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \geq 0 \right) \lesssim \sum_{j \geq M+1} \frac{1}{2^{(2-\alpha)j-\alpha}}.
\]

Noting that $\sum_{j \geq M+1} \frac{1}{2^{(2-\alpha)j-\alpha}}$ tends to 0 as $M$ approaches to infinity, this lemma is concluded. \hfill \square

**Lemma 4.** Suppose that Conditions (C1), (C2), (C4)–(C6), (C8), (C9) and (C11) hold. Define the class

\[
\mathcal{M}_\eta(F) = \{ m(\Lambda, F; X) - m(\Lambda_0, F; X) : \Lambda \in \Phi_n, d_1(\Lambda, \Lambda_0) \leq \eta \}
\]

for each $F \in \mathcal{F}_\delta$. For any $\varepsilon < \eta$ and sufficiently small $\delta$, we have

\[
\log N(\varepsilon, \mathcal{M}_\eta(F), || \cdot ||_{P,B}) \lesssim q_n \log(\eta/\varepsilon),
\]

where the Bernstein norm is defined as $||f||_{P,B} = \{2P(e^{||f||} - 1 - ||f||)\}^{1/2}$.
Proof. By the calculation in Shen and Wong (1994, Page 597), under Condition (C5), for any \( \varepsilon < \eta \), there is a set of brackets

\[
\left\{ [\Lambda^L_i, \Lambda^U_i] : ||\Lambda_i^U(s_1, s_2) - \Lambda_i^L(s_1, s_2)||_{L_2(\mu)} \leq \varepsilon, i = 1, \cdots, (\eta/\varepsilon)^{c_0q_n} \right\}
\]

such that for all \( \Lambda \in \{ \Lambda \in \Phi_n : d_1(\Lambda, \Lambda_0) \leq \eta \} \), we can find an interval \([\Lambda^L_i, \Lambda^U_i] \) satisfying \( \Lambda(s_1, s_2) \in [\Lambda_i^L(s_1, s_2), \Lambda_i^U(s_1, s_2)] \). This implies that

\[
||\Lambda(s_1, s_2) - \Lambda_0(s_1, s_2)||_{L_2(\mu)} \leq \varepsilon,
\]

\[
||\Lambda(s_1, s_2) - \Lambda^L_i(s_1, s_2)||_{L_2(\mu)} \leq \varepsilon.
\]

Noting that \( ||\Lambda(s_1, s_2) - \Lambda_0(s_1, s_2)||_{L_2(\mu)} \leq \eta \), we have

\[
||\Lambda_i^L(s_1, s_2) - \Lambda_0(s_1, s_2)||_{L_2(\mu)} \leq (\varepsilon^2 + \eta^2)^{1/2},
\]

\[
||\Lambda_i^U(s_1, s_2) - \Lambda_0(s_1, s_2)||_{L_2(\mu)} \leq (\varepsilon^2 + \eta^2)^{1/2}.
\]

By Lemma 7.1 of Wellner and Zhang (2007), under Conditions (C8) and (C9), for any \( \Lambda(s_1, s_2) \) satisfying \( ||\Lambda(s_1, s_2) - \Lambda_0(s_1, s_2)||_{L_2(\mu)} \leq \varepsilon^* \), we have \( ||\Lambda(s_1, s_2) - \Lambda_0(s_1, s_2)||_{\infty} \leq (\varepsilon^*/c_1)^{2/3} \) for some constant \( c_1 \).

Hence,

\[
0 \lor ((\varepsilon^2 + \eta^2)^{1/2}/c_1)^{2/3} \leq \Lambda_i^L(s_1, s_2)
\]

\[
\leq \Lambda_i^U(s_1, s_2) \leq \Lambda_0(s_1, s_2) + ((\varepsilon^2 + \eta^2)^{1/2}/c_1)^{2/3},
\]

which implies that \( \Lambda_i^U(s_1, s_2) - \Lambda_i^L(s_1, s_2) \) are uniformly bounded by

\[
2((\varepsilon^2 + \eta^2)^{1/2}/c_1)^{2/3}.
\]
We turn to consider the $\varepsilon$-bracket of $m(\Lambda, F; X) - m(\Lambda_0, F; X)$. Note that

\[ m(\Lambda, F; X) - m(\Lambda_0, F; X) = K \sum_{j=1}^{K} \Delta \left[ (\triangle N_j - \triangle \Lambda_j(Y))^2 - (\triangle N_j - \triangle \Lambda_{0,j}(Y))^2 \right] + \sum_{j=1}^{K} (1 - \Delta) \left[ \int_{Y}^{\infty} \frac{(\triangle N_j - \triangle \Lambda_j(u))^2 - (\triangle N_j - \triangle \Lambda_{0,j}(u))^2}{1 - F(Y)} dF(u) \right] \]

By Conditions (C2) and (C4), $1 - F(Y)$ is positive and has uniform upper and lower bounds. Hence,

\[ m_i^L(\Lambda_i^L, \Lambda_i^U, F; X) \leq m(\Lambda, F; X) - m(\Lambda_0, F; X) \leq m_i^U(\Lambda_i^L, \Lambda_i^U, F; X) \]

where

\[ m_i^L(\Lambda_i^L, \Lambda_i^U, F; X) = K \sum_{j=1}^{K} \Delta \left[ \{\triangle \Lambda_{i,j}^L(Y)\}^2 - 2\triangle N_j \triangle \Lambda_{i,j}^U(Y) - \{\triangle \Lambda_{0,j}(Y)\}^2 + 2\triangle N_j \triangle \Lambda_{0,j}(Y) \right] + \sum_{j=1}^{K} \frac{1 - \Delta}{1 - F(Y)} \int_{Y}^{\infty} \left[ \{\triangle \Lambda_{i,j}^L(u)\}^2 - 2\triangle N_j \triangle \Lambda_{i,j}^U(u) - \{\triangle \Lambda_{0,j}(u)\}^2 + 2\triangle N_j \triangle \Lambda_{0,j}(u) \right] dF(u) \]
and

\[ m_i^U(\Lambda^L_i, \Lambda^U_i, F; X) \]

\[ = \sum_{j=1}^{K} \Delta \left[ \{ \triangle \Lambda_{i,j}^U (Y) \}^2 - 2 \triangle N_j \triangle \Lambda_{i,j}^L (Y) - \{ \triangle \Lambda_{0,j} (Y) \}^2 + 2 \triangle N_j \triangle \Lambda_{0,j} (Y) \right] \]

\[ + \sum_{j=1}^{K} \frac{1 - \Delta}{1 - F(Y)} \int_Y^\infty \left[ \{ \triangle \Lambda_{i,j}^U (u) \}^2 - 2 \triangle N_j \triangle \Lambda_{i,j}^L (u) - \{ \triangle \Lambda_{0,j} (u) \}^2 + 2 \triangle N_j \triangle \Lambda_{0,j} (u) \right] dF(u) \]

for \( i = 1, \ldots, (\eta/\epsilon)^{\alpha_0} \).

We also need to verify

\[ ||m_i^L(\Lambda^L_i, \Lambda^U_i, F; X) - m_i^U(\Lambda^L_i, \Lambda^U_i, F; X)||^2 \leq \varepsilon^2. \]

According to

\[ ||f||^2_{P,B} = 2P(e^{\sqrt{f}} - 1 - |f|) = 2P \left( \sum_{n=2}^{\infty} \frac{|f|^n}{n!} \right) \leq 2P \left( |f|^2 e^{\sqrt{|f|}} \right), \]

\[ ||f||^2_{P,B} \leq \varepsilon^2 \] is followed from \( P(|f|^2 e^{\sqrt{|f|}}) \leq \varepsilon^2 \). Note that

\[ \{ \triangle \Lambda_{i,j}^U (u) \}^2 - \{ \triangle \Lambda_{i,j}^L (u) \}^2 + 2 \triangle N_j (\triangle \Lambda_{i,j}^U (u) - \triangle \Lambda_{i,j}^L (u)) \]

\[ = (\triangle \Lambda_{i,j}^U (u) + \triangle \Lambda_{i,j}^L (u) + 2 \triangle N_j) (\triangle \Lambda_{i,j}^U (u) - \triangle \Lambda_{i,j}^L (u)) \]

\[ \leq \triangle N_j (\triangle \Lambda_{i,j}^U (u) - \triangle \Lambda_{i,j}^L (u)) + (\triangle \Lambda_{i,j}^U (u) - \triangle \Lambda_{i,j}^L (u)). \]

Since \( \Lambda^L_i, \Lambda^U_i \) and \( 1 - F(Y) \) are uniformly bounded, we have

\[ e^{\sqrt{|m_i^U(\Lambda^L_i, \Lambda^U_i, F; X) - m_i^L(\Lambda^L_i, \Lambda^U_i, F; X)|}} \leq e^{cN(T_K)} \]

with some constant \( c \). By Cauchy-Schwarz inequality and Condition (C11),
we obtain

\[
P \left( |m_i^U(\Lambda_i^L, \Lambda_i^U, F; X) - m_i^L(\Lambda_i^L, \Lambda_i^U, F; X)|^2 e^{\|m_i^U(\Lambda_i^L, \Lambda_i^U, F; X) - m_i^L(\Lambda_i^L, \Lambda_i^U, F; X)|} \right)
\]

\[
\lesssim P \left( e^{cN(TK)} |m_i^U(\Lambda_i^L, \Lambda_i^U, F; X) - m_i^L(\Lambda_i^L, \Lambda_i^U, F; X)|^2 \right)
\]

\[
\lesssim P \left[ \left\{ \Delta \sum_{j=1}^K (\triangle N_j + 1) (\triangle \Lambda_{i,j}^U(Y) - \triangle \Lambda_{i,j}^L(Y)) \right. \right.
\]

\[
+ (1 - \Delta) \sum_{j=1}^K \int_Y^{\infty} (\triangle N_j + 1) (\triangle \Lambda_{i,j}^U(u) - \triangle \Lambda_{i,j}^L(u)) dF(u) \left. \right\}^2
\]

\[
\lesssim P \left[ \sum_{j=1}^K \left\{ \Delta (\triangle \Lambda_{i,j}^U(Y) - \Lambda_{i,j}^L(Y))^2 + (1 - \Delta) \int_Y^{\infty} (\triangle \Lambda_{i,j}^U(u) - \triangle \Lambda_{i,j}^L(u))^2 dF(u) \right. \right.
\]

\[
\left. \left. \left. \right. \right. ' \right. \right. \left. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \right. \r
S3. PROOF OF THEOREM 1

\begin{align*}
&= \mathcal{P} \left[ \sum_{j=1}^{K} \Delta (2 \Delta N_j - \Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y)) (\Delta \Lambda_{0,j}(Y) - \Delta \Lambda_j(Y)) \\
&\quad + \sum_{j=1}^{K} (1 - \Delta) \int_{Y}^{\infty} (2 \Delta N_j - \Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u)) (\Delta \Lambda_{0,j}(u) - \Delta \Lambda_j(u)) \, dF_0(u) \right] \\
&= \mathcal{P} \left[ \sum_{j=1}^{K} \Delta \{ \Delta \Lambda_{0,j}(Y) - \Delta \Lambda_j(Y) \}^2 + \sum_{j=1}^{K} (1 - \Delta) \int_{Y}^{\infty} \{ \Delta \Lambda_{0,j}(u) - \Delta \Lambda_j(u) \}^2 \, dF_0(u) \right] \\
&= d_1(\Lambda, \Lambda_0)^2.
\end{align*}

Thus, to draw the conclusion, we only need to consider \( \mathcal{P} m(\Lambda, F_0; X) - \mathcal{P} m(\Lambda_0, F_0; X) \).

Since \( \hat{\Lambda}_n \) is the minimizer of \( \mathbb{P}_n m(\Lambda, \hat{F}_n; X) \) with respect to \( \Lambda \in \Phi_n \), for any direction function \( h \in \Phi_n \), we obtain

\begin{equation}
0 = \lim_{\varepsilon \to 0} \frac{\mathbb{P}_n L(\hat{\Lambda}_n + \varepsilon h, \hat{F}_n; X) - \mathbb{P}_n L(\hat{\Lambda}_n, \hat{F}_n; X)}{\varepsilon} = -2 \mathbb{P}_n \left[ \sum_{j=1}^{K} \left\{ \Delta (\Delta N_j - \Delta \hat{\Lambda}_{n,j}(Y)) \Delta h_j(Y) + (1 - \Delta) \int_{Y}^{\infty} (\Delta N_j - \Delta \hat{\Lambda}_{n,j}(u)) \Delta h_j(u) \, d\hat{F}_n(u) \right\} \right].
\end{equation}

Taking \( h(s) = s \), by Conditions (C1) and (C6), it follows that

\begin{align*}
&\mathbb{P}_n \left[ \sum_{j=1}^{K} \left\{ (T_j - T_{j-1}) \left( \Delta \Delta \hat{\Lambda}_{n,j}(Y) + (1 - \Delta) \int_{Y}^{\infty} \Delta \hat{\Lambda}_{n,j}(u) \, d\hat{F}_n(u) \right) \right\} \right] \\
= &\mathbb{P}_n \left[ \sum_{j=1}^{K} \{ \Delta N_j(T_j - T_{j-1}) \} \right] \xrightarrow{a.s.} E \left[ \sum_{j=1}^{K} \Delta \Lambda_{0,j}(U)(T_j - T_{j-1}) \right] \leq M_3 \Lambda_0(\tau) \tau,
\end{align*}

where the almost surely convergence follows from the strong law of large
number. Moreover, for the left hand side, by Condition (C7), we have

\[
\limsup_{n \to \infty} \mathbb{P}_n \left\{ \sum_{j=1}^K \left( (T_j - T_{j-1}) \left( \Delta \hat{\Lambda}_{n,j} (Y) + (1 - \Delta) \int_0^\infty \frac{\hat{\Lambda}_n (u) d\hat{F}_n (u)}{1 - \hat{F}_n (Y)} \right) \right\} \right.
\]

\[
\geq \limsup_{n \to \infty} \mathbb{P}_n \left[ \Delta \{ \hat{\Lambda}_n (Y) - \hat{\Lambda}_n (Y - T_K) \} + (1 - \Delta) \int_0^\infty 1_{\{T_K \in [0,b_1], u \in [b_2, \tau]\}} d\hat{F}_n (u) \right]
\]

\[
\geq \limsup_{n \to \infty} \Delta \hat{\Lambda}_n (b_1, b_2) \mathbb{P}_n \left[ \Delta1_{\{Y - T_K \in [0,b_1], Y \leq \tau\}} + (1 - \Delta) \int_0^\infty 1_{\{u - T_K \in [0,b_1], u \in [b_2, \tau]\}} d\hat{F}_n (u) \right]
\]

\[
= \limsup_{n \to \infty} \Delta \hat{\Lambda}_n (b_1, b_2) E \left[ 1_{\{U - T_K \in [0,b_1], U \in [b_2, \tau]\}} \right] = \limsup_{n \to \infty} \Delta \hat{\Lambda}_n (b_1, b_2) \mu_2 ([0, b_1] \times [b_2, \tau]).
\]

Hence, for every \( 0 \leq b_1 \leq b_2 \leq \tau \) satisfying \( \mu_2 ([0, b_1] \times [b_2, \tau]) > 0 \), we have

\[
\Delta \hat{\Lambda}_n (s_1, s_2) 1_{\{(s_1, s_2) \in [b_1, b_2] \times [b_1, b_2]\}}
\]

is uniformly bounded. In particular, if \( \mu_2 (\{0\} \times \{\tau\}) > 0 \), then \( \hat{\Lambda}_n (s) \) is uniformly bounded.

By Lemma A1 of [Lu, Zhang, and Huang (2007)], under Condition (C5), there is \( \Lambda^*_n \in \Phi_n \) such that \( \|\Lambda^*_n - \Lambda_0\|_{\infty} = O(n^{-\nu r}) \). This implies that

\[
\mathcal{P}_m(\Lambda^*_n, F_0; X) - \mathcal{P}_m(\Lambda_0, F_0; X) = \mathcal{P} \left[ \sum_{j=1}^K \Delta (\Delta \Lambda_{0,j} (Y) - \Delta \Lambda^*_n, j (Y))^2 \right] + \sum_{j=1}^K (1 - \Delta) \int_0^\infty \frac{(\Delta \Lambda_{0,j} (u) - \Delta \Lambda^*_n, j (u))^2 dF_0 (u)}{1 - F_0 (Y)} \]

\[
\leq \|\Lambda^*_n - \Lambda_0\|_{\infty}^2 = O(n^{-2 \nu r}) = o(1).
\]
Note that
\[ 0 \leq \mathcal{P}m(\hat{\Lambda}_n, F_0; X) - \mathcal{P}m(\Lambda_0, F_0; X) \]
\[ = \mathcal{P}m(\hat{\Lambda}_n, F_0; X) - \mathcal{P}m(\hat{\Lambda}_n, \hat{F}_n; X) + \mathcal{P}m(\hat{\Lambda}_n, \hat{F}_n; X) - \mathbb{P}_n m(\hat{\Lambda}_n, \hat{F}_n; X) \]
\[ + \mathbb{P}_n m(\Lambda_n, \hat{F}_n; X) - \mathbb{P}_n m(\Lambda_n, \hat{F}_n; X) + \mathbb{P}_n m(\Lambda_n, \hat{F}_n; X) - \mathcal{P}m(\Lambda_n, \hat{F}_n; X) \]
\[ + \mathcal{P}m(\Lambda_n, \hat{F}_n; X) - \mathcal{P}m(\Lambda_n, F_0; X) + \mathcal{P}m(\Lambda_n, F_0; X) - \mathcal{P}m(\Lambda_n, F_0; X). \]

By Lemma 1, \( \mathcal{P}m(\hat{\Lambda}_n, \hat{F}_n; X) - \mathcal{P}m(\hat{\Lambda}_n, F_0; X) = o_p(1) \) and \( \mathcal{P}m(\Lambda_n, \hat{F}_n; X) - \mathcal{P}m(\Lambda_n, F_0; X) = o_p(1) \). By Lemma 2, the class of functions
\[ \{m(\Lambda, F; X) : \Lambda \in \Phi_n, F \in F_\delta\} \]
is Donsker. Hence it is Glivenko-Cantelli, and we have
\[ (\mathbb{P}_n - \mathcal{P})m(\Lambda_n^*, \hat{F}_n; X) = o_p(1) \] and \( (\mathbb{P}_n - \mathcal{P})m(\hat{\Lambda}_n, \hat{F}_n; X) = o_p(1) \)
since \( \hat{\Lambda}_n \) is uniformly bounded. According to the definition of \( \hat{\Lambda}_n \),
\[ \mathbb{P}_n m(\hat{\Lambda}_n, \hat{F}_n; X) - \mathbb{P}_n m(\Lambda_n^*, \hat{F}_n; X) \leq 0. \]
Hence, \( d_1(\hat{\Lambda}_n, \Lambda_0) = \mathcal{P}m(\hat{\Lambda}_n, F_0; X) - \mathcal{P}m(\Lambda_0, F_0; X) = o_p(1). \)

**S4 Proof of Theorem 2**

*Proof.* To apply Lemma 3, we need to verify that for every \( \Lambda \in \Phi_n \) and sufficiently large \( n \), the inequalities
\[ \mathcal{P}(m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \leq -d_1(\Lambda, \Lambda_0)^2 + d_1(\Lambda, \Lambda_0)d_2(\hat{F}_n, F_0) + d_2(\hat{F}_n, F_0)^2 \]
and

$$E \sup_{\{\Lambda \in \Phi_n: d_1(\Lambda, \Lambda_0) \leq \eta\}} \left| (\mathbb{P}_n - \mathbb{P}) (m(\Lambda_0, \hat{F}_n; X) - m(\Lambda, \hat{F}_n; X)) \right| \lesssim \frac{\phi_n(\eta)}{\sqrt{n}}$$

hold. By some calculations,

$$\mathcal{P}(m(\Lambda_0, F; X) - m(\Lambda, F; X))$$

$$=\mathcal{P} \left[ \sum_{j=1}^{K} \Delta \left\{ (\Delta N_j - \Delta \Lambda_{0,j}(Y))^2 - (\Delta N_j - \Delta \Lambda_j(Y))^2 \right\} \right]$$

$$+ \sum_{j=1}^{K} (1 - \Delta) \left\{ \int_{Y}^{\infty} \frac{(\Delta N_j - \Delta \Lambda_{0,j}(u))^2 - (\Delta N_j - \Delta \Lambda_j(u))^2}{1 - F(Y)} dF(u) \right\}$$

$$=\mathcal{P} \left[ \sum_{j=1}^{K} \Delta \left\{ 2\Delta N_j - \Delta \Lambda_{0,j}(Y) - \Delta \Lambda_j(Y) \right\} \left\{ \Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y) \right\} \right]$$

$$+ \sum_{j=1}^{K} (1 - \Delta) \int_{Y}^{\infty} \frac{2\Delta N_j - \Delta \Lambda_{0,j}(u) - \Delta \Lambda_j(u)}{1 - F(Y)} \left\{ \Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u) \right\} dF(u)$$

$$= -\mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta(\Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y))^2 + (1 - \Delta) \int_{Y}^{\infty} \frac{(\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2}{1 - F(Y)} dF(u) \right\} \right]$$

$$=\mathcal{P} \left[ \sum_{j=1}^{K} (1 - \Delta) \left\{ \frac{\int_{Y}^{\infty} \frac{(\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2}{1 - F_0(Y)} dF_0(u)}{1 - F(Y)} - \frac{\int_{Y}^{\infty} \frac{(\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2}{1 - F(Y)} dF(u)}{1 - F_0(Y)} \right\} \right]$$

$$-d_1(\Lambda, \Lambda_0)^2.$$
Moreover, by the first part of Lemma 1 under Condition (C1), we have
\[ \begin{align*}
\mathcal{P} \left[ \sum_{j=1}^{K} (1 - \Delta) \left\{ \frac{\int_{Y}^{\infty} (\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2 dF_0(u)}{1 - F_0(Y)} - \frac{\int_{Y}^{\infty} (\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2 dF(u)}{1 - F(Y)} \right\} \right] \\
\lesssim \mathcal{P} \left[ \sum_{j=1}^{K} 2 \left\{ \Delta \Lambda_j(U) - \Delta \Lambda_{0,j}(U) \right\} \left\{ \Delta \Lambda_j(U) - \Delta \Lambda_{0,j}(U) \right\} \right] ||F - F_0||_{\infty} \\
+ \mathcal{P} \left[ \sum_{j=1}^{K} \left( \Delta \Lambda_j(U) - \Delta \Lambda_{0,j}(U) \right)^2 \right] ||F - F_0||_{\infty} \\
\lesssim d_1(\Lambda, \Lambda_0) d_2(F, F_0) + d_1(\Lambda, \Lambda_0)^2 d_2(F, F_0).
\end{align*} \]

This implies that
\[ \mathcal{P} m(\Lambda_0, \hat{F}_n; X) - \mathcal{P} m(\Lambda, \hat{F}_n; X) \lesssim -d_1(\Lambda, \Lambda_0)^2 + d_1(\Lambda, \Lambda_0) d_2(\hat{F}_n, F_0) + d_1(\Lambda, \Lambda_0)^2 d_2(\hat{F}_n, F_0). \]

Second, we need to find a \( \phi_n(\eta) \) such that
\[ E \sup_{\{\Lambda \in \Phi_n : d_1(\Lambda, \Lambda_0) < \eta\}} \left| (\mathbb{P}_n - \mathcal{P})(m(\Lambda, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X)) \right| \lesssim \frac{\phi_n(\eta)}{\sqrt{n}}. \]

By Lemma 4, we have \( \log N_{\|\cdot\|_{P,B}}(\varepsilon, \mathcal{M}_\eta(\hat{F}_n), ||\cdot||_{P,B}) \lesssim q_n \log(\eta/\varepsilon) \), where
\[ \mathcal{M}_\eta(\hat{F}_n) = \{ m(\Lambda, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X) : \Lambda \in \Phi_n, d_1(\Lambda, \Lambda_0) \leq \eta \}. \]

For all \( \Lambda \in \Phi_n, d_1(\Lambda, \Lambda_0) \leq \eta \), note that
\[ |m(\Lambda, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X)| \]
\[ = \sum_{j=1}^{K} \Delta \left| (\Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y)) (\Delta \Lambda_j(Y) + \Delta \Lambda_{0,j}(Y) - 2\Delta N_j) \right| \]
\[ + \sum_{j=1}^{K} (1 - \Delta) \left| \frac{\int_{Y}^{\infty} (\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u)) (\Delta \Lambda_j(u) + \Delta \Lambda_{0,j}(u) - 2\Delta N_j) d\hat{F}_n(u)}{1 - \hat{F}_n(Y)} \right| \]
\[ \lesssim \sum_{j=1}^{K} \left[ (\Delta N_j + 1) \left\{ \Delta |\Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y)| + (1 - \Delta) \frac{\int_{Y}^{\infty} |\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u)| d\hat{F}_n(u)}{1 - \hat{F}_n(Y)} \right\} \right]. \]
Similar to the proof of (S2.4), since \( \hat{\Lambda}_n \) is uniformly bounded, it follows that

\[
e^{m(\hat{\Lambda}_n, F_n; X) - m(\Lambda_0, F_0; X)} \lesssim e^{cN(T_K)} \text{ and}
\]

\[
\mathcal{P} \left[ e^{m(\hat{\Lambda}_n, F_n; X) - m(\Lambda_0, F_0; X)} \right] \lesssim \mathcal{P} \left[ e^{cN(T_K)} \sum_{j=1}^{K} \left\{ \Delta(\Delta N_j + 1)^2(\Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y))^2 \\
+ (1 - \Delta)(\Delta N_j + 1)^2 \int_{Y}^{\infty} (\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2 d\hat{F}_n(u) \right\} \right]\]

\[
\lesssim \mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta(\Delta \Lambda_j(Y) - \Delta \Lambda_{0,j}(Y))^2 + (1 - \Delta) \int_{Y}^{\infty} (\Delta \Lambda_j(u) - \Delta \Lambda_{0,j}(u))^2 d\hat{F}_n(u) \right\} \right]\]

\[
\lesssim d_1^2(\hat{\Lambda}_n, \Lambda_0) + d_1(\hat{\Lambda}_n, \Lambda_0) d_2(\hat{F}_n, F_0).
\]

That means that for sufficiently large \( n \) with \( d_2(\hat{F}_n, F_0) \lesssim \eta \), we have

\[
||m(\hat{\Lambda}_n, F_n; X) - m(\Lambda_0, F_0; X)||_{P,B}^2 \lesssim \eta^2.
\]

Then Lemma 3.4.3 of van der Vaart and Wellner (1996) yields that

\[
E||n^{1/2}(P_n - P)||_{\mathcal{M}_n(\hat{F}_n)} \lesssim J_{[\eta]}(\eta, \mathcal{M}_n(\hat{F}_n), || \cdot ||_{P,B}) \left\{ 1 + \frac{J_{[\eta]}(\eta, \mathcal{M}_n(\hat{F}_n), || \cdot ||_{P,B})}{\eta^2 n^{1/2}} \right\},
\]

where \( J_{[\eta]}(\eta, \mathcal{M}_n(\hat{F}_n), || \cdot ||_{P,B}) := \int_{0}^{\eta} \{ 1 + \log \mathcal{N}_0(\varepsilon, \mathcal{M}_n(\hat{F}_n), || \cdot ||_{P,B}) \}^{1/2} d\varepsilon \lesssim \eta^{1/2} \cdot \eta. \) It follows that

\[
E \sup_{\{A \in \Phi_n : d_1(A, \Lambda_0) \lesssim \eta\}} \sqrt{n}||P_n - P||(m(A, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X)) \lesssim \eta^{1/2} + q_n n^{-1/2}.
\]

Set \( \phi_n(\eta) = q_n^{1/2} \eta + q_n n^{-1/2} \). It is clear that \( \phi_n(\eta)/\eta \) is decreasing about \( \eta \).

Moreover, \( r_n^2 \varphi(1/r_n) = q_n^{1/2} r_n + n^{-1/2} q_n r_n^2 \), where \( r_n = O(n^a) \). Note that
$q_n = O(n^\nu)$ with $0 < \nu < 1/2$. It follows that

$$r_n^2 \phi\left(\frac{1}{r_n} \right) = O\left(n^{a+\nu/2} + n^{2a+\nu - 1/2}\right).$$

Thus, $a \leq (1 - \nu)/2$ ensures $r_n^2 \phi(1/r_n) \lesssim n^{1/2}$. This implies that $r_n = O(n^{(1-\nu)/2})$.

According to the proof of Theorem 1 and the definition of $\hat{\Lambda}_n$, we have

$$\Pr_n m(\hat{\Lambda}_n, \hat{F}_n; X) - \Pr_n m(\Lambda_0, \hat{F}_n; X)$$

$$= \Pr_n m(\hat{\Lambda}_n, \hat{F}_n; X) - \Pr_n m(\Lambda^*_n, \hat{F}_n; X) + \Pr_n m(\Lambda^*_n, \hat{F}_n; X) - \Pr_n m(\Lambda_0, \hat{F}_n; X)$$

$$+ \Pr m(\Lambda_0, \hat{F}_n; X) + \Pr m(\Lambda_0, \hat{F}_n; X) - \Pr_n m(\Lambda_0, \hat{F}_n; X)$$

$$\leq n^{-\nu r + \varepsilon} (\Pr_n - \Pr) \left( m(\Lambda^*_n, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X) \right) + O_p(n^{-2\nu r})$$

for any $0 < \varepsilon < 1/2 - \nu r$. Set the class

$$\tilde{\mathcal{M}}_n = \left\{ \frac{m(\Lambda, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X)}{n^{-\nu r + \varepsilon}} : \Lambda \in \Phi_n, ||\Lambda - \Lambda_0||_{\infty} = O(n^{-\nu r}) \right\}.$$  

Similar to the proof of Theorem 2 in [Lu, Zhang, and Huang (2009)], under Conditions (C2) and (C5), we have $\tilde{\mathcal{M}}_n$ is Donsker, and $\Pr \tilde{m}^2 \to 0$ as $n \to \infty$ for any $\tilde{m} \in \tilde{\mathcal{M}}_n$. Hence,

$$n^{-\nu r + \varepsilon} (\Pr_n - \Pr) \left( m(\Lambda^*_n, \hat{F}_n; X) - m(\Lambda_0, \hat{F}_n; X) \right) = o_p(n^{-\nu r + \varepsilon - 1/2}) = o_p(n^{-2\nu r}).$$

This implies that $\Pr_n m(\hat{\Lambda}_n, \hat{F}_n; X) - \Pr_n m(\Lambda_0, \hat{F}_n; X) \leq O_p(n^{-2\nu r})$. Noting that $\hat{\Lambda}_n$ is needed to satisfy $\Pr_n m(\Lambda_0, \hat{F}_n; X) \geq \Pr_n m(\hat{\Lambda}_n, \hat{F}_n; X) - O_p(r_n^{-2})$, so we should take $\nu$ such that $O_p(n^{-2\nu r}) \leq O_p(r_n^{-2})$. Since $r_n = O(n^{(1-\nu)/2})$,
it follows that $\nu \geq 1/(1 + 2r)$. Taking $\nu = 1/(1 + 2r)$ and by Lemma 3, we have $d_1(\hat{\Lambda}_n, \Lambda_0) = O_p(n^{-r/(1+2r)} + n^{-1/2}) = O_p(n^{-r/(1+2r)})$.

S5 Proof of Theorem 3

Proof. (i) To prove this part, following Zhao and Zhang (2017), we need to verify the following conditions.

(B1) $Q(\Lambda_0, F_0)[h] = 0$ and $Q_n(\hat{\Lambda}_n, \hat{F}_n)[h] = o_p(n^{-1/2})$.

(B2) $\sqrt{n}(Q_n - Q)(\hat{\Lambda}_n, \hat{F}_n)[h] - \sqrt{n}(Q_n - Q)(\Lambda_0, F_0)[h] = o_p(1)$.

(B3) $Q(\hat{\Lambda}_n, \hat{F}_n)[h] - Q(\Lambda_0, F_0)[h] - Q_{\Lambda_0, F_0}^{(1)}(\hat{\Lambda}_n - \Lambda_0)[h] - Q_{\Lambda_0, F_0}^{(2)}(\hat{F}_n - F_0)[h] = o_p(n^{-1/2})$.

For (B1), since $E(\tilde{N}(t)|U = u) = \Lambda_0(u - t)$, we have $Q(\Lambda_0, F_0)[h] = 0$. Note that $\hat{\Lambda}_n$ is obtained by minimizing the loss function $\mathbb{P}_n m(\Lambda, \hat{F}_n; X)$. By (S3.5), we have $Q_n(\hat{\Lambda}_n, \hat{F}_n)[h_n] = 0$ for all $h_n \in \Phi_n$. According to Lemma A1 of Lu, Zhang, and Huang (2007) and the properties of spline functions, for any $h \in \mathcal{H}_r$, there is an $h_n \in \Phi_n$ such that $\|h_n - h\|_\infty = O(n^{-rv}) = O(n^{-r/(1+2r)})$ and $\|h'_n - h'\|_\infty = o(1)$, where $h'$ is the derivative of $h$. Next, to prove $Q_n(\hat{\Lambda}_n, \hat{F}_n)[h] = o_p(n^{-1/2})$, we need to show that
For the first term, by Lemma 1, we have

\[ Q_n(\hat{\Lambda}_n, \hat{F}_n)[h - h_n] = \mathbb{P}_n \psi(\hat{\Lambda}_n, \hat{F}_n; X)[h - h_n] = o_p(n^{-1/2}). \]

Note that

\[
Q_n(\hat{\Lambda}_n, \hat{F}_n)[h - h_n] = \left[ Q_n(\hat{\Lambda}_n, \hat{F}_n)[h - h_n] - Q_n(\hat{\Lambda}_n, F_0)[h - h_n] \right] \\
+ \left[ Q_n(\hat{\Lambda}_n, F_0)[h - h_n] - Q_n(\Lambda_0, F_0)[h - h_n] \right] + Q_n(\Lambda_0, F_0)[h - h_n] \\
= I_{1n} + I_{2n} + I_{3n}.
\]

For the first term, by Lemma 1 we have

\[
\mathbb{P}|I_{1n}| = \mathbb{P}|Q_n(\hat{\Lambda}_n, \hat{F}_n)[h - h_n] - Q_n(\hat{\Lambda}_n, F_0)[h - h_n]| \\
\leq \mathbb{P} \left[ \sum_{j=1}^{K} (1 - \Delta) \left| \int_{Y} \left( \triangle N_j - \triangle \hat{\Lambda}_{n,j}(u) \right) \cdot \left( \triangle h_j(u) - \triangle h_{n,j}(u) \right) dF_0(u) \right| \right. \\
\leq \left. \int_{Y} \left( \triangle N_j - \triangle \hat{\Lambda}_{n,j}(u) \right) \cdot \left( \triangle h_j(u) - \triangle h_{n,j}(u) \right) d\hat{F}_n(u) \right| \\
\leq d_2(\hat{F}_n, F_0)(||h - h_n||_\infty + ||h' - h'_n||_\infty) = o_p(n^{-1/2}).
\]

For \( I_{2n} \), by the Cauchy-Schwarz inequality, we have

\[
\mathbb{P}|I_{2n}| = \mathbb{P}|Q_n(\hat{\Lambda}_n, F_0)[h - h_n] - Q_n(\Lambda_0, F_0)[h - h_n]| \\
\leq \mathbb{P} \left[ \sum_{k=1}^{K} \left| \triangle \Lambda_{0,j}(Y) - \triangle \hat{\Lambda}_{n,j}(Y) \right| \cdot \left| \triangle h_j(Y) - \triangle h_{n,j}(Y) \right| \right] \\
+ (1 - \Delta) \int_{Y} \left| \triangle \Lambda_{0,j}(u) - \triangle \hat{\Lambda}_{n,j}(u) \right| \cdot \left| \triangle h_j(u) - \triangle h_{n,j}(u) \right| dF_0(u) \\
\leq \mathbb{P} \left[ \sum_{k=1}^{K} \left\{ \left| \triangle \Lambda_{0,j}(Y) - \triangle \hat{\Lambda}_{n,j}(Y) \right| \right. \right. \\
\left. \left. + (1 - \Delta) \int_{Y} \left| \triangle \Lambda_{0,j}(u) - \triangle \hat{\Lambda}_{n,j}(u) \right| dF_0(u) \right| \right) \right] ||h - h_n||_\infty \\
\leq d_1(\hat{\Lambda}_n, \Lambda_0)||h - h_n||_\infty = o_p(n^{-1/2}).
\]
For the third term, note that \( P\psi(\Lambda_0, F_0; X_i)[h - h_n] = 0 \). By the independence between \( X_i \) and \( X_j \) when \( i \neq j \), we obtain

\[
P I_{3n}^2 = P \left( \frac{1}{n} \sum_{i=1}^{n} \psi(\Lambda_0, F_0; X_i)[h - h_n] \right)^2 = n^{-1} P \left( \frac{1}{n} \sum_{i=1}^{n} \psi^2(\Lambda_0, F_0; X_i)[h - h_n] \right)
\]

\[
\leq n^{-1} P \left[ \sum_{j=1}^{K} \left\{ \Delta|\Delta N_j - \Delta \Lambda_{0,j}(Y)| + (1 - \Delta) \int_{Y}^{\infty} \frac{\Delta N_j - \Delta \Lambda_{0,j}(u)}{1 - F_0(Y)} dF_0(u) \right\} \right]^2 ||h - h_n||^2_{\infty}
\]

\[
\leq n^{-1} ||h - h_n||^2_{\infty}.
\]

Thus, \( Q_n(\hat{\Lambda}_n, \hat{F}_n)[h - h_n] = o_p(n^{-1/2}) \).

For (B2), note that

\[
\sqrt{n}(Q_n - Q)(\hat{\Lambda}_n, \hat{F}_n)[h] - \sqrt{n}(Q_n - Q)(\Lambda_0, F_0)[h] = \sqrt{n}(\mathbb{P}_n - P)(\psi(\hat{\Lambda}_n, \hat{F}_n; X)[h] - \psi(\Lambda_0, F_0; X)[h]).
\]

For a bounded function \( h \in \mathcal{H}_r \), define

\[
\tilde{\Psi}_\eta(h) = \{ \psi(\Lambda, F; X)[h] - \psi(\Lambda_0, F_0; X)[h] : \Lambda \in \Phi_n, F \in \mathcal{F}, d_1(\Lambda, \Lambda_0) < \eta, d_2(F, F_0) < \eta, \Lambda \text{ is uniformly bounded} \}.
\]

Similar to the proof of Lemma 2, \( \tilde{\Psi}_\eta(h) \) is Donsker. According to Condition
\(\text{(C6)}\) and Lemma \[\text{[1]}\) we obtain

\[
P(\psi(\Lambda, F; X)[h] - \psi(\Lambda_0, F_0; X)[h])^2
= \mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta(\Delta \lambda_{a,j}(Y) - \Delta \lambda_j(Y)) \Delta h_j(Y) + (1 - \Delta) \frac{\int_{Y}^{\infty} \{\Delta N_j - \Delta \lambda_j(u)\} \Delta h_j(u) dF_0(u)}{1 - F_0(Y)} \right\} \right]^2
\]

+ (1 - \Delta) \left( \frac{\int_{Y}^{\infty} \{\Delta N_j - \Delta \lambda_j(u)\} \Delta h_j(u) dF_0(u)}{1 - F_0(Y)} \right)^2
\]

\[\leq \mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta(\Delta \lambda_{a,j}(Y) - \Delta \lambda_j(Y))^2 \Delta h^2_j(Y) + (1 - \Delta) \frac{\int_{Y}^{\infty} \{\Delta N_j - \Delta \lambda_j(u)\} \Delta h_j(u) dF_0(u)}{1 - F_0(Y)} \right\} \right]^2
\]

\[+ \mathcal{P} \left[ (1 - \Delta) \left( \frac{\int_{Y}^{\infty} \{\Delta N_j - \Delta \lambda_j(u)\} \Delta h_j(u) dF_0(u)}{1 - F_0(Y)} \right)^2 \right]^2
\]

\[\leq d_1(\Lambda, \Lambda_0)^2 + d_2(F, F_0)^2.
\]

It follows that for any function \(\tilde{\psi} \in \tilde{\Psi}(h),\)

\[
\sup_{\tilde{\psi} \in \tilde{\Psi}_\eta(h)} \rho_\mathcal{P}(\tilde{\psi}) \leq \sup_{\tilde{\psi} \in \tilde{\Psi}_\eta(h)} \mathcal{P}(\tilde{\psi}^2)^{1/2} \approx \eta \to 0
\]

as \(\eta \to 0\) for the seminorm \(\rho_\mathcal{P}(\tilde{\psi}) = \{\mathcal{P}(\tilde{\psi} - \mathcal{P}\tilde{\psi})^2\}^{1/2}\). Then by Corollary 2.3.12 of \cite{van_der_Vaart}, we have

\[
\sqrt{n}(\mathbb{P}_n - \mathcal{P})(\psi(\hat{\Lambda}_n, \hat{F}_n; X)[h] - \psi(\Lambda_0, F_0; X)[h]) = o_p(1), \quad \text{(S5.6)}
\]

and (B2) holds.

Next, we verify (B3). Note that

\[
\hat{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h]
= \mathcal{P} \left[ (1 - \Delta) \frac{1 - \hat{F}_n(Y)}{1 - F_0(Y)} \sum_{j=1}^{K} \left\{ \int_{Y}^{\infty} \{\Delta N_j - \Delta \lambda_{a,j}(u)\} \Delta h_j(u) d\hat{F}_n(u) \right\} \right].
\]
By the conclusion of Lemma 1, we have

\[
|Q(\Lambda_0, \hat{F}_n)[h] - Q(\Lambda_0, F_0)[h] - \dot{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h]| = 0
\]

Thus, (B1)-(B3) hold. By (B1) and (B3), we have

\[
\text{Combining the above two equations, it follows that}
\]

\[
Q(\hat{\Lambda}_n, \hat{F}_n)[h] - Q(\Lambda_0, \hat{F}_n)[h] - \dot{Q}^{(1)}_{\Lambda_0, \hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] - \dot{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h] = 0.
\]

Moreover, \(Q(\hat{\Lambda}_n, \hat{F}_n)[h] - Q(\Lambda_0, \hat{F}_n)[h] - \dot{Q}^{(1)}_{\Lambda_0, \hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] = 0\). It follows that

\[
Q(\hat{\Lambda}_n, \hat{F}_n)[h] - Q(\Lambda_0, F_0)[h] - \dot{Q}^{(1)}_{\Lambda_0, \hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] - \dot{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h] = o_p(n^{-1/2}).
\]

Thus, (B1)-(B3) hold. By (B1) and (B3), we have

\[
Q(\hat{\Lambda}_n, \hat{F}_n)[h] - \dot{Q}^{(1)}_{\Lambda_0, \hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] - \dot{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h] = 0.
\]

By (B1) and (B2), we have

\[
-\sqrt{n}\dot{Q}^{(1)}_{\Lambda_0, \hat{F}_n}(\hat{\Lambda}_n - \Lambda_0)[h] = \sqrt{n}\dot{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h] + \sqrt{n}Q_n(\Lambda_0, F_0)[h] + o_p(1).
\]

(ii) To prove the second part, we need to rewrite \(\dot{Q}^{(2)}_{\Lambda_0, F_0}(\hat{F}_n - F_0)[h] + Q_n(\Lambda_0, F_0)[h]\). Let

\[
\tilde{\varphi}_{\Lambda, F}(u, a; X)[h] = (1 - F(u))\tilde{\varphi}_{\Lambda, F}(u; X)[h] - \int_u^a \tilde{\varphi}_{\Lambda, F}(s; X)[h]dF(s).
\]
Note that $\hat{F}_n$ is the KM estimator based on the data $\{(\bar{Y}_i, \bar{\Delta}_i) : i = 1, \cdots, n\}$. Setting $G_0(s) = \mathcal{P}1_{\{\bar{Y} \geq s\}}$, then for any constant $a$, by Propositions 3 and 4 of \cite{Akritas}, we have

$$
\int_0^a \phi_{\Lambda_0, F_0}(u, X)[h]d(\hat{F}_n - F_0)(u) = \frac{1}{n} \sum_{i=1}^n \int_0^a \frac{\tilde{\phi}_{\Lambda_0, F_0}(u, a; X)[h]}{G_0(u)}d\tilde{M}_i(u) + o_p(n^{-1/2}),
$$

where $\tilde{M}_i(u) = 1_{\{Y_i \leq u, \bar{\Delta}_i = 1\}} - \int_{-\infty}^u 1_{\{Y_i \geq s\}}/(1 - F_0(s))dF_0(s)$. It follows that

$$
\int_0^a \phi_{\Lambda_0, F_0}(u, X)[h]d(\hat{F}_n - F_0)(u)
\geq \frac{1}{n} \sum_{i=1}^n \left\{ \int_0^a \frac{\tilde{\phi}_{\Lambda_0, F_0}(u, a; X)[h]}{G_0(u)}d\tilde{M}_i(u) - \int_0^a \frac{\tilde{\phi}_{\Lambda_0, F_0}(u, Y; X)[h]}{G_0(u)}d\tilde{M}_i(u) \right\} + o_p(n^{-1/2})
$$

$$
\geq \frac{1}{n} \sum_{i=1}^n \left\{ 1_{\{\bar{\Delta}_i = 1\}} \phi_{\Lambda_0, F_0}(\bar{Y}_i, \infty; X)[h] - \int_0^{\bar{Y}_i} \frac{\phi_{\Lambda_0, F_0}(u, \infty; X)[h]}{G_0(u)(1 - F_0(u))}dF_0(u)

\quad + \int_0^{\bar{Y}_i \wedge Y} \frac{\phi_{\Lambda_0, F_0}(u, Y; X)[h]}{G_0(u)(1 - F_0(u))}dF_0(u) - 1_{\{\bar{\Delta}_i = 1, Y_i \geq \bar{Y}_i\}} \frac{\tilde{\phi}_{\Lambda_0, F_0}(\bar{Y}_i, Y; X)[h]}{G_0(\bar{Y}_i)} \right\} + o_p(n^{-1/2}).
$$

Hence, we obtain $\dot{Q}_{\Lambda_0, F_0}^{(2)}(\hat{F}_n - F_0)[h] = \mathbb{P}_n\{\mathcal{P}_n \phi(\Lambda_0, F_0; X; \bar{Y}, \bar{\Delta})[h]\}\approx o_p(n^{-1/2})$,

where

$$
\phi(\Lambda, F; X; \bar{Y}, \bar{\Delta})[h]
\geq 1_{\{\bar{\Delta} = 1\}} \frac{\tilde{\phi}_{\Lambda, F}(\bar{Y}, \infty; X)[h]}{G_0(\bar{Y})} - \int_0^{\bar{Y}} \frac{\tilde{\phi}_{\Lambda, F}(u, \infty; X)[h]}{G_0(u)(1 - F_0(u))}dF(u)

\quad + \int_0^{\bar{Y} \wedge Y} \frac{\tilde{\phi}_{\Lambda, F}(u, Y; X)[h]}{G_0(u)(1 - F_0(u))}dF(u) - 1_{\{\bar{\Delta} = 1, Y \geq \bar{Y}\}} \frac{\tilde{\phi}_{\Lambda, F}(\bar{Y}, Y; X)[h]}{G_0(\bar{Y})}.
$$

Noting that $Q_n(\Lambda_0, F_0)[h] = \mathbb{P}_n \psi(\Lambda_0, F_0; X)[h]$, we have

$$
-\sqrt{n} \dot{Q}_{\Lambda_0, \hat{F}_n}^{(1)}(\tilde{\Lambda}_n - \Lambda_0)[h] \xrightarrow{d} N(0, \sigma_0^2),
$$
where $\sigma_0^2 = E[\{P\varphi(\Lambda_0, F_0; X; \tilde{Y}, \Delta)[h] + \psi(\Lambda_0, F_0; X)[h]\}^2]$. \hfill \square

S6 Proof of Theorem 4

Proof. Setting $U_n^{(l)} = \sqrt{n}(\mathbb{P}_n - \mathcal{P})\varsigma(\hat{\Lambda}_l, \hat{F}_n; X)[h_n]$ for $l = 1, 2$, we have $U_n = U_n^{(1)} - U_n^{(2)}$ and $U_n^{(l)} = U_{1n}^{(l)} + U_{2n}^{(l)} + U_{3n}^{(l)}$, where

\begin{align*}
U_{1n}^{(l)} &= \sqrt{n}(\mathbb{P}_n - \mathcal{P})\varsigma(\hat{\Lambda}_l, \hat{F}_n; X)[h_n], \\
U_{2n}^{(l)} &= \sqrt{n}\mathcal{P}\varsigma(\hat{\Lambda}_l, \hat{F}_n; X)[h_n - h], \\
U_{3n}^{(l)} &= \sqrt{n}\mathcal{P}\varsigma(\hat{\Lambda}_l, \hat{F}_n; X)[h].
\end{align*}

For $U_{1n}^{(l)}$, similar to the proof of (S5.6), we have

$$\sqrt{n}(\mathbb{P}_n - \mathcal{P})(\psi(\hat{\Lambda}_l, \hat{F}_n; X)[h_n] - \psi(\Lambda_0, F_0; X)[h_n]) = o_p(1)$$

and

$$\sqrt{n}(\mathbb{P}_n - \mathcal{P})(\psi(\Lambda_0, \hat{F}_n; X)[h_n] - \psi(\Lambda_0, F_0; X)[h_n]) = o_p(1).$$

Noting that $\varsigma(\Lambda, F; X)[h_n] = \psi(\Lambda_0, F; X)[h_n] - \psi(\Lambda, F; X)[h_n]$, it follows that $U_{1n}^{(l)} = o_p(1)$. 

For $U_{2n}^{(l)}$, by Cauchy-Schwarz inequality, we obtain

$$U_{2n}^{(l)} \leq \sqrt{n} \mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta \left( \hat{\Lambda}_{l,j}(Y) - \Lambda_{0,j}(Y) \right)^2 \right. \\
+ (1 - \Delta) \left[ \int_{Y}^{\infty} \left( \hat{\Lambda}_{l,j}(u) - \Lambda_{0,j}(u) \right)^2 d\hat{F}_n(u) \right] \frac{1}{1 - \hat{F}_n(Y)} \right]^{1/2} \\
\times \mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta \left( h_{n,j}(Y) - h_j(Y) \right)^2 \\
+ (1 - \Delta) \left[ \int_{Y}^{\infty} \left( h_{n,j}(u) - h_j(u) \right)^2 d\hat{F}_n(u) \right] \frac{1}{1 - \hat{F}_n(Y)} \right] \right]^{1/2}.$$

By Lemma 1, similar to the proof of the first inequality in Theorem 2, we have

$$\mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta \left( \hat{\Lambda}_{l,j}(Y) - \Lambda_{0,j}(Y) \right)^2 \\
+ (1 - \Delta) \left[ \int_{Y}^{\infty} \left( \hat{\Lambda}_{l,j}(u) - \Lambda_{0,j}(u) \right)^2 d\hat{F}_n(u) \right] \frac{1}{1 - \hat{F}_n(Y)} \right] \right] \lesssim d_1(\hat{\Lambda}_l, \Lambda_0)^2 + d_1(\hat{\Lambda}_l, \Lambda_0) d_2(\hat{F}_n, F_0) = O_p(n^{-2r/(1+2r)})$$

and

$$\mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta \left( h_{n,j}(Y) - h_j(Y) \right)^2 \\
+ (1 - \Delta) \left[ \int_{Y}^{\infty} \left( h_{n,j}(u) - h_j(u) \right)^2 d\hat{F}_n(u) \right] \frac{1}{1 - \hat{F}_n(Y)} \right] \right] \lesssim d_1(h_n, h)^2 + d_1(h_n, h) d_2(\hat{F}_n, F_0) = o_p(n^{-1/(1+2r)}).$$

Hence, $U_{2n}^{(l)} = o_p(1)$.

For $U_{3n}^{(l)}$, by Theorem 3, we have

$$U_{3n}^{(l)} = \sqrt{n} \mathcal{Q}_{\hat{\Lambda}_l, F_0}^{(2)}(\hat{F}_n - F_0)[h] + \sqrt{n} \mathbb{P}_{n} \psi(\Lambda_0, F_0; X)[h] + o_p(1),$$
where $P_{nl}$ is the empirical measure based on group $l$. Thus,

$$U_n = U_{3n}^{(1)} - U_{3n}^{(2)} + o_p(1)$$

$$= \sqrt{\frac{n}{n_1}} \sqrt{n_1} P_{n_1} \psi(\Lambda_0, F_0; X)[h] - \sqrt{\frac{n}{n_2}} \sqrt{n_2} P_{n_2} \psi(\Lambda_0, F_0; X)[h] + o_p(1).$$

Note that $P_{n_1}$ and $P_{n_2}$ are independent, and $\sqrt{\frac{n}{n_1}} P_{n_1} \psi(\Lambda_0, F_0; X)[h] \xrightarrow{d} N(0, \sigma_0^2)$. Thus, we have $U_n \xrightarrow{d} N(0, (\frac{1}{p} + \frac{1}{1-p})\sigma_0^2)$.

Finally, we need to prove that $\hat{\sigma}_n^2 - \sigma_0^2 = o_p(1)$. Note that $\sigma_0^2 = P \psi^2(\Lambda_0, F_0; X)[h]$ and $\hat{\sigma}_n^2 = P_n \psi^2(\hat{\Lambda}_n, \hat{F}_n; X)[h_n]$. Thus, we have

$$\hat{\sigma}_n^2 - \sigma_0^2 = P_n \{\psi^2(\hat{\Lambda}_n, \hat{F}_n; X)[h_n] - \psi^2(\Lambda_0, F_0; X)[h_n]\}$$

$$+ P_n \{\psi^2(\Lambda_0, F_0; X)[h_n] - \psi^2(\Lambda_0, F_0; X)[h]\} + (P_n - P) \psi^2(\Lambda_0, F_0; X)[h].$$

By the consistency of $(\hat{\Lambda}_n, \hat{F}_n)$ and the law of large numbers, we have

$$P_n \{\psi^2(\hat{\Lambda}_n, \hat{F}_n; X)[h_n] - \psi^2(\Lambda_0, F_0; X)[h_n]\} = o_p(1)$$

and

$$(P_n - P) \psi^2(\Lambda_0, F_0; X)[h] = o_p(1).$$

Then we only need to consider $P_n \{\psi^2(\Lambda_0, F_0; X)[h_n] - \psi^2(\Lambda_0, F_0; X)[h]\}$. Since $\Lambda_0$, $F_0$, $h$ and $h_n$ are bounded functions, we have

$$|\psi(\Lambda_0, F_0; X)[h_n] + \psi(\Lambda_0, F_0; X)[h]| = |\psi(\Lambda_0, F_0; X)[h_n + h]|$$

$$\lesssim \sum_{j=1}^K \left[ \Delta \{\Delta N_j - \Delta \Lambda_0, 0, j(Y)\} + (1 - \Delta) \int_{\gamma_{\Delta Y}} \{\Delta N_j - \Delta \Lambda_0, 0, j(u)\} dF_0(u) \right]$$

$$\lesssim (N(T_K) + \Lambda_0(\tau))$$
S7 Proof of Theorem 5

with probability 1. Thus, by Cauchy-Schwarz inequality, we have

\[ \mathcal{P} \left| \psi^2(\Lambda_0, F_0; X)[h_n] - \psi^2(\Lambda_0, F_0; X)[h] \right| \]

\[ \lesssim \mathcal{P} \left[ \psi(\Lambda_0, F_0; X)[h_n] - \psi(\Lambda_0, F_0; X)[h] \right] (N(T_K) + \Lambda_0(\tau)) \]

\[ \lesssim \mathcal{P} \left[ (N(T_K) + \Lambda_0(\tau)) \sum_{j=1}^{K} \left\{ \Delta|\Delta N_j - \Delta \Lambda_{0,j}(Y)| |\Delta h_{n,j}(Y) - \Delta h_j(Y)| \right. \right. \]

\[ + \left. \left. (1 - \Delta) \frac{f_{Y}^{\infty} |\Delta N_j - \Delta \Lambda_{0,j}(u)| |\Delta h_{n,j}(u) - \Delta h_j(u)| dF_0(u)}{1 - F_0(Y)} \right\} \right] \]

\[ \lesssim \mathcal{P} \left[ (N(T_K) + \Lambda_0(\tau))^2 \sum_{j=1}^{K} \{ \Delta|\Delta h_{n,j}(Y) - \Delta h_j(Y)| \right. \right. \]

\[ + \left. \left. (1 - \Delta) \frac{f_{Y}^{\infty} |\Delta h_{n,j}(u) - \Delta h_j(u)| dF_0(u)}{1 - F_0(Y)} \right\} \right] \]

\[ \leq \mathcal{P} \left[ \sum_{j=1}^{K} \left\{ \Delta(\Delta h_{n,j}(Y) - \Delta h_j(Y))^2 + (1 - \Delta) \frac{f_{Y}^{\infty} (\Delta h_{n,j}(u) - \Delta h_j(u))^2 dF_0(u)}{1 - F_0(Y)} \right\} \right]^{\frac{1}{2}} \]

\[ \times \mathcal{P} \left[ (N(T_K) + \Lambda_0(\tau))^4 \right]^{\frac{1}{2}} \]

\[ \lesssim d_1(h_n, h) = o(1). \]

Therefore, \( \mathbb{P}_n \{ \psi^2(\Lambda_0, F_0; X)[h_n] - \psi^2(\Lambda_0, F_0; X)[h] \} = o_p(1) \) and \( \hat{\sigma}_n^2 - \sigma_0^2 = o_p(1). \) \]

\( \square \)

S7 Proof of Theorem 5

Proof. (i) First we write

\[ \tilde{U}_n = \sum_{l=1}^{2} \frac{n_l}{\sqrt{n}} \mathbb{P}_{n_l} \left( \zeta(\hat{\Lambda}_1, \hat{F}_1; X^{(l)}[h_n] - \zeta(\hat{\Lambda}_2, \hat{F}_1; X^{(l)}[h_n]) \right) \]
for \( l = 1, 2 \). Note that

\[
\mathbb{P}_{nl}(\varsigma(\Lambda, \hat{F}_l; X^{(l)})[h_n]) = (\mathbb{P}_{nl} - \mathcal{P}_l)\varsigma(\Lambda, \hat{F}_l; X^{(l)})[h_n] + \mathcal{P}_l\varsigma(\Lambda, \hat{F}_l; X^{(l)})[h_n - h] + \mathcal{P}_l(\varsigma(\Lambda, \hat{F}_l; X^{(l)})[h] - \varsigma(\Lambda, F_l; X^{(l)})[h]) + \mathcal{P}_l\varsigma(\Lambda, F_l; X^{(l)})[h].
\]

We next show that

\[
\mathbb{P}_{nl}(\varsigma(\hat{\Lambda}_1, \hat{F}_l; X^{(l)})[h_n]) = \mathcal{P}_l\varsigma(\hat{\Lambda}_1, F_l; X^{(l)})[h] + o_p(n^{-1/2}) \tag{S7.7}
\]

and

\[
\mathbb{P}_{nl}(\varsigma(\hat{\Lambda}_2, \hat{F}_l; X^{(l)})[h_n]) = \mathcal{P}_l\varsigma(\hat{\Lambda}_2, F_l; X^{(l)})[h] + o_p(n^{-1/2}). \tag{S7.8}
\]

According to the proof of Theorem 4, we have

\[
\sqrt{n_l}(\mathbb{P}_{nl} - \mathcal{P}_l)\varsigma(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[h_n] = o_p(1) \text{ and } \sqrt{n_l}\mathcal{P}_l\varsigma(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[h_n - h] = o_p(1).
\]

Moreover, according to Lemma 1, this implies that

\[
\mathfrak{P}_l \left( \varsigma(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[h] - \varsigma(\hat{\Lambda}_l, F_l; X^{(l)})[h] \right) \\
\lesssim \left( \mathcal{P}_l \left[ \sum_{j=1}^{K(l)} \left| \Delta\hat{\Lambda}_{l,j}(U^{(l)}) - \Delta\Lambda_0,j(U^{(l)}) \right| \right] \\
+ \mathcal{P}_l \left[ \sum_{j=1}^{K(l)} \left| \Delta\hat{\Lambda}_{l,j}'(U^{(l)}) - \Delta\Lambda_0,j'(U^{(l)}) \right| \right] \right) \| \hat{F}_l - F_l \|_{\infty}
\lesssim (|\hat{\Lambda}_l - \Lambda_0|_{L_2(\mu_l)} + |\hat{\Lambda}_l' - \Lambda_0'|_{L_2(\mu_l)}) \| \hat{F}_l - F_l \|_{\infty}.
\]

Applying Lemma 3.5 and Corollary 2.1 of [He and Shi (1994)], we have \( |\hat{\Lambda}_l' - \Lambda_0'|_{L_2(\mu_l)} = O_p(n^{-(r-1)/(1+2r)}) = o_p(1). \) This gives \( \mathcal{P}_l(\varsigma(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[h] - \varsigma(\hat{\Lambda}_l, F_l; X^{(l)})[h]) = o_p(n^{-1/2}). \) Thus, (S7.7) and (S7.8) hold.
Moreover, we have

\[
P_l \left[ \sum_{j=1}^{K(l)} \triangle h_j(U^{(l)}) \right] = \int E \left[ \sum_{j=1}^{K(l)} \triangle h_j(u) \right| U^{(l)} = u] dF_l(u)
\]

\[
= \int E \left[ \sum_{j=1}^{K(r)} \triangle h_j(u) \right| U^{(r)} = u] f_r(u) dF_r(u) = P_r \left[ \sum_{j=1}^{K(r)} \triangle h_j(U^{(r)}) \right] f_r(U^{(r)})
\]

(S7.10)

Hence, (S7.7), (S7.8) and (S7.10) yield

\[
\tilde{U}_n = \sum_{l=1}^{2} \frac{m_l}{\sqrt{n}} P_l \left( \varsigma(\hat{\Lambda}_1, F_l; X^{(l)})[h] - \varsigma(\hat{\Lambda}_2, F_l; X^{(l)})[h] \right) + o_p(1)
\]

\[
\Rightarrow \sqrt{n} \sum_{l=1}^{2} \frac{m_l}{\sqrt{n}} P_l \varsigma(\hat{\Lambda}_1, F_l; X^{(l)})[w_1] - \sqrt{n} \sum_{l=1}^{2} \frac{m_l}{\sqrt{n}} P_l \varsigma(\hat{\Lambda}_2, F_l; X^{(l)})[w_2] + o_p(1)
\]

\[
\Rightarrow \frac{\sqrt{n_1}}{\sqrt{p_1}} P_1 \varsigma(\hat{\Lambda}_1, F_1; X^{(1)})[w_1] - \frac{\sqrt{n_2}}{\sqrt{p_2}} P_2 \varsigma(\hat{\Lambda}_2, F_2; X^{(2)})[w_2] + o_p(1).
\]

Set \( G_l(s) = P_l 1_{\{Y^{(l)} \geq s}\} \) and

\[
\varphi_l(\Lambda, F; X; \bar{Y}, \hat{\Delta})[w] = 1_{\{\Delta = 1\}} \frac{\tilde{\varphi}_{\Lambda,F}(\bar{Y}, \infty; X)[w]}{G_l(\bar{Y})} - \int_0^{\bar{Y}} \frac{\tilde{\varphi}_{\Lambda,F}(u, \infty; X)[w]}{G_l(u)(1 - F(u))} dF(u)
\]

\[
+ \int_0^{\bar{Y}} \frac{\tilde{\varphi}_{\Lambda,F}(u, Y; X)[w]}{G_l(u)(1 - F(u))} dF(u) - 1_{\{\hat{\Delta} = 1, Y > \bar{Y}\}} \frac{\tilde{\varphi}_{\Lambda,F}(\bar{Y}, Y; X)[w]}{G_l(\bar{Y})}.
\]

By Theorem 3 and (S7.9), we can get

\[
\sqrt{m_l} P_l \varsigma(\hat{\Lambda}_l, F_l; X^{(l)})[w_l] = \sqrt{m_l} P_l \varsigma(\hat{\Lambda}_l, \tilde{F}_l; X^{(l)})[w_l] + o_p(1) \sim N(0, \sigma_l^2).
\]

Since \( \hat{\Lambda}_1 \) and \( \hat{\Lambda}_2 \) are independent, it then follows that \( \tilde{U}_n \sim N(0, (\sigma_1^2/p_1 + \sigma_2^2/p_2)) \).
(ii) Set

\[
\varphi_n^*(\Lambda, F; X; \tilde{Y}, \tilde{\Delta})[w] = 1_{\{\tilde{\Delta} = 1\}} \frac{\tilde{\varphi}_{\Lambda, F}(\tilde{Y}, \infty; X)[w]}{G_n^((\tilde{Y})}} - \int_0^{\tilde{Y}} \frac{\tilde{\varphi}_{\Lambda, F}(u, \infty; X)[w]}{G_n^((u)(1 - F(u))} dF(u)
\]

\[+ \int_0^{\tilde{Y}} \frac{\tilde{\varphi}_{\Lambda, F}(u, Y; X)[w]}{G_n^((u)(1 - F(u))} dF(u) - 1_{\{\tilde{\Delta} = 1, \tilde{Y} > Y\}} \frac{\tilde{\varphi}_{\Lambda, F}(\tilde{Y}, Y; X)[w]}{G_n^((Y)}
\]

where \(G_n^((s) = \mathbb{P}_n 1_{\{Y(n) \geq s\}} \). Then

\[
\sigma_i^2 - \sigma_i^2 = \mathbb{P}_n \left[ \left\{ \mathbb{P}_n \varphi_n^*(\hat{\Lambda}_n, \hat{F}_n; X^{(n)}; \tilde{Y}^{(n)}, \tilde{\Delta}^{(n)})[w_n^{(n)}] + \psi(\hat{\Lambda}_n, \hat{F}_n; X^{(n)})[w_n^{(n)}] \right\}^2
\]

\[- \left\{ \mathcal{P} \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)})[w_n^{(l)}] + \psi(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[w_n^{(l)}] \right\}^2
\]

\[+ \mathbb{P}_n \left[ \mathcal{P} \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)})[w_n^{(l)}] + \psi(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[w_n^{(l)}] \right]^2
\]

\[- \mathcal{P} \left[ \mathcal{P} \varphi_l(\Lambda_0, F_l; X^{(l)}; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)})[w_l] + \psi(\Lambda_0, F_l; X^{(l)})[w_l] \right]^2.
\]

Note that

\[
\mathbb{P}_n \left[ \left\{ \mathbb{P}_n \varphi_n^*(\hat{\Lambda}_n, \hat{F}_n; X^{(n)}; \tilde{Y}^{(n)}, \tilde{\Delta}^{(n)})[w_n^{(n)}] + \psi(\hat{\Lambda}_n, \hat{F}_n; X^{(n)})[w_n^{(n)}] \right\}^2
\]

\[- \left\{ \mathcal{P} \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)})[w_n^{(l)}] + \psi(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[w_n^{(l)}] \right\}^2
\]

\[= \frac{1}{n} \sum_{i=1}^{n} \left[ \left\{ \mathbb{P}_n \varphi_n^*(\hat{\Lambda}_n, \hat{F}_n; X^{(n)}; \tilde{Y}_i^{(n)}, \tilde{\Delta}_i^{(n)})[w_n^{(n)}] - \mathcal{P} \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \tilde{Y}_i^{(l)}, \tilde{\Delta}_i^{(l)})[w_n^{(l)}] \right\} \right]
\]

\[\times \left\{ \mathbb{P}_n \varphi_n^*(\hat{\Lambda}_n, \hat{F}_n; X^{(n)}; \tilde{Y}_i^{(n)}, \tilde{\Delta}_i^{(n)})[w_n^{(n)}] + \mathcal{P} \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \tilde{Y}_i^{(l)}, \tilde{\Delta}_i^{(l)})[w_n^{(l)}] \right\}
\]

\[+ 2\psi(\hat{\Lambda}_l, \hat{F}_l; X^{(l)})[w_l^{(l)}] \].
For each \( i \), we obtain

\[
\mathbb{P}_n \varphi_n(l)(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] = \mathcal{P}\varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)]
\]

\[
= (\mathbb{P}_n - \mathcal{P}) \varphi_n(l)(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)]
+ \mathcal{P}\left( \varphi_n(l)(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] - \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] \right).
\]

Since \( \sup_{s \in [0, T]} |G_n^{(l)}(s) - G_l(s)| = O_p(n^{-1/2}) \), we get

\[
\mathcal{P}\left( \varphi_n(l)(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] - \varphi_l(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] \right) = o_p(1).
\]

Similar to the proof of Lemma 2, under Conditions (C2), (C4), (C6) and (C10),

\[
\{ \varphi_n(l)(\Lambda, F; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] : \Lambda \in \Phi, \ F \in \mathcal{F}, \ d_2(F, F_l) \leq \delta, \ \Lambda \text{ is uniformly bounded} \}
\]

is Donsker and it is Glivenko-Cantelli. It follows that

\[
(\mathbb{P}_n - \mathcal{P}) \varphi_n(l)(\hat{\Lambda}_l, \hat{F}_l; X^{(l)}; \hat{\Delta}_i^{(l)}; \Delta_i^{(l)})[w_n(l)] = o_p(1)
\]

and the first term of \( \sigma_1^2 - \sigma_2^2 \) is \( o_p(1) \). Denote

\[
v_l(\Lambda, F; \hat{Y}, \hat{\Delta}; X)[w] = \mathcal{P}\varphi_l(\Lambda, F; X^{(l)}; \hat{Y}, \hat{\Delta})[w] + \psi(\Lambda, F; X)[w].
\]

To verify that the second term of \( \sigma_1^2 - \sigma_2^2 \) is \( o_p(1) \), we only need to prove

\[
E[|v^2_1(\Lambda_0, F_l; \hat{Y}^{(l)}, \hat{\Delta}^{(l)}; X^{(l)})[w_n(l)] - v^2_1(\Lambda_0, F_l; \hat{Y}^{(l)}, \hat{\Delta}^{(l)}; X^{(l)})[w_l]|] = o(1).
\]
By the definition of $w_n^{(l)}(u, t)$, we obtain

\[
\left| \left\{ w_n^{(l)}(u - t_1) - w_n^{(l)}(u - t_2) \right\} - \left\{ w_l(u - t_1) - w_l(u - t_2) \right\} \right|
\]

\[
= \left| \left\{ (h_n(u - t_1) - h(u - t_1)) - (h_n(u - t_2) - h(u - t_2)) \right\} \left( \frac{n_l}{n} + \frac{n_r \hat{f}_r(u)}{f_l(u)} \right) \right|
\]

\[
+ \left( \frac{n_l}{n} + \frac{n_r \hat{f}_r(u)}{f_l(u)} - p_r \frac{f_r(u)}{f_l(u)} - p_l \right) \left| h(u - t_1) - h(u - t_2) \right|
\]

\[
\leq \left| (h_n(u - t_1) - h(u - t_1)) - (h_n(u - t_2) - h(u - t_2)) \right| \left( \frac{n_l}{n} + \frac{n_r \hat{f}_r(u)}{f_l(u)} \right)
\]

\[
+ \left( \frac{n_l}{n} + \frac{n_r \hat{f}_r(u)}{f_l(u)} - p_r \frac{f_r(u)}{f_l(u)} - p_l \right) \left| h(u - t_1) - h(u - t_2) \right|
\]

\[
\leq c \left\{ |h_n(u - t_1) - h(u - t_1)| + |h_n(u - t_2) - h(u - t_2)| + \left| \frac{\hat{f}_r(u)}{f_l(u)} - \frac{f_r(u)}{f_l(u)} \right| \right\}
\]

(S7.11)

with probability 1 for some constant $c$, where $l, r = 1, 2$ and $l \neq r$. According to Theorem 2.2 of Földes, Réjtő, and Winter (1981), we have

\[
\sup_u \left| \frac{\hat{f}_r(u)}{f_l(u)} - \frac{f_r(u)}{f_l(u)} \right| \xrightarrow{a.s.} 0.
\]

(S7.12)

Furthermore, by the definition of $\varphi_l(\Lambda, F; X; Y, \Delta)[w]$, $\varphi_{\Lambda, F}(u; X)[w]$ and $\tilde{\varphi}_{\Lambda, F}(u, a; X)[w]$, after some calculations, we have

\[
|\nu_l(\Lambda_0, F_l; \tilde{Y}^{(l)}, \tilde{\Delta}^{(l)}, X^{(l)})[w_n^{(l)} + w_l]| \lesssim N^{(l)}(T_{K^{(l)}}) + \Lambda_0(0)
\]

(S7.13)
and

\[ |v_l(\Lambda_0, F_l; \bar{Y}^{(l)}, \bar{\Delta}^{(l)}, X^{(l)})[w_n^{(l)} - w_l]| \]

\[ \lesssim \left( N^{(l)}(T^{(l)}_K) + \Lambda_0(\tau) \right) \sum_{j=1}^{K^{(l)}} \left[ \Delta^{(l)} \left| \Delta w_{n,j}^{(l)}(Y^{(l)}) - \Delta w_{l,j}(Y^{(l)}) \right| \right] \]

\[ + (1 - \Delta^{(l)}) \frac{\int_{Y^{(l)}}^\infty \left| \Delta w_{n,j}^{(l)}(u) - \Delta w_{l,j}(u) \right| dF_l(u)}{1 - F_l(Y^{(l)})} \]  

\[ + |P\varphi_l(\Lambda_0, F_l; X^{(l)}, \bar{Y}^{(l)}, \bar{\Delta}^{(l)})[w_n^{(l)} - w_l]| \]  

with probability 1. Combining (S7.11)–(S7.14) and using the Cauchy-Schwarz inequality, we obtain

\[ E|v_l^2(\Lambda_0, F_l; \bar{Y}^{(l)}, \bar{\Delta}^{(l)}; X^{(l)})[w_n^{(l)}] - v_l^2(\Lambda_0, F_l; \bar{Y}^{(l)}, \bar{\Delta}^{(l)}; X^{(l)})[w_l]| \]

\[ \lesssim E \left[ \left( N^{(l)}(T^{(l)}_K) + \Lambda_0(\tau) \right) \sum_{j=1}^{K^{(l)}} \left\{ \Delta^{(l)} \left| \Delta w_{n,j}^{(l)}(Y^{(l)}) - \Delta w_{l,j}(Y^{(l)}) \right| \right\} \right] \]

\[ + (1 - \Delta^{(l)}) \frac{\int_{Y^{(l)}}^\infty \left| \Delta w_{n,j}^{(l)}(u) - \Delta w_{l,j}(u) \right| dF_l(u)}{1 - F_l(Y^{(l)})} \]  

\[ + E \left[ \left( N^{(l)}(\tau) + \Lambda_0(\tau) \right) |P\varphi_l(\Lambda_0, F_l; X^{(l)}, \bar{Y}^{(l)}, \bar{\Delta}^{(l)})[w_n^{(l)} - w_l]| \right] \]

\[ \lesssim E \left[ |P\varphi_l(\Lambda_0, F_l; X^{(l)}, \bar{Y}^{(l)}, \bar{\Delta}^{(l)})[w_n^{(l)} - w_l]| \right] + ||h^{(l)}_n - h_l||_{L^2(\mu_l)} \]

\[ \lesssim ||h^{(l)}_n - h_l||_{L^2(\mu_l)} \to 0. \]
S8 Two-Sample Test with Crossed Distribution Functions of Terminal Event Times

At this section, we conducted the simulation studies for two groups with crossed distribution functions of the terminal event times. We generated $U_i^{(1)}$ and $U_i^{(2)}$ from $6 + \exp(1)$ and $6 + \text{weibull}(1,2)$, respectively. Then the distribution functions of the terminal event times in the two groups were $F_1(u) = 1 - \exp(-u)$ and $F_2(u) = 1 - \exp(-u^2)$, which crossed over on $u \in [0, \tau]$. The censoring times $C_i^{(1)}$ and $C_i^{(2)}$ were from $6 + \kappa_1 \exp(2)$ and $6 + \kappa_2 \exp(2)$, where $(\kappa_1, \kappa_2) = (0.748, 0.801)$ and $(0.334, 0.421)$ to result in 40% and 60% censoring rate, respectively. We still considered two cases with sample size $n_1 = n_2 = 100, 150$ or 200. For Case 1, we set $\Lambda_1(s) = s$ and $\Lambda_2(s) = \beta s$ with $\beta = 1, 1.1, 1.2, 1.3$. For Case 2, we considered partly overlapped two functions in the following two scenarios by

Scenario 1: $\Lambda_1(s) = s$ and $\Lambda_2(s) = 1_{\{s \geq 10 - \beta\}}s + 1_{\{s < 10 - \beta\}} \frac{10s}{s + \beta}$,

Scenario 2: $\Lambda_1(s) = s$ and $\Lambda_2(s) = 1_{\{s \geq 10 - \beta\}} \frac{10s}{s + \beta} + 1_{\{s < 10 - \beta\}} s$

with $\beta = 3, 4, 5, 6$.

Figure[1] displays the true mean functions in Case 2 for the two scenarios with different values of $\beta$, which reveals that $\Lambda_1$ and $\Lambda_2$ overlap on $[\tau - \beta, \tau]$ in Scenario 1 and on $[0, \tau - \beta]$ in Scenario 2. We also considered the three
weight processes $h_n^{(j)}(t), j = 1, \cdots, 3$ in Subsection 5.2, and the simulation results were evaluated with 1000 replications.
Figure 2: Q-Q plots for $n_1=n_2=200$, $\beta=1$, and censoring rate = 60% when the distribution functions of the terminal event time for the two groups cross over.

We also used the quantile plots against the standard normal distribution to justify the normality of the test statistics. Figure 2 shows the plot for
### S8. Two-Sample Test with Crossed Distribution Functions of Terminal Event Times

Table 1: Simulation results of the two-sample tests with different weights for Case 1 when the distribution functions of the terminal event time for the two groups cross over.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Censoring rate 40%</th>
<th>Censoring rate 60%</th>
<th>$n_1 = n_2 = 100$</th>
<th>$n_1 = n_2 = 150$</th>
<th>$n_1 = n_2 = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_n(h_n^{(1)})$</td>
<td>$T_n(h_n^{(2)})$</td>
<td>$T_n(h_n^{(3)})$</td>
<td>$T_n(h_n^{(1)})$</td>
<td>$T_n(h_n^{(2)})$</td>
</tr>
<tr>
<td>1</td>
<td>0.057</td>
<td>0.057</td>
<td>0.059</td>
<td>0.066</td>
<td>0.060</td>
</tr>
<tr>
<td>1.1</td>
<td>0.268</td>
<td>0.235</td>
<td>0.267</td>
<td>0.253</td>
<td>0.249</td>
</tr>
<tr>
<td>1.2</td>
<td>0.713</td>
<td>0.683</td>
<td>0.710</td>
<td>0.730</td>
<td>0.690</td>
</tr>
<tr>
<td>1.3</td>
<td>0.956</td>
<td>0.944</td>
<td>0.960</td>
<td>0.953</td>
<td>0.948</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n_1 = n_2 = 150$</th>
<th>$n_1 = n_2 = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.059</td>
</tr>
<tr>
<td>1.1</td>
<td>0.368</td>
</tr>
<tr>
<td>1.2</td>
<td>0.869</td>
</tr>
<tr>
<td>1.3</td>
<td>0.998</td>
</tr>
</tbody>
</table>

Table 2: Simulation results of the two-sample tests with different weights for Scenario 1 in Case 2 when the distribution functions of the terminal event time for the two groups cross over.

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>Censoring rate 40%</th>
<th>Censoring rate 60%</th>
<th>$n_1 = n_2 = 100$</th>
<th>$n_1 = n_2 = 150$</th>
<th>$n_1 = n_2 = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_n(h_n^{(1)})$</td>
<td>$T_n(h_n^{(2)})$</td>
<td>$T_n(h_n^{(3)})$</td>
<td>$T_n(h_n^{(1)})$</td>
<td>$T_n(h_n^{(2)})$</td>
</tr>
<tr>
<td>3</td>
<td>0.970</td>
<td>1.000</td>
<td>0.828</td>
<td>0.976</td>
<td>0.999</td>
</tr>
<tr>
<td>4</td>
<td>0.817</td>
<td>0.957</td>
<td>0.693</td>
<td>0.856</td>
<td>0.966</td>
</tr>
<tr>
<td>5</td>
<td>0.446</td>
<td>0.523</td>
<td>0.402</td>
<td>0.491</td>
<td>0.562</td>
</tr>
<tr>
<td>6</td>
<td>0.188</td>
<td>0.153</td>
<td>0.177</td>
<td>0.213</td>
<td>0.179</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n_1 = n_2 = 150$</th>
<th>$n_1 = n_2 = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.994</td>
</tr>
<tr>
<td>4</td>
<td>0.925</td>
</tr>
<tr>
<td>5</td>
<td>0.627</td>
</tr>
<tr>
<td>6</td>
<td>0.248</td>
</tr>
</tbody>
</table>
Table 3: Simulation results of the two-sample tests with different weights for Scenario 2 in Case 2 when the distribution functions of the terminal event time for the two groups cross over.

<table>
<thead>
<tr>
<th>β</th>
<th>Censoring rate 40%</th>
<th></th>
<th>Censoring rate 60%</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_n(h_n^{(1)})$</td>
<td>$T_n(h_n^{(2)})$</td>
<td>$T_n(h_n^{(3)})$</td>
<td>$T_n(h_n^{(1)})$</td>
</tr>
<tr>
<td>3</td>
<td>0.073</td>
<td>0.057</td>
<td>0.078</td>
<td>0.081</td>
</tr>
<tr>
<td>4</td>
<td>0.309</td>
<td>0.390</td>
<td>0.274</td>
<td>0.295</td>
</tr>
<tr>
<td>5</td>
<td>0.829</td>
<td>0.941</td>
<td>0.750</td>
<td>0.796</td>
</tr>
<tr>
<td>6</td>
<td>0.986</td>
<td>1.000</td>
<td>0.973</td>
<td>0.987</td>
</tr>
</tbody>
</table>

$n_1 = n_2 = 100$, $\beta = 1$ and censoring rate 60% with the three weight processes in Case 1, and the plots for other situations are similar and not presented here. Tables 1-3 summarize the sizes and powers of the test with the three weight processes at significance level 0.05 for Cases 1 and 2, respectively. For Case 1, the sizes and powers yield similar conclusions to those in Subsection 5.2 and 5.3 even for a high censoring rate, say 60%. For Case 2, the simulation results reveal that the less the overlap, the higher power the test has, regardless of where the overlap locates. The test power increases as the sample size increases for all values of $\beta$. For the three weight processes, the powers with $h_n^{(3)}$ are the smallest when $\beta = 3, 4,$ and
5 in Scenario 1; while the powers with $h_{m}^{(2)}$ are the largest when $\beta = 4$, 5, and 6 in Scenario 2.

**S9  Real Data Analysis for Data from 2002**

In this section, we treated the survey in 2002 as the baseline survey and analyzed all the 967 individuals who were still alive and contacted in the survey in 2005 among the 4362 elders. We still divided them into three groups according to their living area, which yields 346 elders living in urban, 319 elders living in rural, and 302 elders living in both areas. The censoring rate was 30.61%.

Figure 3 also justifies that urban-area residents may experience more serious illnesses than rural-area residents and dual-area residents. The patterns of the estimated mean functions for the four groups are fundamentally the same as those shown in Figure 5 in Section 6. The log-rank test revealed that the survival functions of death for any two groups were not significantly different, where $p$-value$= 0.2$ for the test between urban-area residents and rural-area residents, $p$-value$= 0.6$ for the test between urban-area residents and dual-area residents, and $p$-value$= 0.4$ for the test between rural-area residents and dual-area residents. Figure 4 shows the KM estimates for the survival functions of death. Therefore, we used the statistic in Theorem
Figure 3: The estimates of the mean function for the CLHLS data from 2002.

4 to test the null hypotheses $H_0^{(1)}$: $\Lambda_U(s) = \Lambda_R(s)$, $H_0^{(2)}$: $\Lambda_U(s) = \Lambda_D(s)$, and $H_0^{(3)}$: $\Lambda_R(s) = \Lambda_D(s)$. Table 4 presents the test results, which yield the same conclusions as those for the whole data from 1998. However, since the sample size (only 967 individuals had at least one observation after 2002) is largely reduced, the powers for detecting the group differences as shown in Table 4 are significantly lost as expected.
Figure 4: The KM estimates of the survival function for the CLHLS data from 2002.

Bibliography


He, X., and Shi, P. (1994). Convergence rate of B-spline estimators of
Table 4: Two-sample test results with the three weights for the CLHLS data from 2002.

<table>
<thead>
<tr>
<th>$H_0^{(1)}$: $\Lambda_U(s) = \Lambda_R(s)$</th>
<th>$H_0^{(2)}$: $\Lambda_U(s) = \Lambda_D(s)$</th>
<th>$H_0^{(3)}$: $\Lambda_R(s) = \Lambda_D(s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_n$</td>
<td>209.299</td>
<td>3.239</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>97.905</td>
<td>1.669</td>
</tr>
<tr>
<td>$T_n$</td>
<td>2.138</td>
<td>1.941</td>
</tr>
<tr>
<td>p-value</td>
<td>0.033**</td>
<td>0.052*</td>
</tr>
<tr>
<td>$U_n$</td>
<td>200.766</td>
<td>2.494</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>100.329</td>
<td>1.678</td>
</tr>
<tr>
<td>$T_n$</td>
<td>2.001</td>
<td>1.486</td>
</tr>
<tr>
<td>p-value</td>
<td>0.045**</td>
<td>0.137</td>
</tr>
<tr>
<td>$U_n$</td>
<td>-15.854</td>
<td>-0.735</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>89.800</td>
<td>1.562</td>
</tr>
<tr>
<td>$T_n$</td>
<td>-0.177</td>
<td>-0.470</td>
</tr>
<tr>
<td>p-value</td>
<td>0.860</td>
<td>0.638</td>
</tr>
</tbody>
</table>

$T_n(h)$: the observed value of the test statistic with different weight functions;

* represents significance level of 0.1; ** represents significance level of 0.05.


