Web-based Supplementary Materials for “Estimating subject-specific dependent competing risk profile with censored event time observations”

by Li, Tian and Wei
Appendix A: The Asymptotical Properties of $\hat{\eta}(v)$

Recalling that $\hat{\beta}$ is a solution from the estimating equation
\[
n^{-1} \sum_{i=1}^{n} \frac{w_i}{G(T_i \wedge t_0)} X_i \{ I(T_i \leq t_0, \epsilon = 1) - g(\beta' X_i) \} = 0,
\]
it follows from the similar arguments used in Tian et al. (2007) that $\hat{\beta}$ converges to a deterministic limit $\beta_0$ and
\[
\hat{\beta} - \beta_0 = n^{-1} \xi_i + o_p(n^{-1})
\]
where $\beta_0$ is the solution of $r(\beta) = 0$,
\[
\xi_i = [E\{\hat{g}(\beta_0'X_i)X_i^2\}]^{-1} \left( X_i \{ I(\hat{T}_i \leq t_0, \epsilon = 1) - g(\beta' X_i) \} - \int_0^{t_0} \mathbb{K}(X_i I(\hat{T}_i \leq t_0, \epsilon = 1) - g(\beta' X_i), u) \frac{dM_i^C(u)}{G(T_i \wedge t_0)} \right),
\]
$\mathbb{K}(W, u) = W - E\{WI(\hat{T} \wedge t_0 \geq u)\}/\text{pr}(\hat{T} \wedge t_0 \geq u)$ for any random vector $W$ and $M_i^C(u)$ is the martingale process associated with the censoring time $C_i$. Let $V_i = \beta_0'X_i$ and $\hat{V}_i = \hat{\beta}'X_i$.

With slightly abuse of notation, we let $\{\hat{a}(v)', \hat{b}(v)\}'$ be the maximizer of
\[
\sum_{i=1}^{n} \frac{w_iK_h(\hat{V}_i - v)}{G(T_i \wedge t_0)} \left( \sum_{k=1}^{K-1} Y_{ik}\{a_k + b_k(\hat{V}_i - v)\} - \log \left[ 1 + \sum_{k=1}^{K-1} \exp\{a_k + b_k(\hat{V}_i - v)\} \right] \right),
\]
and then it is the solution to the estimating equation
\[
\hat{S}(a, b; v) = \{ \hat{S}_1'(a, b; v), \ldots, \hat{S}_{K-1}'(a, b; v) \}' = 0
\]
where
\[
\hat{S}_k(a, b; v) = n^{-1} \sum_{i=1}^{n} \frac{w_iK_h(\hat{V}_i - v)}{G(T_i \wedge t_0)} \left( \frac{1}{\hat{V}_i - v} \right) \left\{ Y_{ik} - \frac{\exp\{a_k + b_k(\hat{V}_i - v)\}}{1 + \sum_{k=1}^{K-1} \exp\{a_k + b_k(\hat{V}_i - v)\}} \right\}.
\]

To study the asymptotical properties of $\hat{a}(v)$, we let $\hat{\Delta}_a(v) = \{ \hat{a}_1(v) - m_1(v), \ldots, \hat{a}_{K-1}(v) - m_{K-1}(v) \}'$ and $\hat{\Delta}_b(v) = h\{ \hat{b}_1(v) - \hat{m}_1(v), \ldots, \hat{b}_{K-1}(v) - \hat{m}_{K-1}(v) \}$, where $m_j(v) = \log\{\eta_j(v)/\eta_{K}(v)\}$ and $\hat{m}_j(v) = dm_j(v)/dv, j = 1, \ldots, K - 1$. Therefore, $\{\hat{\Delta}_a(v)', \hat{\Delta}_b(v)\}'$ is the solution to the estimating equation
\[
\hat{Q}(\Delta_a, \Delta_b; v) = \{ \hat{Q}_1'(\Delta_a, \Delta_b; v), \ldots, \hat{Q}_{K-1}'(\Delta_a, \Delta_b; v) \}' = 0
\]
where \( \hat{Q}_k(\Delta_a, \Delta_b; v) \) is
\[
n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{V}_i - v)}{G(T_i \wedge t_0)} \left( \frac{1}{(\hat{V}_i - v)/h} \right) \left\{ Y_{ik} - \frac{e^{\Delta_{a_k} + \Delta_{b_k}(\hat{V}_i - v)/h + \bar{m}_k(\hat{V}_i, v)}}{1 + \sum_{k=1}^{K-1} e^{\Delta_{a_k} + \Delta_{b_k}(\hat{V}_i - v)/h + \bar{m}_k(\hat{V}_i, v)}} \right\},
\]
\[
\Delta_a = (\Delta_{a_1}, \cdots , \Delta_{a_{K-1}})' , \quad \Delta_b = (\Delta_{b_1}, \cdots , \Delta_{b_{K-1}})' \text{ and } \bar{m}_k(u, v) = m_k(u) + m_k(v)(u - v).
\]
Following the similar arguments used in Cai et al. (2008), one may show that the changes in \( \hat{Q}_k(\Delta_a, \Delta_b; v) \) by replacing \( \hat{G}(\cdot) \) and \( \hat{V}_i \) by \( G(\cdot) \) and \( V_i \), respectively, are asymptotically negligible. Let \( Q_k(\Delta_a, \Delta_b; h, v) \) be
\[
n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{G(T_i \wedge t_0)} \left( \frac{1}{(V_i - v)/h} \right) \left\{ Y_{ik} - \frac{e^{\Delta_{a_k} + \Delta_{b_k}(V_i - v)/h + \bar{m}_k(V_i, v)}}{1 + \sum_{k=1}^{K-1} e^{\Delta_{a_k} + \Delta_{b_k}(V_i - v)/h + \bar{m}_k(V_i, v)}} \right\},
\]
and write difference \( \hat{Q}_k(\Delta_a, \Delta_b; v) - Q_k(\Delta_a, \Delta_b; v) \) as
\[
n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{V}_i - v)}{G(T_i \wedge t_0) G(T_i \wedge t_0)} \left( \frac{1}{(\hat{V}_i - v)/h} \right) \left\{ Y_{ik} - \frac{e^{\Delta_{a_k} + \Delta_{b_k}(\hat{V}_i - v)/h + \bar{m}_k(\hat{V}_i, v)}}{1 + \sum_{k=1}^{K-1} e^{\Delta_{a_k} + \Delta_{b_k}(\hat{V}_i - v)/h + \bar{m}_k(\hat{V}_i, v)}} \right\} I(s < \beta' X) \left( \frac{1}{(\beta' X - v)/h} \right) \left\{ Y_k - \frac{e^{\Delta_{a_k} + \Delta_{b_k}(\beta' X - v)/h + \bar{m}_k(\beta' X, v)}}{1 + \sum_{k=1}^{K-1} e^{\Delta_{a_k} + \Delta_{b_k}(\beta' X - v)/h + \bar{m}_k(\beta' X, v)}} \right\}
\]
\[
- I(s < \beta'_0 X) \left( \frac{1}{(\beta'_0 X - v)/h} \right) \left\{ Y_k - \frac{e^{\Delta_{a_k} + \Delta_{b_k}(\beta'_0 X - v)/h + \bar{m}_k(\beta'_0 X, v)}}{1 + \sum_{k=1}^{K-1} e^{\Delta_{a_k} + \Delta_{b_k}(\beta'_0 X - v)/h + \bar{m}_k(\beta'_0 X, v)}} \right\}
\]
\[
\times \frac{w}{G(T \wedge t_0)},
\]
which is bounded by
\[
\sup |\hat{G}(t) - G(t)| \times n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{V}_i - v)}{G(T_i \wedge t_0) G(T_i \wedge t_0)} \left( \frac{1}{(\hat{V}_i - v)/h} \right) \left\{ Y_{ik} - \frac{\exp(\Delta_{a_k} + \Delta_{b_k}(\hat{V}_i - v)/h + \bar{m}_k(\hat{V}_i, v)) \frac{1}{1 + \sum_{k=1}^{K-1} \exp(\Delta_{a_k} + \Delta_{b_k}(\hat{V}_i - v)/h + \bar{m}_k(\hat{V}_i, v))} \right\} + O_P\left(h^{-1}n^{-1/2}\right) \times \frac{w}{G(T \wedge t_0)} + O_P\left(h^2 + n^{-1/2}\right) = O_P\left(n^{-1/2} + (nh^2)^{-3/4} + (nh)^{-1}\right),
\]
for some small \( \delta_0 > 0 \), where \( P_n \) and \( P \) are the expectation operator with respect to the
uniformly converges to distributed independent random functions, it follows from the standard arguments that it respectively, and $G_n = n^{1/2}(P_n - P)$. Furthermore, since
\[ Q(\Delta_a, \Delta_b; v) = \{Q_1(\Delta_a, \Delta_b; v), \cdots, Q_{K-1}(\Delta_a, \Delta_b; v)\}' \]
can be written as sum of $n$ identically distributed independent random functions, it follows from the standard arguments that it uniformly converges to $q(\Delta_a, \Delta_b; v) = \{q_1(\Delta_a, \Delta_b; v), \cdots, q_{K-1}(\Delta_a, \Delta_b; v)\}'$, where
\[
q_k(\Delta_a, \Delta_b; v) = \left( g_0(v) \int K(x) \left[ \eta_k(v) - \frac{\exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}} \right] dx \right) \left( g_0(v) \int x K(x) \left[ \eta_k(v) - \frac{\exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}} \right] dx \right),
\]
and $g_0(\cdot)$ is the density function of the random variable $\beta_0 \mathbf{X}$. Since $(\Delta_a', \Delta_b') = (0', 0')$ is the unique solution of $q(\Delta_a, \Delta_b; v) = 0$. $\hat{\Delta}_a(v)$ and $\hat{\Delta}_b(v)$ converge to zero uniformly in $v$, assuming that the "slope" matrix of $q(\Delta_a, \Delta_b; v)$ is nonsingular. Coupled with the consistency of $\hat{\Delta}_a$ and $\hat{\Delta}_b$, the Taylor series expansion can be used to show that
\[
\hat{\Delta}_a(v) = A(u)^{-1} n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{Y}_i - v)}{G(T_i \wedge t_0)} \left( \begin{array}{c} Y_{i1} - \frac{\exp\{\bar{m}_1(V_i, v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\bar{m}_k(V_i, v)\}} \\ \vdots \\ Y_{i(K-1)} - \frac{\exp\{\bar{m}_{K-1}(V_i, v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\bar{m}_k(V_i, v)\}} \end{array} \right) + o_p\{(nh)^{-1/2}\},
\]
where
\[
A(u) = g_0(v) \left( \begin{array}{cccc} \eta_1(u)\{1 - \eta_1(u)\} & -\eta_1(u)\eta_2(u) & \cdots & -\eta_1(u)\eta_{K-1}(u) \\
-\eta_2(u)\eta_1(u) & \eta_2(u)\{1 - \eta_2(u)\} & \cdots & -\eta_2(u)\eta_{K-1}(u) \\
\vdots & \vdots & \ddots & \vdots \\
-\eta_{K-1}(u)\eta_1(u) & -\eta_{K-1}(u)\eta_2(u) & \cdots & \eta_{K-1}(u)\{1 - \eta_{K-1}(u)\} \end{array} \right).
\]
Again, following the similar arguments for estimating the asymptotical order of (B.2) in Appendix B of Cai et al. (2008), one may show that
\[
n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{Y}_i - v)}{G(T_i \wedge t_0)} \left\{ Y_{iK} - \frac{\exp\{\bar{m}_K(V_i, v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\bar{m}_k(V_i, v)\}} \right\}
\]
can be approximated by
\[
n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{G(T_i \wedge t_0)} \left\{ Y_{iK} - \frac{\exp\{\bar{m}_K(V_i, v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\bar{m}_k(V_i, v)\}} \right\}
\]
uniformly in $v$ up to an order of $o_p\{(nh)^{-1/2}\}$ for $h = n^{-\nu}, \nu \in (1/5, 1/2)$. Noting that the
consistent estimator for $\eta_k(v)$ is

$$\hat{\eta}_k(v) = \frac{\exp\{\hat{a}_k(v)\}}{1 + \sum_{j=1}^{K-1} \exp\{\hat{a}_j(v)\}},$$

by $\delta$-method we have

$$f\{\eta^*(v)\} - f\{\eta(v)\} = D(v)A(u)\Delta_u(v)/g_0(v) + o_P\{(nh)^{-1/2}\}$$

$$= D(v)n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{g_0(v) G(T_i \land t_0)} \zeta_i(v) + o_P\{(nh)^{-1/2}\}$$

where $D(v) = \text{diag}[\hat{f}_1(v), \ldots, \hat{f}_K(v)]$, $\hat{f}(\cdot)$ is the derivative of $f(\cdot)$, and

$$\zeta_i(v) = \begin{pmatrix} Y_{i1} - \frac{\exp\{m_1(V_i, v)\}}{1 + \sum_{k=1}^{K} \exp\{m_k(V_i, v)\}} \\ \vdots \\ Y_{iK-1} - \frac{\exp\{m_{K-1}(V_i, v)\}}{1 + \sum_{k=1}^{K} \exp\{m_k(V_i, v)\}} \end{pmatrix}.$$ 

Therefore by the central limit theorem

$$(nh)^{1/2}[f\{\hat{\eta}(v)\} - f\{\eta(v)\}] \rightarrow N\{0, \Sigma(v)\},$$

in distribution as $n \rightarrow \infty$.

To justify the consistency of the variance-covariance matrix estimator $\hat{\Sigma}(v)$ based on the resampling method, we first note that following the same arguments above, we have

$$f\{\eta^*(v)\} - f\{\eta(v)\} = D(v)n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{g_0(v) G(T_i \land t_0)} \zeta_i(v) B_i + o_P\{(nh)^{-1/2}\},$$

where the probability measure $P^*$ is on the joint product spaces of the random data and $\{B_i\}$. Therefore $(nh)^{1/2}[f\{\eta^*(v)\} - f\{\eta(v)\}]$ is asymptotically equivalent to

$$D(v)n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{g_0(v) G(T_i \land t_0)} \zeta_i(v) (B_i - 1)$$

whose conditional variance is

$$D(v)n^{-1} \sum_{i=1}^{n} h \left\{ \frac{w_i K_h(V_i - v)}{g_0(v) G(T_i \land t_0)} \right\}^2 \zeta_i(v)^2 \odot 2D(v)$$

which converges to $\Sigma(v)$ in probability as $n \rightarrow \infty$. Therefore, we have shown that $P(|\hat{\Sigma}(v) - \Sigma(v)| > \epsilon | \text{data})$ converges to 0 for any $\epsilon > 0$. 
Appendix B: The Justification of the Resampling Methods

To justify the resampling-based variance estimator, note that the variance estimator \( \tilde{\sigma}_k^2(v) \) can be approximated by

\[
\frac{f^2\{\hat{\eta}_k(v)\}}{g_0^2(v)} n^{-1} \sum_{i=1}^{n} E_{B_i} \left[ \frac{w_i K_h(V_i^* - v)}{G(T_i \wedge t_0)} \{Y_{ik} - \frac{e_{\hat{a}_k(v) + b_k(V_i^* - v)}}{1 + \sum_{k=1}^{K-1} e_{\hat{a}_k(v) + b_k(V_i^* - v)}} B_i \} \right]^2
\]

which uniformly converges to \( \sigma_k^2 \), the asymptotical variance of \( n^{1/2} [f\{\hat{\eta}_k(v)\} - f\{\eta_k(v)\}] \), in probability as \( n \to \infty \), where \( E_{B_i} \) is the expectation with respect to the random weights \( \{B_1, \ldots, B_n\} \), which are independent of the observed data. The first approximation follows from the fact that \( |\hat{\beta}^* - \hat{\beta}| + \sup_t |G^*(t) - \hat{G}(t)| \) is in the order of \( O_p(n^{-1/2}) \) and similar arguments used to bound the difference between \( \hat{Q}_k(\Delta_a, \Delta_b; v) \) and \( Q_k(\Delta_a, \Delta_b; v) \).

To justify the proposed procedure for constructing the simultaneous confidence band of \( \eta_k(v), v \in I \), first note that we have already established that uniformly in \( v \),

\[
(nh)^{1/2} [f\{\hat{\eta}_k(v)\} - f\{\eta_k(v)\}]
\]

\[
=(nh)^{1/2} \frac{f\{\eta_k(v)\}}{g_0(v)} \sum_{i=1}^{n} \left[ \frac{w_i K_h(V_i - v)}{G(T_i \wedge t_0)} \left\{ Y_{ik} - \frac{e_{\hat{m}_k(V_i, v)}}{1 + \sum_{k=1}^{K-1} e_{\hat{m}_k(V_i, v)}} \right\} \right] + o_p(n^{-\delta_0})
\]

\[
= (nh)^{-1/2} \sum_{i=1}^{n} K \left( \frac{V_i - v}{h} \right) \xi_{ki} + o_p(n^{-\delta_0})
\]

\[
=h^{-1/2} \int K \left( \frac{x - v}{h} \right) ydZ_{nk}(x, y) + o_p(n^{-\delta_0})
\]

for some \( \delta_0 > 0 \), where

\[
\xi_{ki} = \frac{f\{\eta_k(V_i)\} w_i}{g_0(V_i) G(T_i \wedge t_0)} \left\{ Y_{ik} - \eta_k(V_i) \right\},
\]

\[
Z_{kn}(x, y) = n^{1/2} \left\{ n^{-1} \sum_{i=1}^{n} I(V_i, \xi_{ki}) - F_k(x, y) \right\},
\]

and \( F_k(x, y) \) is the CDF of \( (V, \xi_k) \). From the strong approximation theorem (Tusnady, 1977), one may construct a sequence of standard bivariate Brownian bridge processes \( B_n(x, y) \) such that

\[
\sup_{x, y} \|B_n(M(x, y)) - Z_{nk}(x, y)\| = O_p(n^{-1/2} \{|\log(n)|^2\}),
\]

where \( M(x, y) \) is the Rosenblatt transformation such that \( M(V_i, \xi_i) \) is uniformly distributed.
on the unit square. Note that the similar strong approximation result for empirical process with more than 2 dimensions is not established yet. Therefore, by integration by part,

\[(nh)^{1/2}[f \{ \hat{\eta}_k(v) \} - f \{ \eta_k(v) \}] / \hat{\sigma}_k(v)\]
can be further approximately uniformly by

\[
\mathcal{Y}_{1,n}(v) = \frac{1}{h^{1/2} \sigma_k(v)} \int K \left( \frac{x - v}{h} \right) ydB_n \{ M(x, y) \}.
\]

Let

\[
\mathcal{Y}_{2,n}(v) = \frac{1}{h^{1/2} \sigma_k(v)} \int K \left( \frac{x - v}{h} \right) ydW_n \{ M(x, y) \},
\]

\[
\mathcal{Y}_{3,n}(v) = \frac{1}{h^{1/2} \sigma_k(v)} \int \sigma_k(x)K \left( \frac{x - v}{h} \right) dW(x)
\]

and \(W_n(\cdot, \cdot)\) be a sequence of bivariate Wiener processes satisfying that

\[
\mathbb{P} \{ -2 \log(h) \leq s \} = \exp(-2e^{-s}) + o(1),
\]

as \(n \to \infty\), where

\[
d_h = \left\{ -2 \log(h) \right\}^{1/2} + \frac{1}{\left\{ -2 \log(h) \right\}^{1/2}} \log \left\{ \int K(t)^2 dt / 4\pi \right\}.
\]

Unlike the supremum value of tight processes, \(S\) itself does not converge in distribution, since \(d_h \to \infty\) as \(n \to \infty\). In parallel arguments \(S^*\), the resampling counterpart of \(S\), is equivalent to

\[
\sup_{\mathcal{I}} \left| \frac{1}{(nh)^{1/2} \sigma_k(v)} \sum_{i=1}^n K \left( \frac{V_i - v}{h} \right) \hat{\xi}_{ki}B_i \right| + o_p(n^{-\delta_0})
\]

and

\[
\sup_{\mathcal{I}} \left| \frac{1}{(nh)^{1/2} \sigma_k(v)} \int K \left( \frac{x - v}{h} \right) ydW_n^* \{ M^*(x, y) \} \right| + o_p(n^{-\delta_0})
\]
for some $\delta_0 > 0$ where

$$\hat{\xi}_{ki} = \frac{\hat{f}\{\hat{\eta}_k(\hat{V}_i)\}w_i}{\hat{g}_0(\hat{V}_i)G(T_i \wedge t_0)}\{Y_{ik} - \hat{\eta}_k(\hat{V}_i)\}$$

and $W_n^*\{M^*(x, y)\}$ is a sequence of mean zero Gaussian processes, whose covariance function is identical to that of $W_n\{M(x, y)\}$ conditional on the observed data. Let $T^* = \{(-2 \log(h))^{1/2}(S^* - d_h)\}$ and $T = \{(-2 \log(h))^{1/2}(S - d_h)\}$. It follows that

$$|\Pr_B(T^* \leq s) - \Pr(T \leq s)| = o_p(n^{-\delta_0}),$$

which implies that we can use the conditional distribution of $S^*$ to approximate that of $S$, where $\Pr_B$ is conditional on the observed data.