Web-based Supplementary Materials for "Additive Mixed Effect Model for Clustered Failure Time Data" by Jianwen Cai and Donglin Zeng

Web Appendix

Technical conditions for Theorems 1 and 2 To establish the asymptotic distributions of $(\widehat{\boldsymbol{\beta}}, \widehat{\theta})$ and $\widehat{\Lambda}$, we need the following conditions, where any subscript 0 means the true values.

(C.1) For any constant $\boldsymbol{\alpha}$ and a deterministic function $\mu(t)$ satisfying $\boldsymbol{\alpha} \neq 0$ or $\mu(t) \neq 0$, $P(\boldsymbol{\alpha}^T \mathbf{X}_{ij}(t) + \mu(t) = 0, j = 1, ..., n_i) < 1$.

(C.2) The true parameter $H_0(t) = \Lambda_0(t) + G(t; \theta_0)$ is a right-continuous increasing function and satisfies $H_0(\tau) < \infty$ and $P(C_{ij} \geq \tau, \text{for some } j = 1, ..., n_i) > 0$.

(C.3) The true parameter θ_0 belongs to a known bounded set Θ. Moreover,

$$E\left\{ \sum_{j \neq l, j, l=1}^{n_i} \int_0^\tau \int_0^\tau Y_{ij}(t) Y_{il}(s) \nabla_\theta Q(t, s; \theta) dt ds \right\}$$

is non-singular in a neighborhood of θ_0, where $\nabla_\theta Q(t, s; \theta)$ denotes the derivative of $Q(t, s; \theta)$ with respect to θ.

(C.4) The cluster size, n_i, is bounded and independent of $Y_{ij}, \mathbf{X}_{ij}$ and Δ_{ij}. Additionally, the censoring time is assumed to be independent of T_{ij} and ξ_i given $\mathbf{X}_{ij}$.

Conditions (C.1) and (C.2) imply that the matrix $n^{-1} \sum_{i=1}^n \sum_{j=1}^{n_i} \int_0^\tau Y_{ij}(t) \{\mathbf{X}_{ij}(t) - \overline{\mathbf{X}}(t)\}^{\otimes 2} dt$ is positive definite when n is large enough. Thus, $\widehat{\boldsymbol{\beta}}$ is well defined. This also implies that $\boldsymbol{\Sigma}$ is invertible. Otherwise, suppose $\boldsymbol{\alpha}^T \boldsymbol{\Sigma} \boldsymbol{\alpha} = 0$ for some constant vector $\boldsymbol{\alpha}$. Then with probability one, $\boldsymbol{\alpha}^T \{\mathbf{X}_{ij}(t) - \mu(t)\} = 0$ for any t and $j = 1, ..., n_k$. However, from condition (C.3) this implies $\boldsymbol{\alpha} = 0$. Thus, $\boldsymbol{\Sigma}$ is positive definite. Additionally, condition (C.2) gives

$n^{-1}\sum_{i=1}^n\sum_{j=1}^{n_i} Y_{ij}(t)$ is bounded away from zero. On the other hand, condition (C.3) ensures that the estimating equation for θ has a solution which is a consistent estimator for θ_0.

Asymptotic expansion of $\widehat{\theta}$

To obtain the asymptotic distribution for $\widehat{\theta}$, we perform the first-order Taylor expansion at $\theta = \theta_0$ on the left-hand side of (7). It yields

$$\mathbf{P}_n\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{d\widehat{\epsilon}_{ij}(t)d\widehat{\epsilon}_{il}(s) - Q(t,s;\theta_0)dtds\right\}\right]$$
$$-\left[E\left\{\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\nabla_\theta Q(t,s;\theta_0)dtds\right\} + o_p(1)\right](\widehat{\theta}-\theta_0) = 0.$$

Note

$$\begin{aligned}
&\mathbf{P}_n\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{d\widehat{\epsilon}_{ij}(t)d\widehat{\epsilon}_{il}(s) - Q(t,s;\theta_0)dtds\right\}\right]\\
&= (\mathbf{P}_n - \mathbf{P})\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{d\epsilon_{ij0}(t)d\epsilon_{il0}(s) - Q(t,s;\theta_0)dtds\right\}\right]\\
&\quad -E\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{d\widehat{H}(t) - dH_0(t) + \mathbf{X}_{ij}(t)^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)dt\right\}d\epsilon_{il0}(s)\right]\\
&\quad -E\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)d\epsilon_{ij0}(t)\left\{d\widehat{H}(s) - dH_0(s) + \mathbf{X}_{il}(t)^T(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)ds\right\}\right].
\end{aligned}$$

Thus, from equations (A.1) and (A.2), we obtain

$$\widehat{\theta} - \theta_0 = (\mathbf{P}_n - \mathbf{P})S_\theta(\mathbf{O}_i) + o_p(n^{-1/2}),$$

where

$$\begin{aligned}
S_\theta(\mathbf{O}_i) &= \left[E\left\{\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\nabla_\theta Q(t,s;\theta_0)dtds\right\}\right]^{-1}\\
&\quad\times\left\{\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{d\epsilon_{ij0}(t)d\epsilon_{il0}(s) - Q(t,s;\theta_0)dtds\right\}\right.\\
&\qquad -E\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{dS_H(\mathbf{O}_i;t) + \mathbf{X}_{ij}(t)^T S_{\boldsymbol{\beta}}(\mathbf{O}_i)dt\right\}d\epsilon_{il0}(s)\right]\\
&\qquad \left.-E\left[\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)d\epsilon_{ij0}(t)\left\{dS_H(\mathbf{O}_i;s) + \mathbf{X}_{il}(t)^T S_{\boldsymbol{\beta}}(\mathbf{O}_i)ds\right\}\right]\right\}.
\end{aligned}$$

Consistent estimation of asymptotic variance. Clearly, the asymptotic covariance of $(\widehat{\boldsymbol{\beta}}, \widehat{H}, \widehat{\theta})$ is given by the covariance of the corresponding influence function $(\mathbf{S}_{\boldsymbol{\beta}}, S_H, S_\theta)$. Thus, a consistent estimator of the asymptotic covariance can be obtained from the empirical covariance of $(\widehat{\mathbf{S}}_{\boldsymbol{\beta}}, \widehat{S}_H, \widehat{S}_\theta)$, where the latter are the consistent estimators of $(\mathbf{S}_{\boldsymbol{\beta}}, S_H, S_\theta)$. Particularly, we can choose

$$\widehat{S}_{\boldsymbol{\beta}}(\mathbf{O}_i) = \left[n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n_i}\int_0^\tau Y_{ij}(t)\{\mathbf{X}_{ij}(t) - \overline{\mathbf{X}}(t)\}^{\otimes 2}dt\right]^{-1}\sum_{j=1}^{n_i}\int_0^\tau Y_{ij}(t)\{\mathbf{X}_{ij}(t) - \overline{\mathbf{X}}(t)\}d\widehat{\epsilon}_{ij}(t),$$

$$\widehat{S}_H(\mathbf{O}_i; s) = \int_0^s \frac{\sum_{j=1}^{n_i} Y_{ij}(t)d\widehat{\epsilon}_{ij}(t)}{n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n_i} Y_{ij}(t)} - \int_0^s \overline{\mathbf{X}}(t)^T dt \widehat{S}_{\boldsymbol{\beta}}(\mathbf{O}_i),$$

and

$$\begin{aligned}
\widehat{S}_\theta(\mathbf{O}_i) \;=\; & \left\{n^{-1}\sum_{i=1}^{n}\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\nabla_\theta Q(t,s;\widehat{\theta})dtds\right\}^{-1} \\
& \times\left\{\sum_{j\neq l,j,l=1}^{n_i}\int_0^\tau\int_0^\tau Y_{ij}(t)Y_{il}(s)\left\{d\widehat{\epsilon}_{ij}(t)d\widehat{\epsilon}_{il}(s) - Q(t,s;\widehat{\theta})dtds\right\}\right. \\
& \quad - n^{-1}\sum_{k=1}^{n}\left[\sum_{j\neq l,j,l=1}^{n_k}\int_0^\tau\int_0^\tau Y_{kj}(t)Y_{kl}(s)\left\{d\widehat{S}_H(\mathbf{O}_i;t) + \mathbf{X}_{kj}(t)^T\widehat{S}_{\boldsymbol{\beta}}(\mathbf{O}_i)dt\right\}d\widehat{\epsilon}_{kl}(s)\right] \\
& \quad \left. - n^{-1}\sum_{k=1}^{n}\left[\sum_{j\neq l,j,l=1}^{n_k}\int_0^\tau\int_0^\tau Y_{kj}(t)Y_{kl}(s)d\widehat{\epsilon}_{kj}(t)\left\{d\widehat{S}_H(\mathbf{O}_i;s) + \mathbf{X}_{kl}(t)^T\widehat{S}_{\boldsymbol{\beta}}(\mathbf{O}_i)ds\right\}\right]\right\}.
\end{aligned}$$