Web-based Supplementary Materials for

Semiparametric Regression Analysis of Length-Biased Interval-Censored Data

by

Fei Gao* **and Kwun Chuen Gary Chan Chan****

Department of Biostatistics, University of Washington, Seattle, Washington, U.S.A.

**email:* feigao@uw.edu

***email:* kcgchan@u.washington.edu

Web Appendix A

Nonparametric Maximum Likelihood Estimation with Left-continuous $\widehat{\Lambda}$

We consider the nonparametric maximum likelihood estimation approach where the estimator for Λ is a left-continuous function with potential discontinuous points at the ends of the intervals that bracket the failure times. Specifically, we let $\lambda_0, \lambda_1, \ldots, \lambda_k$ be the respective jump sizes such that $\Lambda(t) = \sum_{l=1}^{j} \lambda_l$ for $t \in (t_{j-1}, t_j]$, where $\lambda_0 = 0$. Write $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k)$. We maximize the objective function

$$l_n(\boldsymbol{\beta}, \boldsymbol{\lambda}) \equiv \sum_{i=1}^{n} \left(\log \left[\exp\left\{ -\sum_{t_j \le L_i} \lambda_j \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \right\} - I(R_i < \infty) \exp\left\{ -\sum_{t_j \le R_i} \lambda_j \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \right\} \right] - \log \int_0^\tau \frac{1}{\tau} \exp\left\{ -\sum_{t_{j-1} < a} \lambda_j \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \right\} da \right).$$

Following from the argument in Section 2.2, the number of truncated samples n_i follows a negative binomial distribution with parameter

$$\pi_i = P(T^*_{im} < A^*_{im} | \boldsymbol{Z}_i) = \sum_{j=1}^{k} (1 - t_{j-1}/\tau) \lambda_j \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \exp\left\{ -\sum_{l=1}^{j} \lambda_l \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \right\},$$

and

$$p_{ij} = P(T^*_{im} = t_j | T^*_{im} < A^*_{im}, \boldsymbol{Z}_i) = \frac{(1 - t_{j-1}/\tau) \lambda_j \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \exp\left\{ -\sum_{l=1}^{j} \lambda_l \exp\left(\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_i\right) \right\}}{\pi_i}.$$

A similar EM algorithm can be constructed with $\widehat{E}(n_{ij})$ replaced by

$$\widehat{E}(n_{ij}) = \frac{(1 - t_{j-1}/\tau)\lambda_j \exp\left(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{Z}_i\right) \exp\left\{-\sum_{l=1}^{j} \lambda_l \exp\left(\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{Z}_i\right)\right\}}{1 - \pi_{ij}}.$$

To see the numerical difference between different versions, we analyzed the simulated data sets in the first set of simulation studies (with length-biased assumption) using the proposed methods. The results are summarized in Web Table 1 and the difference to the right-continuous is small, especially for large n.

Web Appendix B

Proof of Lemmas

Proof of Lemma 1. Since $\mathcal{D}_M$ consists of increasing and uniformly bounded functions on $\mathcal{U}$, Lemma 2.2 of van der Geer (2000) implies that for any $\epsilon > 0$, the bracketing number satisfies

$$N_{[]}(\epsilon, \mathcal{D}_M, \|\cdot\|_{L_2}) \lesssim \epsilon^{-1},$$

where $\|\cdot\|_{L_2}$ denote the L_2-norm with respect to the Lebesgue measure on $\mathcal{U}$, and $A \lesssim B$ means that $A \leq cB$ for a positive constant c. For $\epsilon > 0$, we can find $\exp\{O(1/\epsilon)\}$ number of brackets $\{[\Lambda_j^L, \Lambda_j^U]\}$ with $\|\Lambda_j^L - \Lambda_j^U\|_{L_2} \leq \epsilon$ and $|\Lambda_j^L(\tau) - \Lambda_j^U(\tau)| < \epsilon$ to cover $\mathcal{D}_M$. In addition, there are $O(\epsilon^{-p})$ number of brackets $\{[\boldsymbol{\beta}_j^L, \boldsymbol{\beta}_j^U]\}$ covering $\mathcal{B}$, such that two $\|\boldsymbol{\beta}_j^L - \boldsymbol{\beta}_j^U\| \leq \epsilon$. Hence, there are in total $\exp\{O(1/\epsilon)\} \times O(\epsilon^{-p})$ brackets that covers $\mathcal{B} \times \mathcal{D}_M$. For any pair of parameters $(\boldsymbol{\beta}_1, \Lambda_1)$ and $(\boldsymbol{\beta}_2, \Lambda_2)$, there exists some constant c such that

$$\begin{aligned} |m(\boldsymbol{\beta}_1, \Lambda_1) - m(\boldsymbol{\beta}_2, \Lambda_2)| &\leq |m(\boldsymbol{\beta}_1, \Lambda_1) - m(\boldsymbol{\beta}_2, \Lambda_1)| + |m(\boldsymbol{\beta}_2, \Lambda_1) - m(\boldsymbol{\beta}_2, \Lambda_2)| \\ &\leq c\|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\| + c\sum_{m=0}^{M} \Delta_m |\Lambda_1(U_m) - \Lambda_2(U_m)| \\ &\quad + c\int_0^\tau |\Lambda_1(a) - \Lambda_2(a)|\, da. \end{aligned}$$

Therefore,

$$\|m(\boldsymbol{\beta}_1, \Lambda_1) - m(\boldsymbol{\beta}_2, \Lambda_2)\|_{L_2(\mathbb{P})} \leq O\{\|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\| + \|\Lambda_1 - \Lambda_2\|_{L_2} + |\Lambda_1(\tau) - \Lambda_2(\tau)|\} = O(\epsilon).$$

The bracketing number of $\mathcal{M}$ then satisfies

$$N_{[]} \leq \exp\{O(1/\epsilon)\}O(\epsilon^{-p}),$$

such that the entropy integral is finite. The class $\mathcal{M}$ is then $\mathbb{P}$-Donsker.

Proof of Lemma 2. By Theorem 1, $\widehat{\Lambda}$ is consistent for Λ_0. Therefore, there exists a finite constant M such that $\widehat{\Lambda}(\tau) \leq M$. By Lemma 1, $m(\widehat{\boldsymbol{\beta}}, \widehat{\Lambda})$ belongs to a Donsker class with bracketing integral

$$J_{[]}(\delta, \mathcal{M}, L_2(\mathbb{P})) = \int_0^\delta \sqrt{1 + \log N_{[]}(\epsilon, \mathcal{M}, L_2(\mathbb{P}))} \leq O(\delta^{1/2}).$$

In addition, by Lemma 1.3 of van der Geer (2000) and the mean-value theorem,

$$\mathbb{P}\left\{m\left(\widehat{\boldsymbol{\beta}}, \widehat{\Lambda}\right) - m\left(\boldsymbol{\beta}_0, \widetilde{\Lambda}\right)\right\} \lesssim H^2\left\{\left(\widehat{\boldsymbol{\beta}}, \widehat{\Lambda}\right), \left(\boldsymbol{\beta}_0, \widetilde{\Lambda}\right)\right\},$$

where $H(\cdot, \cdot)$ is the Hellinger distance defined as

$$H\left\{(\boldsymbol{\beta}_1, \Lambda_1), (\boldsymbol{\beta}_2, \Lambda_2)\right\} = \left[\int \left\{L(\boldsymbol{\beta}_1, \Lambda_1) - L(\boldsymbol{\beta}_2, \Lambda_2)\right\}^2 d\mu\right]^{1/2},$$

with respect to the dominating measure μ. By Theorem 3.4.1 of van der Vaart and Wellner (1996), there exists r_n with $r_n^2\phi(1/r_n) \sim n^{1/2}$ such that $H\{(\widehat{\boldsymbol{\beta}}, \widehat{\Lambda}), (\boldsymbol{\beta}_0, \widetilde{\Lambda})\} = O_P(1/r_n)$, where

$$\phi_n(\delta) = J_{[]}(\delta, \mathcal{M}, H(\cdot, \cdot))\left\{1 + \frac{J_{[]}(\delta, \mathcal{M}, H(\cdot, \cdot))}{\delta^2/\sqrt{n}}\right\}.$$

In particular, we can choose r_n in the order of $n^{1/3}$ such that $H\{(\widehat{\boldsymbol{\beta}}, \widehat{\Lambda}), (\boldsymbol{\beta}_0, \widetilde{\Lambda})\} = O_P(n^{-1/3})$.

Therefore, there exists finite constants c_1 and c_2 such that

$$\begin{aligned} & O_P(n^{-2/3}) \\ = \ & E\left[\left\{\frac{\sum_{m=0}^M \Delta_m \left[\exp\left\{-\widehat{\Lambda}(U_m)\exp\left(\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\boldsymbol{Z}\right)\right\} - \exp\left\{-\widehat{\Lambda}(U_{m+1})\exp\left(\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\boldsymbol{Z}\right)\right\}\right]}{\int_0^\tau \frac{1}{\tau}\exp\left\{-\widehat{\Lambda}(a)\exp\left(\widehat{\boldsymbol{\beta}}^{\mathrm{T}}\boldsymbol{Z}\right)\right\} da}\right.\right. \\ & \left.\left. - \frac{\sum_{m=0}^M \Delta_m \left[\exp\left\{-\Lambda_0(U_m)\exp\left(\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{Z}\right)\right\} - \exp\left\{-\Lambda_0(U_{m+1})\exp\left(\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{Z}\right)\right\}\right]}{\int_0^\tau \frac{1}{\tau}\exp\left\{-\Lambda_0(a)\exp\left(\boldsymbol{\beta}_0^{\mathrm{T}}\boldsymbol{Z}\right)\right\} da}\right\}^2\right] \end{aligned}$$

$$= c_1 \left\| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \right\|^2 + c_2 E \left(\left[L(\boldsymbol{\beta}_0, \Lambda_0) \left\{ \sum_{m=0}^{M} \Delta_m \int_0^\tau Q(t, U_m, U_{m+1}; \boldsymbol{\beta}_0, \Lambda_0) d(\widehat{\Lambda} - \Lambda_0)(t) \right\} \right]^2 \right),$$

where the last equality follows from the mean-value theorem. We define a norm in $BV(\mathcal{U})$ such that for any $f \in BV(\mathcal{U})$,

$$\|f\|_1 = \left[E \left\{ \sum_{m=0}^{M} f(U_m)^2 \right\} \right]^{1/2}.$$

In addition, we define a seminorm

$$\|f\|_2 = E \left(\left[L(\boldsymbol{\beta}_0, \Lambda_0) \left\{ \sum_{m=0}^{M} \Delta_m \int_0^\tau Q(t, U_m, U_{m+1}; \boldsymbol{\beta}_0, \Lambda_0) df(t) \right\} \right]^2 \right)^{1/2}.$$

Note that if $\|f\|_2 = 0$ for some $f \in BV(\mathcal{U})$, then

$$L(\boldsymbol{\beta}_0, \Lambda_0) \left\{ \sum_{m=0}^{M} \Delta_m \int_0^\tau Q(t, U_m, U_{m+1}; \boldsymbol{\beta}_0, \Lambda_0) df(t) \right\} = 0$$

with probability 1.

For any $m \in \{0, \ldots, M\}$, we sum over all possible $\Delta_{m'}$ with $m' = m, \ldots, M$ to obtain

$$- \int_0^{U_m} df(t) + \int \frac{\int_0^a \exp\left\{-\Lambda_0(a) \exp\left(\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{Z}\right)\right\} da}{\int_0^\tau \exp\left\{-\Lambda_0(a) \exp\left(\boldsymbol{\beta}_0^{\mathrm{T}} \boldsymbol{Z}\right)\right\} da} df(t) = 0.$$

Because m is arbitrary, we can replace U_m with any $t \in \mathcal{U}$. We differentiate both sides with respect to t to obtain $f'(t) = 0$, such that $f(t) = 0$ for $t \in \mathcal{U}$, implying that $\| \cdot \|_2$ is a norm in $BV(\mathcal{U})$.

By the Cauchy-Schwarz inequality, for any $f \in BV(\mathcal{U})$,

$$\begin{aligned} \|f\|_2 &\leq \left(E \left[L(\boldsymbol{\beta}_0, \Lambda_0) \left\{ \sum_{m=0}^{M} \Delta_m \int_0^\tau Q(t, U_m, U_{m+1}; \boldsymbol{\beta}_0, \Lambda_0) dt \right\} \right]^2 E \left\{ \sum_{m=0}^{M} f(U_m)^2 \right\} \right)^{1/2} \\ &\leq c_3 \|f\|_1, \end{aligned}$$

where c_3 is a finite constant. By the bounded inverse theorem in the Banach space, we have $\|f\|_2 \geq c_3' \|f\|_1$ for some constant c_3'. Therefore,

$$O_P(n^{-2/3}) + O\left(\left\| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0 \right\|^2 \right) \geq c_2 {c_3'}^2 E \left[\sum_{m=0}^{M} \left\{ \widehat{\Lambda}(U_m) - \Lambda_0(U_m) \right\}^2 \right].$$

The lemma thus holds.

References

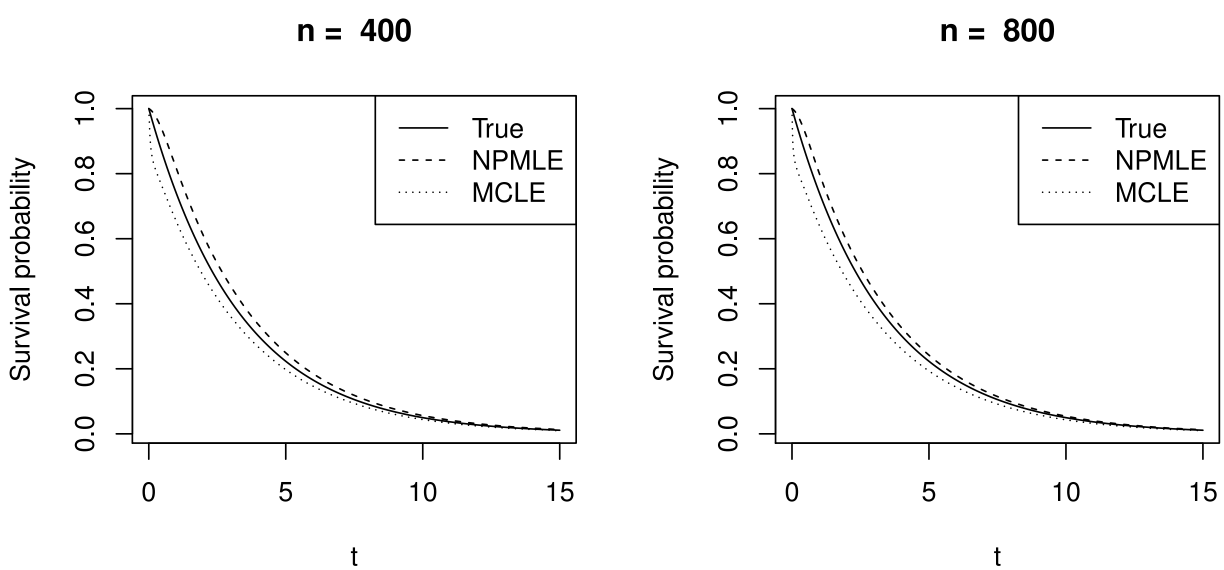
van der Geer, S. A. (2000). *Empirical Processes in M-estimation.* Cambridge: Cambridge University Press.

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

Web Table 1: Summary statistics for the simulation studies with left-continuous $\widehat{\Lambda}$.

		Bias	SE	SEE	RMSE	CP	RMSD
$n = 100$	β_1	0.006	0.168	0.171	0.168	0.957	0.003
	β_2	0.010	0.293	0.316	0.294	0.966	0.006
$n = 200$	β_1	0.003	0.117	0.116	0.117	0.950	0.001
	β_2	0.003	0.205	0.212	0.205	0.957	0.003
$n = 400$	β_1	0.002	0.082	0.081	0.082	0.944	0.001
	β_2	0.001	0.144	0.146	0.144	0.951	0.001

Note: SE, SEE, RMSE, CP, and RMSD are the empirical standard error, mean standard error estimator, root mean squared error, empirical coverage probability of the 95% confidence interval, and root mean squared difference to the right-continuous version, respectively.



Web Figure 1: Estimated baseline survival functions in simulation studies with length-biased assumption. The solid, dashed, and dotted curves pertain to the true value, the nonparametric maximum likelihood estimation and conditional likelihood estimation, respectively.