Supplementary material for 'Generalized Odds Rate Frailty Models for Current Status Data with Informative Censoring'

Yang Xu[a], Shishun Zhao[a], Tao Hu[b*] and Jianguo Sun[c]

a. Center for Applied Statistical Research and College of Mathematics, Jilin University, Changchun, China

b. School of Mathematical Sciences, Capital Normal University, Beijing, China

c. Department of Statistics, University of Missouri, Columbia, MO, USA

Abstract: This supplementary material contains the proofs of the two theorems.

1. Proofs of the Asymptotic Properties

In this appendix, we will sketch the proofs for the asymptotic properties of the proposed estimator $\hat{\boldsymbol{\theta}}_n$. For the proof, we will mainly employ the empirical process theory and some nonparametric techniques. Let $l(\boldsymbol{\theta}_n, \boldsymbol{O})$ denote the log-likelihood function based on a single observation $\boldsymbol{O} = (\tilde{C}, \delta, \Delta, X)$. Define $Pf = \int f(y)dP$ and $P_n f = n^{-1}\sum_{i=1}^{n} f(Y_i)$ to be the expectation of

*Corresponding author. E-mail: hutaomath@foxmail.com

$f(Y)$ under the probability measure P and the expectation of $f(Y)$ under the empirical measure P_n, respectively. Also let K represent some universal positive constant that may vary from place to place.

Proof the Theorem 1.

We prove the consistency by using the idea in Theorem 5.7 of Van der Vaart (2000). Firstly, we need to show the condition $\lim_n \sup_{\boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n} |P_n l(\boldsymbol{\theta}_n, \boldsymbol{O}) - Pl(\boldsymbol{\theta}_n, \boldsymbol{O})| = o_p(1)$ is satisfied, we need to verify that $\mathcal{E}_1 = \{l(\boldsymbol{\theta}_n, \boldsymbol{O}), \boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n\}$ is a Euclidean class (Definition 2.7 in Pakes and Pollard (1989) for its envelope function $\max_{\boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n} l(\boldsymbol{\theta}_n | \boldsymbol{O})$. By (**C2**) and (**C3**) and Lemma 2.14 in Pakes and Pollard (1989), it is easy to see that class $\mathcal{E}_1$ is a Euclidean class. Hence, we have

$$\sup_{\boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n} |P_n l(\boldsymbol{\theta}_n, \boldsymbol{O}) - Pl(\boldsymbol{\theta}_n, \boldsymbol{O})| \to 0, a.s. \tag{A1}$$

Let $M(\boldsymbol{\theta}_n, \boldsymbol{O}) = -l(\boldsymbol{\theta}_n, \boldsymbol{O})$, define $K_\epsilon = \{\boldsymbol{\theta}_n : d(\boldsymbol{\theta}_n, \boldsymbol{\theta}_0) \geq \epsilon, \boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n\}$ and

$$\zeta_{1n} = \sup_{\boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n} |P_n M(\boldsymbol{\theta}_n, \boldsymbol{O}) - PM(\boldsymbol{\theta}_n, \boldsymbol{O})|, \zeta_{2n} = P_n M(\boldsymbol{\theta}_0, \boldsymbol{O}) - PM(\boldsymbol{\theta}_0, \boldsymbol{O}).$$

Then,

$$\begin{aligned} \inf_{K_\epsilon} PM(\boldsymbol{\theta}_n, \boldsymbol{O}) &= \inf_{K_\epsilon} \{PM(\boldsymbol{\theta}_n, \boldsymbol{O}) - P_n M(\boldsymbol{\theta}_n, \boldsymbol{O}) + P_n M(\boldsymbol{\theta}_n, \boldsymbol{O})\} \\ &\leq \zeta_{1n} + \inf_{K_\epsilon} P_n M(\boldsymbol{\theta}_n, \boldsymbol{O}). \end{aligned} \tag{A2}$$

If $\hat{\boldsymbol{\theta}}_n \in K_\epsilon$, we have

$$\inf_{K_\epsilon} P_n M(\boldsymbol{\theta}_n, \boldsymbol{O}) = P_n M(\hat{\boldsymbol{\theta}}_n, \boldsymbol{O}) \leq P_n M(\boldsymbol{\theta}_0, \boldsymbol{O}) = \zeta_{2n} + PM(\boldsymbol{\theta}_0, \boldsymbol{O}). \quad \text{(A3)}$$

By (A2) and (A3), we have

$$\inf_{K_\epsilon} PM(\boldsymbol{\theta}_n, \boldsymbol{O}) \leq \zeta_{1n} + \zeta_{2n} + PM(\boldsymbol{\theta}_0, \boldsymbol{O}) = \zeta_n + PM(\boldsymbol{\theta}_0, \boldsymbol{O})$$

with $\zeta_n = \zeta_{1n} + \zeta_{2n}$. By Condition (**C4**), adopting similar proofs of Theorem 2.1 in Chang et al. (2007) and applying inverse function theorem, we can show the identifiability of the model parameters. Thus, we have

$$\inf_{K_\epsilon} PM(\boldsymbol{\theta}_n, \boldsymbol{O}) - PM(\boldsymbol{\theta}_0, \boldsymbol{O}) = \delta_\epsilon > 0.$$

Then, we get $\zeta_n \geq \delta_\epsilon$ and this gives $\{\hat{\boldsymbol{\theta}}_n \in K_\epsilon\} \subseteq \{\zeta_n \geq \delta_\epsilon\}$. By (A1) and the strong law of large numbers, we have both $\zeta_{1n} = o(1)$ and $\zeta_{2n} = o(1)$ almost surely. Therefore, $\cup_{k=1}^{\infty} \cap_{n=k}^{\infty} \{\hat{\boldsymbol{\theta}}_n \in K_\epsilon\} \subseteq \cup_{k=1}^{\infty} \cap_{n=k}^{\infty} \{\zeta_n \geq \delta_\epsilon\}$, which proves that $d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_0) \to 0$ almost surely.

In the following, we will show the convergence rate of $\hat{\boldsymbol{\theta}}_n$ by using Theorem 3.2.5 of Van der Vaart and Wellner (1996). For any $\epsilon > 0$, define $\mathcal{F}_\epsilon = \{l(\boldsymbol{\theta}_n, \boldsymbol{O}) - l(\boldsymbol{\theta}_{n0}, \boldsymbol{O}) : \boldsymbol{\theta}_n \in \boldsymbol{\Theta}_n, d(\boldsymbol{\theta}_n, \boldsymbol{\theta}_{n0}) \leq \epsilon\}$ with $\boldsymbol{\theta}_{n0} = (\boldsymbol{\beta}_{10}, \boldsymbol{\beta}_{20}, \eta_0, \Lambda_{1n0}, \Lambda_{20})$. Following the calculation in Shen and Wong (1994, p.597), we can establish that for $0 < \rho < \epsilon$, $\log N_{[\,]}(\rho, \mathcal{F}_\epsilon, \|\cdot\|_2) \leq KN \log(\epsilon/\rho)$ with $N = K_n$, where $N_{[\,]}(\varepsilon, \mathcal{F}, d)$ denotes the bracketing number with respect to the metric or semi-metric d of a function class $\mathcal{F}$. Moreover, some algebraic manipulations yield that $\|l(\boldsymbol{\theta}_n, \boldsymbol{O}) - l(\boldsymbol{\theta}_{n0}, \boldsymbol{O})\|_2^2 \le K\epsilon^2$ for any $l(\boldsymbol{\theta}_n, \boldsymbol{O}) - l(\boldsymbol{\theta}_{n0}, \boldsymbol{O}) \in \mathcal{F}_\epsilon$. Under Conditions (**C2**) and (**C3**), it is easy to see that $\mathcal{F}_\epsilon$ is uniformly bounded. Therefore, by Lemma 3.4.2 of Van der Vaart and Wellner (1996), we obtain

$$E_P \left\| n^{1/2} (P_n - P) \right\|_{\mathcal{F}_\epsilon} \le K J_\epsilon (\rho, \mathcal{F}_\epsilon, \|\cdot\|_2) \left\{ 1 + \frac{J_\epsilon (\rho, \mathcal{F}_\epsilon, \|\cdot\|_2)}{\epsilon^2 n^{1/2}} \right\},$$

where $J_\epsilon (\rho, \mathcal{F}_\epsilon, \|\cdot\|_2) = \int_0^\epsilon \{1 + \log N_{[\,]} (\rho, \mathcal{F}_\epsilon, \|\cdot\|_2)\}^{1/2} d\rho \le \int_0^\epsilon \{1 + [KN \log(\epsilon/\rho)]^{1/2}\} d\rho$ $\le KN^{1/2}\epsilon$. This yields $\phi_n(\epsilon) = K \left(N^{1/2}\epsilon + N/n^{1/2} \right)$. It is easy to see that $\phi_n(\epsilon)/\epsilon$ is decreasing in ϵ, and $r_n^2 \phi_n (1/r_n) = r_n N^{1/2} + r_n^2 N/n^{1/2} \le 2n^{1/2}$, where $r_n = N^{-1/2} n^{1/2} = n^{(1-v)/2}, 0 < v < 1/2$. Thus, by applying Theorem 3.2.5 of Van der Vaart and Wellner (1996), we have $n^{(1-v)/2} d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_{n0}) = O_P(1)$. This together with $d(\boldsymbol{\theta}_{n0}, \boldsymbol{\theta}_0) = O_p(n^{-\kappa v})$ using the results of Lemma A1 in Lu et al. (2007), and yields that $d(\hat{\boldsymbol{\theta}}_n, \boldsymbol{\theta}_0) = O_p \left(n^{-(1-v)/2} + n^{-\kappa v} \right)$, which completes the proof.

Proof the Theorem 2.

To prove Theorem 2, we need following notations. Let $\delta_n = n^{-(1-v)/2} + n^{-\kappa v}$ denote the rate of convergence obtained in Theorem 1 and let V denote the linear span of $\boldsymbol{\Theta} - \boldsymbol{\theta}_0$, where $\boldsymbol{\theta}_0$ denotes the true value of $\boldsymbol{\theta}$ and $\boldsymbol{\Theta}_0$ denotes the true parameter space. Then for any $\boldsymbol{\theta} \in \{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0 : d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) = O(\delta_n)\}$, define the first order directional derivative of $l(\boldsymbol{\theta}, \boldsymbol{O})$ at the direction

$v \in V$ as

$$\dot{l}(\boldsymbol{\theta}, \boldsymbol{O})[v] = \left. \frac{\mathrm{d}l(\boldsymbol{\theta} + sv, \boldsymbol{O})}{\mathrm{d}s} \right|_{s=0}.$$

Also define the Fisher inner product for $v, \tilde{v} \in V$ as $< v, \tilde{v} >= P\left\{\dot{l}\left(\boldsymbol{\theta}, \boldsymbol{O}\right)[v]\dot{l}\left(\boldsymbol{\theta}, \boldsymbol{O}\right)[\tilde{v}]\right\}$ and the Fisher norm for $v \in V$ as $\|v\|^2 =< v, v >$. Let $\overline{V}$ be the closed linear span of V under the Fisher norm, then $(\overline{\mathbf{V}}, \|\cdot\|)$ is a Hilbert space. Furthermore, for a vector of $(2p+1)$-dimension $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1', \boldsymbol{\alpha}_2', \alpha_3)'$ with $\|\boldsymbol{\alpha}\|_E \leq 1$ and any $v \in V$, define a smooth functional of $\boldsymbol{\theta}$ as $h(\boldsymbol{\theta}) = \boldsymbol{\alpha}_1'\boldsymbol{\beta}_1 + \boldsymbol{\alpha}_2'\boldsymbol{\beta}_2 + \alpha_3\eta$ and

$$\dot{h}\left(\boldsymbol{\theta}_0\right)[v] = \left. \frac{\mathrm{d}h\left(\boldsymbol{\theta}_0 + sv\right)}{\mathrm{d}s} \right|_{s=0}$$

whenever the right hand-side limit is well defined. Then by the Riesz representation theorem, there exists $v^* \in \bar{V}$ such that $\dot{h}\left(\boldsymbol{\theta}_0\right)[v] =< v, v^* >$ for all $v \in \bar{V}$ and $\|v^*\|^2 = \|\dot{h}\left(\boldsymbol{\theta}_0\right)\|^2 = \|\dot{l}\left(\boldsymbol{\theta}_0, \boldsymbol{O}\right)[v^*]\|^2$. Therefore, by Theorem 1 of Shen (1997), we obtain that

$$\boldsymbol{\alpha}'\left((\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_{10})', (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20})', (\hat{\eta} - \eta_0)\right)' + \int_0^{\tau_c} g(t)d\left(\hat{\Lambda}_2(t) - \Lambda_{20}(t)\right))$$
$$= \frac{1}{n}\sum_{i=1}^{n}\dot{l}\left(\boldsymbol{\theta}_0, \boldsymbol{O}_i\right)[v^*] + o_p(n^{-1/2}).$$

Furthermore, the asymptotic normality is guaranteed by the central limits theorem and we have

$$n^{1/2}\boldsymbol{\alpha}'\left((\hat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_{10})', (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20})', (\hat{\eta} - \eta_0)\right)' + \int_0^{\tau_c} g(t)d\left(\hat{\Lambda}_2(t) - \Lambda_{20}(t)\right))$$

$$= n^{-1/2} \sum_{i=1}^{n} \dot{l}\left(\boldsymbol{\theta}_0, \boldsymbol{O}_i\right)[v^*] + o_p(1) \xrightarrow{D} N(0, \Sigma).$$

The semiparametric efficiency can be established by applying the result of Bickel and Kwon (2001) or Theorem 4 in Shen (1997). This completes the proof.

References

Bickel, P. J. and Kwon, J. (2001). Inference for semiparametric models: some questions and an answer. *Statistica Sinica* **94**, 863-886.

Chang, I. S., Wen, C. C. and Wu, Y. J. (2007). A profile likelihood theory for the correlated gamma-frailty model with current status family data. *Statistica Sinica* **17**, 1023-1046.

Lu, M., Zhang, Y. and Huang, J. (2007). Estimation of the mean function with panel count data using monotone polynomial splines. *Biometrika* **84**, 705-718.

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators *Econometrica: Journal of the Econometric Society* **57**, 1027-1057.

Shen, X. (1997). On methods of sieves and penalization. *The Annals of Statistics* **25**, 2555-2591.

Shen, X. and Wong, W. H. (1994). Convergence rate of sieve estimates. *The Annals of Statistics* **22**, 580-615.

Van der Vaart, A. W. (2000). *Asymptotic statistics (Vol. 3)*. New York: Cambridge University Press.

Van der Vaart, A. W. and Wellner, J. (1996) *Weak convergence and empirical processes*. New York: Springer.