Erratum to: "Why does a human die? A structural approach to cohort-wise mortality prediction under Survival Energy Hypothesis" [ASTIN Bulletin, 2021, 51, (1) 191-219.]

Yasutaka Shimizu

Department of Applied Mathematics
Waseda University

May 9, 2022

In Theorem 3.2, the expression of the asymptotic variance includes many typos. We shall give the complete version of the statement as well as the proof.

All the notations used below are the same as in the paper Shimizu *et al.* [1].

Theorem 3.2. Suppose the same assumptions as in Theorem 3.1. Moreover, suppose $q_c \in C^2(\Theta)$ and Θ is a convex subset of $\mathbb{R}^m$. Then,

$$\sqrt{n_c}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} R_d^{-1} Q_d \cdot N_d(0, \Sigma), \quad n_c \to \infty$$

where

$$Q_d = (\partial_\theta q_c(t_1, \theta_0|S), \dots, \partial_\theta q_c(t_d, \theta_0|S) \in \mathbb{R}^m \otimes \mathbb{R}^d,$$
$$R_d = \sum_{i=1}^d (\partial_\theta q_c)(\partial_\theta^\top q_c)(t_i, \theta_0|S) \in \mathbb{R}^m \otimes \mathbb{R}^m,$$

and the variance-covariance matrix $\Sigma = (\sigma_{ij})_{1 \le i,j \le m}$ is given by

$$\sigma_{ij} = \frac{1}{\overline{q}_c^2(S)}\Lambda(t_i, t_j) + \frac{\overline{q}_c(t_i)\overline{q}_c(t_j)}{\overline{q}_c^4(S)}\Lambda(S, S) - \frac{\overline{q}_c(t_i)}{\overline{q}_c^3(S)}\Lambda(t_j, S) - \frac{\overline{q}_c(t_j)}{\overline{q}_c^3(S)}\Lambda(t_i, S)$$

with $\Lambda(x, y) = q_c(x \wedge y) - q_c(x)q_c(y)$ and $\overline{q}_c = 1 - q_c$.

Proof. By the usual *"sub-sub sequence argument"*, we can assume that the convergence $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta_0$ can be replaced with the almost sure convergence without loss of generality. Hence, for n large enough, we may assume that $\widehat{\theta}_n$ belongs to the interior of the parameter space Θ; $\partial_\theta \ell_n(\widehat{\theta}_n) = 0$, and that $\det \partial_\theta^2 \ell_n(\theta) \neq 0$ around θ_0. These assumptions can make the proof simple.

Under the above assumptions, we use Taylor's formula to obtain that

$$\sqrt{n_c}(\widehat{\theta}_n - \theta_0) = \left[\int_0^1 \partial_\theta^2 \ell_n(\theta_0 + u(\widehat{\theta}_n - \theta_0))\, \mathrm{d}u\right]^{-1} \cdot \sqrt{n_c}\partial_\theta \ell_n(\theta_0)), \qquad (3.1)$$

where

$$\sqrt{n_c}\partial_\theta \ell_n(\theta_0) = 2\sum_{i=1}^d \sqrt{n_c}\left(\widehat{q}_c(t_i|S) - q_c(t_i,\theta_0|S)\right)\partial_\theta q_c(t_i,\theta_0|S);$$

$$\partial_\theta^2 \ell_n(\theta) = 2\sum_{i=1}^d \left[(\partial_\theta q_c)(\partial_\theta^\top q_c)(t_i,\theta|S) + \{q_c(t_i,\theta|S) - \widehat{q}_c(t_i|S)\}\partial_\theta^2 q_c(t_i,\theta|S)\right].$$

Note that $\widehat{q}_c(t_i|S)$ is the empirical distribution that is consistent with $q_c(t_i|S)$:

$$\widehat{q}_c(t_i|S) = \frac{\widehat{F}_n(t_i) - \widehat{F}_n(S)}{1-\widehat{F}_n(S)}, \quad q_c(t_i|S) = \frac{F(t_i)-F(S)}{1-F(S)},$$

where F is the distribution function of τ_i^c and

$$\widehat{q}_c(t) := \frac{1}{n_c}\sum_{i=1}^{n_c} \mathbf{1}_{\{\tau_i^c \le t\}}$$

and that it follows from the Donsker-type theorem that, for any $x_i \in \mathbb{R}$,

$$(\sqrt{n_c}(\widehat{q}_c(x_1) - q_c(x_1)), \dots, \sqrt{n_c}(\widehat{q}_c(x_K) - q_c(x_K))) \xrightarrow{\mathcal{D}} (Z_{x_1}, \dots, Z_{x_k}) \sim N_K(0,\Lambda),$$

as $n_c \to \infty$, where $\Lambda = (\Lambda(x_i,x_j))_{1\le i,j\le K}$ with $\Lambda(x_i,x_j) = q_c(x_i \wedge x_j) - q_c(x_i)q_c(x_j)$ (see van der Vaart [2], Theorem 19.3). Hence, it follows that

$$\begin{aligned}\sqrt{n_c}(\widehat{q}_c(t_i|S) - q_c(t_i|S)) &= \frac{\sqrt{n_c}(\widehat{q}_n(t_i) - q_c(t_i)) - \sqrt{n_c}(\widehat{q}_n(S) - q_c(S))}{1-\widehat{q}_n(S)} \\ &\quad + (q_c(t_i) - q_c(S))\sqrt{n_c}\left(\frac{1}{1-\widehat{q}_n(S)} - \frac{1}{1-q_c(S)}\right) \\ &\xrightarrow{\mathcal{D}} \frac{1}{\overline{q}_c(S)}Z_{t_i} - \frac{\overline{q}_c(t_i)}{\overline{q}_c^2(S)}Z_S\end{aligned}$$

jointly for $i = 1,2,\dots,d$, and therefore,

$$\begin{aligned}\sqrt{n_c}\partial_\theta \ell_n(\theta_0) &\xrightarrow{\mathcal{D}} -2\sum_{i=1}^d \left[\frac{1}{\overline{q}_c(S)}Z_{t_i} - \frac{\overline{q}_c(t_i)}{\overline{q}_c^2(S)}Z_S\right]\partial_\theta q_c(t_i,\theta_0|S) \\ &= -2Q_d \cdot N_d(0,\Sigma),\end{aligned}$$

where $\Sigma = (\sigma_{ij})_{1\le i,j\le m}$ is given by

$$\sigma_{ij} = \frac{1}{\overline{q}_c^2(S)}\Lambda(t_i,t_j) + \frac{\overline{q}_c(t_i)\overline{q}_c(t_j)}{\overline{q}_c^4(S)}\Lambda(S,S) - \frac{\overline{q}_c(t_i)}{\overline{q}_c^3(S)}\Lambda(t_j,S) - \frac{\overline{q}_c(t_j)}{\overline{q}_c^3(S)}\Lambda(t_i,S)$$

because, by putting $Y_i := \frac{1}{\overline{q}_n(S)}Z_{t_i} - \frac{\overline{q}_n(t_i)}{\overline{q}_n^2(S)}Z_S$, for which $\mathbb{E}[Y_i] = 0$, it follows that

$$Cov(Y_i,Y_j) = \mathbb{E}\left[\left(\frac{1}{\overline{q}_c(S)}Z_{t_i} - \frac{\overline{q}_c(t_i)}{\overline{q}_c^2(S)}Z_S\right)\left(\frac{1}{\overline{q}_c(S)}Z_{t_j} - \frac{\overline{q}_c(t_j)}{\overline{q}_c^2(S)}Z_S\right)\right] = \sigma_{ij}.$$

Moreover, it follows by the same argument as in Theorem 3.1 for the proof of (a): the uniform convergence of ℓ_n, we can show that

$$\sup_{\theta\in\Theta} |\partial_\theta^2 \ell_n(\theta) - \partial_\theta^2 \ell(\theta)| \xrightarrow{\mathbb{P}} 0, \quad n_c \to \infty.$$

Hence, with the consistency of $\widehat{q}_c$ to q_c, we have that

$$\begin{aligned}\partial_\theta^2 \ell_n(\theta_0 + u(\widehat{\theta}_n - \theta_0)) &\xrightarrow{\mathbb{P}} \sum_{i=1}^d \Big[(\partial_\theta q_c)(\partial_\theta^\top q_c)(t_i, \theta_0|S) + \{q_c(t_i, \theta_0|S) - q_c(t_i|S)\}\partial_\theta^2 q_c(t_i, \theta_0|S)\Big] \\ &= \sum_{i=1}^d (\partial_\theta q_c)(\partial_\theta^\top q_c)(t_i, \theta_0|S) =: R_d,\end{aligned}$$

by the continuous mapping theorem and the assumption of the model specification: $q_c(\cdot, \theta_0|S) \equiv q_c(\cdot|S)$. Finally, we obtain from (3.1) that

$$\sqrt{n_c}(\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{D}} -(2R_d)^{-1} \cdot 2Q_d \cdot N_d(0, \Sigma) =^d R_d^{-1} Q_d \cdot N_d(0, \Sigma), \quad n_c \to \infty.$$

This completes the proof. □

We finally point out a typo in the first display in page 200, [1]:

$$q_c(t|S) := \mathbb{P}(\tau^c \le t | \tau^c > S) = \frac{q_c(t) - q_c(S)}{1 - q_c(S)}.$$

That's all.

References

[1] Shimziu, Y; Minami, Y. and Ito, R. (2021). Why does a human die? A structural approach to cohort-wise mortality prediction under Survival Energy Hypothesis, *ASTIN Bulletin*, **51**, (1) 191-219.

[2] van der Vaart, A. W. (1998). *Asymptotic statistics.* Cambridge University Press, Cambridge.