Web-based Supplementary Materials for "Identifying Predictive Markers for Personalized Treatment Selection"

by

Yuanyuan Shen and Tianxi Cai

This supplementary material contains the proof of convergence of the proposed test statistics.

Convergence of the Proposed Test Statistic

We can write the statistic in terms of stochastic process:

$$\widehat{Q}_{\psi} = n^{-1}\widehat{\mathbf{\Delta}}^{\mathsf{T}}\mathbb{K}_n\widehat{\mathbf{\Delta}} = \int\int k(\mathbf{x},\mathbf{x}')d\widehat{\Theta}(\mathbf{x})d\widehat{\Theta}(\mathbf{x}') \tag{1}$$

where

$$\widehat{\Theta}(\mathbf{x}) = n^{-\frac{1}{2}}\sum_{i=1}^{n}\frac{(Y_i-\bar{Y}_1)I(T_i=1)}{\hat{\pi}_1}I(\mathbf{x}_i\leqslant\mathbf{x}) - n^{-\frac{1}{2}}\sum_{i=1}^{n}\frac{(Y_i-\bar{Y}_0)I(T_i=0)}{\hat{\pi}_0}I(\mathbf{x}_i\leqslant\mathbf{x}). \tag{2}$$

To derive the influence function of $\widehat{\Theta}(\mathbf{x})$, we write the first part of $\widehat{\Theta}(\mathbf{x})$ in the following way:

$$\begin{aligned}
&n^{-\frac{1}{2}}\sum_{i=1}^{n}\frac{(Y_i-\bar{Y}_1)I(T_i=1)}{\hat{\pi}_1}I(\mathbf{X}_i\leqslant\mathbf{x})\\
=&n^{-\frac{1}{2}}\sum_{i=1}^{n}\frac{(Y_i-\mu_1)I(T_i=1)}{\pi_1}\left[I(\mathbf{X}_i\leqslant\mathbf{x}) - n^{-1}\sum_{i=1}^{n}I(T_i=1)I(\mathbf{X}_i\leqslant\mathbf{x})\right] + o_P(1)\\
=&n^{-\frac{1}{2}}\sum_{i=1}^{n}\frac{(Y_i-\mu_1)I(T_i=1)}{\pi_1}\left[I(\mathbf{X}_i\leqslant\mathbf{x}) - \mathcal{F}(\mathbf{x})\right] + o_P(1)
\end{aligned}$$

where $\mathcal{F}(\mathbf{x}) = P(\mathbf{X}_i\leqslant\mathbf{x})$. Since by a uniform law of large numbers (ULLN) (Pollard, 1990), $n^{-1}\sum_{i=1}^{n}\frac{I(T_i=1)I(\mathbf{X}_i\leqslant\mathbf{x})}{\pi_1}$ converges in probability to its limit, $\mathcal{F}(\mathbf{x})$, uniformly in $\mathbf{x}$. Therefore,

$$\widehat{\Theta}(\mathbf{x}) = n^{-\frac{1}{2}}\sum_{i=1}^{n}\theta_i(\mathbf{x}) + o_P(1), \tag{3}$$

where

$$\theta_i = \left[\frac{(Y_i-\mu_1)I(T_i=1)}{\pi_1} - \frac{(Y_i-\mu_0)I(T_i=1)}{\pi_0}\right]\left[I(\mathbf{X}_i\leqslant\mathbf{x}) - \mathcal{F}(\mathbf{x})\right] \tag{4}$$

It is not hard to show that $E\{\theta_i(\mathbf{x})\} = 0$. In addition, it follows from a functional central limit theorem (Pollard, 1990) that $\widehat{\Theta}(\mathbf{x})$ converges jointly to a zero mean Gaussian process $G(\mathbf{x})$. By

Lemma A.3 of Bilias et al. (1997) and the strong representation theorem, we have

$$(1) \to \int\int k(\mathbf{x}, \mathbf{x}')dG(\mathbf{x})dG(\mathbf{x}')$$

References

Bilias, Y., Gu, M., Ying, Z., et al. (1997). Towards a general asymptotic theory for cox model with staggered entry. *The Annals of Statistics* **25,** 662–682.

Pollard, D. (1990). *Empirical processes: theory and applications.* NSF-CBMS Reginal Conference Series in Probability and Statistics 2. Institute of Mathematical Statistics and American Statistical Association.