Statistica Sinica Preprint No: SS-2016-0185.R1

Title	A Further Study of Propensity Score Calibration in Missing Data Analysis
Manuscript ID	SS-2016-0185.R1
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202016.0185
Complete List of Authors	Peisong Han
Corresponding Author	Peisong Han
E-mail	peisonghan@uwaterloo.ca
Notice: Accepted version subject to English editing.	

Statistica Sinica: Preprint
doi:10.5705/ss.202016.0185

is a model for $E(Y \mid \boldsymbol{X})$ and $\boldsymbol{\gamma}$ is a finite-dimensional unknown parameter. Because $R \perp Y \mid \boldsymbol{X}$ from the MAR mechanism, $\boldsymbol{\gamma}$ is conventionally estimated by $\hat{\boldsymbol{\gamma}}$ based on a complete-case analysis. When $a(\boldsymbol{\gamma}; \boldsymbol{X})$ is a reasonable model for $E(Y \mid \boldsymbol{X})$, $\hat{\mu}_{\text{aipw}}(\hat{\boldsymbol{\gamma}})$ usually has better efficiency than $\hat{\mu}_{\text{ipw}}$. When $a(\boldsymbol{\gamma}; \boldsymbol{X})$ is correctly specified in that $a(\boldsymbol{\gamma}_0; \boldsymbol{X}) = E(Y \mid \boldsymbol{X})$ for some $\boldsymbol{\gamma}_0$, $\hat{\mu}_{\text{aipw}}(\hat{\boldsymbol{\gamma}})$ attains the semiparametric efficiency bound.

When $a(\boldsymbol{\gamma}; \boldsymbol{X})$ is incorrectly specified, $\hat{\mu}_{\text{aipw}}(\hat{\boldsymbol{\gamma}})$ can be quite inefficient (Chen, Leung, and Qin 2008; Rubin and van der Laan 2008). There have been many recent developments on gaining efficiency in this case. Two main gains have been achieved: intrinsic efficiency and improved efficiency. Estimators that are intrinsically efficient have influence functions of the form

$$\text{Resid}\left\{\frac{R(Y-\mu_0)}{\pi(\boldsymbol{X})}, \frac{R-\pi(\boldsymbol{X})}{\pi(\boldsymbol{X})}h(\boldsymbol{X})\right\}.$$

Hereafter, for any random variable ξ and finite-dimensional random vector $\boldsymbol{\phi}$ with mean zero and finite second moments, $\text{Resid}(\xi, \boldsymbol{\phi}) = \xi - E(\xi\boldsymbol{\phi}^{\text{T}})\{E(\boldsymbol{\phi}\boldsymbol{\phi}^{\text{T}})\}^{-1}\boldsymbol{\phi}$ denotes the residual of the projection of ξ onto span$\{\boldsymbol{\phi}\}$, the linear space spanned by components of $\boldsymbol{\phi}$. Apparently the projection residual has the smallest variance among the class of influence functions $R(Y-\mu_0)/\pi(\boldsymbol{X}) - c\{R-\pi(\boldsymbol{X})\}h(\boldsymbol{X})/\pi(\boldsymbol{X})$ with an arbitrary c. Various intrinsically efficient estimators have been proposed and studied by Tan (2006; 2008; 2010), Chen, Leung, and Qin (2008), Chan (2012) and Rotnitzky et al. (2012). Improved efficiency, on the other hand, is achieved by using an estimator $\tilde{\boldsymbol{\gamma}}$ instead of $\hat{\boldsymbol{\gamma}}$, where $\tilde{\boldsymbol{\gamma}}$ converges in probability to the minimizer of the asymptotic variance of $\hat{\mu}_{\text{aipw}}(\boldsymbol{\gamma})$. Estimators of μ_0 with improved efficiency have been proposed and studied by Rubin and van der Laan (2008), Tan (2008; 2010), Cao, Tstiatis, and Davidian (2009) and Rotnitzky et al. (2012). Many of these estimators are doubly robust: they are still consistent if $\pi(\boldsymbol{\alpha}; \boldsymbol{X})$ is misspecified but $a(\boldsymbol{\gamma}; \boldsymbol{X})$ is not.