The Canadian Journal of Statistics

Vol. xx, No. yy, 20??, Pages 1–25

La revue canadienne de statistique

Web-based Supplementary Materials for "Estimating treatment effects in observational studies with both prevalent and incident cohorts"

Jing Ning[1]*, Chuan Hong[2], Liang Li[1], Xuelin Huang[1] and Yu Shen[1]

[1]*Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston TX, 77030, USA*

[2]*Department of Biostatistics, Harvard School of Public Health, Boston MA, 02115, USA*

1. SENSITIVITY ANALYSIS

We have conducted some sensitivity studies to evaluate the robustness of the proposed method with violations of model assumptions. We considered two scenarios: Scenario S1) the true survival model in the data generation was a function of Z and X, but the fitted outcome model used in the estimation procedure was specified as a function of Z and the propensity score $e(X)$; and Scenario S2) the true propensity score model in the data generation contained nonlinear terms and the fitted regression model only included linear terms.

* *Author to whom correspondence may be addressed.*

E-mail: jning@mdanderson.org

Under Scenario S1, the failure time T_i was generated from an exponential distribution with a rate of $\exp\{0.5Z_i + 0.4X_{i1} + 0.4X_{i2} + 0.4X_{i3}\}$. The simulation results are summarized in Table S1. Although the outcome model was misspecified, the estimated treatment effect had a small empirical bias and appropriate coverage probability. Note that the regression coefficient of the propensity score under the Cox model is a nuisance parameter here. We did not know the true value of the regression coefficient of the propensity score under the Cox model, and hence do not include its estimators in Table S1.

Under Scenario S2, the true propensity score model was $e(\mathbf{x_i}) = P(Z_i|\mathbf{X_i}) =$ expit$(0.1 + 0.4X_{i1} + 0.4X_{i2} + 0.4X_{i3} + \beta_s X_{i1}^2)$, where $\beta_s = 0.2, 0.4, 0.6, 0.8$. The simulation results are summarized in Table S2. When the regression coefficient of the square term was relatively small (e.g., ≤ 0.4), the estimators of both the propensity score model and the Cox model had small empirical biases and reasonable coverage probabilities. As expected, with the increasing regression coefficient of the square term, the coefficient of X_1 under the propensity score model had increasing biases and decreasing coverage probabilities. However, the regression coefficients of the Cox model had reasonable biases and coverage probabilities even when the coefficient of the square term was twice as large as the regression coefficient of the linear term.

Table S1: Summary of $1,000$ simulations when the true survival model in the data generation was a function of Z and X: empirical bias (Bias), empirical standard error (ESE), asymptotic standard error (ASE), and coverage probability (CP).

Sample Size		Scenario I: $C_1\% = 68\%$, $C_2\% = 67\%$					Scenario II: $C_1\% = 52\%$, $C_2\% = 46\%$				
(m,n)		$\boldsymbol{\gamma} = (0.1, 0.4, 0.4, 0.4)$				$\alpha = 0.5$	$\boldsymbol{\gamma} = (0.1, 0.4, 0.4, 0.4)$				$\alpha = 0.5$
(200,200)	Bias	-.003	.021	.008	.018	.041	-.003	.021	.008	.018	.032
	ESE	.163	.157	.156	.155	.205	.163	.157	.156	.155	.156
	ASE	.152	.158	.159	.159	.200	.152	.158	.159	.159	.159
	CP	.938	.958	.956	.970	.940	.938	.958	.956	.970	.958
(400,200)	Bias	.000	.014	-.003	-.005	.015	.000	.014	-.003	-.005	.009
	ESE	.102	.114	.118	.114	.163	.102	.114	.118	.114	.128
	ASE	.106	.110	.111	.110	.161	.106	.110	.111	.110	.128
	CP	.960	.946	.912	.944	.944	.960	.946	.912	.944	.960
(600,200)	Bias	.002	.010	.008	.007	.022	.002	.010	.008	.007	.018
	ESE	.086	.090	.090	.086	.148	.086	.090	.090	.086	.116
	ASE	.087	.090	.090	.090	.140	.087	.090	.090	.090	.111
	CP	.948	.954	.948	.958	.942	.948	.954	.948	.958	.940

$C_1\%$ and $C_2\%$: the respective censoring rates of the incident cohort and prevalent cohort

2. ADDITIONAL SIMULATION STUDIES FOR EFFICIENCY COMPARISON

We have compared the small sample performance of the estimators from the estimation equations $U_1(\boldsymbol{\gamma})$ and $U_2(\boldsymbol{\gamma}, \alpha, \beta)$, and the estimation equations $U_{LB1}(\boldsymbol{\gamma})$

Table S2: Summary of $1,000$ simulations in the presence of mis-specification of the propensity score function : empirical bias (Bias), empirical standard error (ESE), asymptotic standard error (ASE), and coverage probability (CP).

β_s		Scenario I: $C_1 \sim (0, 0.25), C_2 \sim (0, 0.52)$						Scenario II: $C_1 \sim (0, 0.5), C_2 \sim (0, 1.3)$					
		$\boldsymbol{\gamma} = (0.1, 0.4, 0.4, 0.4)$				$(\alpha, \beta) = (0.5, 1)$		$\boldsymbol{\gamma} = (0.1, 0.4, 0.4, 0.4)$				$(\alpha, \beta) = (0.5, 1)$	
0.2	Bias	.173	-.018	.005	.012	.038	-.039	.173	-.018	.005	.012	.037	-.057
	ESE	.165	.161	.157	.167	.160	.589	.165	.161	.157	.167	.132	.493
	MBSE	.153	.157	.159	.159	.158	.600	.153	.157	.159	.159	.133	.516
	CP	.798	.958	.956	.948	.948	.962	.798	.958	.956	.948	.946	.956
0.4	Bias	.337	-.075	-.006	-.007	.066	-.085	.337	-.075	-.006	-.007	.060	-.083
	ESE	.161	.160	.155	.160	.166	.646	.161	.160	.155	.160	.137	.528
	MBSE	.154	.153	.160	.159	.157	.638	.154	.153	.160	.159	.132	.552
	CP	.426	.892	.956	.948	.922	.940	.426	.892	.956	.948	.928	.934
0.6	Bias	.475	-.135	-.014	-.018	.081	-.084	.475	-.135	-.014	-.018	.081	-.078
	ESE	.165	.152	.161	.160	.160	.729	.165	.152	.161	.160	.134	.620
	MBSE	.156	.148	.161	.160	.157	.708	.156	.148	.161	.160	.133	.619
	CP	.136	.824	.944	.952	.926	.936	.136	.824	.944	.952	.908	.922
0.8	Bias	.591	-.181	-.021	-.032	.114	-.120	.591	-.181	-.021	-.032	.107	-.102
	ESE	.167	.154	.163	.165	.163	.766	.167	.154	.163	.165	.138	.652
	MBSE	.158	.143	.163	.162	.158	.771	.158	.143	.163	.162	.134	.676
	CP	.026	.730	.942	.934	.886	.928	.026	.730	.942	.934	.846	.934

C_1 and C_2: the respective censoring times of the incident and prevalent cohorts.

The sample size is (200, 200).

and $U_{LB2}(\boldsymbol{\gamma}, \alpha, \beta)$. We generated data from Scenario I and Scenario II, in which the sample sizes of the incident cohort and prevalent cohort were 400 and 200, respectively. Table S3 lists the empirical biases and empirical standard errors based on $1,000$ simulations. Both methods have small empirical biases. As expected, by incorporating the distribution information of the truncation time, the estimators obtained from estimation equations $U_{LB1}(\boldsymbol{\gamma})$ and $U_{LB2}(\boldsymbol{\gamma}, \alpha, \beta)$ are more efficient than those from estimation equations $U_1(\boldsymbol{\gamma})$ and $U_2(\boldsymbol{\gamma}, \alpha, \beta)$, which are for general left-truncated data. Note that the covariate and treatment information of the uncensored subjects is utilized in $U_{LB1}(\boldsymbol{\gamma})$ to improve the statistical efficiency for the estimated propensity score. When the censoring percentage decreases, the relative efficiency gain of the estimation equations $U_{LB1}(\boldsymbol{\gamma})$ and $U_{LB2}(\boldsymbol{\gamma}, \alpha, \beta)$ to $U_1(\boldsymbol{\gamma})$ and $U_2(\boldsymbol{\gamma}, \alpha, \beta)$ increases.

Table S3: Summary of $1,000$ simulations for efficiency comparison: empirical bias (Bias) and empirical standard error (ESE)

	Scenario I: $C_1\% = 60\%$, $C_2\% = 40\%$				Scenario II: $C_1\% = 40\%$, $C_2\% = 20\%$			
	U_1 and U_2		U_{LB1} and U_{LB2}		U_1 and U_2		U_{LB1} and U_{LB2}	
	Bias	ESE	Bias	ESE	Bias	ESE	Bias	ESE
γ_0	.000	.102	-.004	.099	.000	.102	-.007	.094
γ_1	.014	.114	.014	.106	.014	.114	.018	.101
γ_2	-.003	.118	.000	.112	-.003	.118	.004	.106
γ_3	-.005	.114	-.004	.108	-.005	.114	.001	.104
α	.012	.136	.015	.131	.012	.112	.014	.106
β	-.013	.463	-.012	.422	.006	.395	-.003	.348