Summary of the simulation result for improving efficiency of WCLS

Tianchen Qian

2018.07.17

1 Estimator

Consider four estimators: WCLS, WCLS with $\tilde{p}_t S_t$ included in $g(H_t)$, WCLS with A_t not centered in the residual part, and WCLS with A_t not centered in the residual part and a special weight

WCLS-1: WCLS is the solution to the following estimating equation:

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\{Y_{t+1} - g(H_t)^T\alpha - (A_t - \tilde{p}_t)S_t^T\beta\}W_t \begin{bmatrix} g(H_t) \\ (A_t - \tilde{p}_t)S_t \end{bmatrix}.$$

WCLS-2: WCLS with $\tilde{p}_t S_t$ included in $g(H_t)$ is self-explanatory.

WCLS-3: WCLS with A_t not centered in the residual part is the solution to the following estimating equation:

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\{Y_{t+1} - g(H_t)^T\alpha - A_t S_t^T\beta\}W_t \begin{bmatrix} g(H_t) \\ (A_t - \tilde{p}_t)S_t \end{bmatrix}.$$

WCLS-4: WCLS with A_t not centered in the residual part is the solution to the following estimating equation:

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\{Y_{t+1} - g(H_t)^T\alpha - A_t S_t^T\beta\}W_t \frac{1}{\tilde{p}_t(1-\tilde{p}_t)} \begin{bmatrix} g(H_t) \\ (A_t - \tilde{p}_t)S_t \end{bmatrix}.$$

WCLS-5: WCLS with $p_t S_t$ included in $g(H_t)$ is self-explanatory. (Peng suggests this.)

Here, the weight variable W_t equals

$$W_t = \left(\frac{\tilde{p}_t}{p_t}\right)^{A_t} \left(\frac{1-\tilde{p}_t}{1-p_t}\right)^{1-A_t}.$$

Theorem 1. *If our working model (the g part) accidentally is correct and our treatment model is correct conditional on entire history and residual variance is constant, (which implies that we can set $\tilde{p}_t = p_t$ and $W_t = 1$), then WCLS-3 is semiparametric efficient.*

Proof. See the note "note_20180730 - EIF alternative form (action centering) for continuous and binary outcomes.pdf" (copied from the folder of binary outcome project). □

Conjecture: WCLS-2 has similar performance to WCLS-3 when S_t is low-dimensional compared to n so that including $\tilde{p}_t S_t$ in $g(H_t)$ does not result in a big loss of degrees of freedom.

Note: Susan says that WCLS-4 is efficient in the above theorem situation instead of WCLS-3. I don't think so.

2 Simulation

2.1 Generative model

- Covariate Z_t is an exogenous AR(1) process with auto-correlation 0.5.
- The randomization probability is $p_t(H_t) = \min[0.8, \max\{0.2, \text{expit}(0.5Z_t)\}]$.
- The outcome Y_{t+1} is generated as Gaussian with mean $\alpha_0 + \alpha_1 Z_t + A_t(\beta_0 + \beta_1 Z_t)$ and variance 1.
- The parameter value is $\beta_0 = 0.5$, $\beta_1 = 1$, $\alpha_0 = -1$, $\alpha_1 = 1$.

2.2 Simulation result

We correctly specify all the models for all estimators (so that Z_t is included in both the control part and the treatment effect part, and we set $\tilde{p}_t = p_t(H_t)$).

Result is in Table 1. Observations:

- All three estimators have close to 0 bias. (And by theory we know they are all consistent.)
- WCLS-2 and WCLS-3 have almost the same SD, whereas WCLS-1 is less efficient than the other two in estimating β_1.
- WCLS-4 is slightly less efficient than WCLS-3 and WCLS-2, but more efficient than WCLS-1.

Table 1: Consistency and relative efficiency among the three estimators, based on 10,000 simulations

		β_0		β_1	
		Bias	SD	Bias	SD
$n = 100, T = 30$	WCLS-1	3.8×10^{-4}	0.0386	1.7×10^{-4}	0.0376
	WCLS-2	2.5×10^{-4}	0.0380	-1.9×10^{-4}	0.0354
	WCLS-3	2.5×10^{-4}	0.0380	-1.8×10^{-4}	0.0354
	WCLS-4	2.5×10^{-4}	0.381	-1.6×10^{-4}	0.0356
$n = 30, T = 210$	WCLS-1	8.7×10^{-4}	0.0259	6.3×10^{-4}	0.0260
	WCLS-2	7.9×10^{-4}	0.0258	4.5×10^{-4}	0.0242
	WCLS-3	7.9×10^{-4}	0.0258	5.0×10^{-4}	0.0241
	WCLS-4				