Notes on function AAIW

Kohei Yata

August 12, 2017

Arguments

- Y: $n \times 1$ numpy matrix
- U: $n \times m$ numpy matrix
- EU: $n \times m$ numpy matrix
- EUU: $n \times m \times m$ numpy array
- Z: $n \times l$ numpy matrix
- n: population size, which is equal to sample size in our case

This function maps these arguments into the AAIW standard error for $\hat{\theta}_n$ defined below.

Set Up

We use the notation of AAIW (2017). For each i, we observe $(Y_{n,i}, U_{n,i}, Z_{n,i})$. In our case, $U_{n,i}$ is a vector of m binary elements and the k-th element of $U_{n,i}$ is equal to D_{i,t_k}, where m is the number of treatments. $Z_{n,i}$ is a vector of l fixed attributes including the propensity scores. For each i, $E[U_{n,i}]$ is known because we know the propensity of each treatment. Besides, $E[U_{n,i}U'_{n,i}]$ is known because $U^2_{n,i,k} = U_{n,i,k}$ and $U_{n,i,k}U_{n,i,j} = 0$ when $k \neq j$ so that $E[U_{n,i}U'_{n,i}]$ is an $m \times m$ diagonal matrix with diagonal entries $E[U_{n,i,1}], \ldots, E[U_{n,i,m}]$, where $U_{n,i,k}$ is the k-th element of $U_{n,i}$. The function AAIW includes $E[U_{n,i}]$ and $E[U_{n,i}U'_{n,i}]$ in its arguments, and can be used when they are known like in our case. To implement this function, we need to stack $(Y_{n,i}, U_{n,i}, E[U_{n,i}], E[U_{n,i}U'_{n,i}], Z_{n,i})$ for all i, i.e., let $Y = (Y_{n,1}, \ldots, Y_{n,n})'$, $U = (U_{n,1}, \ldots, U_{n,n})'$, $EU = (EU_{n,1}, \ldots, EU_{n,n})'$, $Z = (Z_{n,1}, \ldots, Z_{n,n})'$, and EUU be an $n \times m \times m$ array made by stacking $m \times m$ matrices $E[U_{n,1}U'_{n,1}], \ldots, E[U_{n,n}U'_{n,n}]$ so that $EUU[i,k,j] = E[U_{n,i,k}U_{n,i,j}]$.

Formulas

- $\Lambda_n = (\sum_{i=1}^n E[U_{n,i}]Z'_{n,i})(\sum_{i=1}^n Z_{n,i}Z'_{n,i})^{-1} = (EU'Z)(Z'Z)^{-1}$.
- $X_{n,i} = U_{n,i} - \Lambda_n Z_{n,i}$ and $X = (X_{n,1}, \ldots, X_{n,n})' = U - Z\Lambda'_n$.
- $W = [X \;\; Z]$.
- $(\hat{\theta}_n, \hat{\gamma}_n)' = \underset{(\theta,\gamma)'}{\arg\min} \sum_{i=1}^n (Y_{n,i} - X'_{n,i}\theta - Z'_{n,i}\gamma)^2 = (W'W)^{-1}(W'Y)$.
- $H = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n E[X_{n,i}X'_{n,i}]$.
- $\epsilon_{n,i} = Y_{n,i} - X'_{n,i}\theta_n^{\text{causal}} - Z'_{n,i}\gamma_n^{\text{causal}}$.
- $\Delta^{\text{cond}} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \text{var}(X_{n,i}\epsilon_{n,i})$.

By Theorem 3 (ii),

$$\sqrt{N}(\hat{\theta}_n - \theta_n^{\text{causal,sample}}) \xrightarrow{d} \mathcal{N}(0, H^{-1}\Delta^{\text{cond}}H^{-1}).$$

The estimator for H is[1]

$$\begin{aligned}
H_n &= \frac{1}{n}\sum_{i=1}^n E[X_{n,i}X'_{n,i}] \\
&= \frac{1}{n}\sum_{i=1}^n (E[U_{n,i}U'_{n,i}] - E[U_{n,i}]Z'_{n,i}\Lambda'_n - \Lambda_n Z_{n,i}E[U_{n,i}] + \Lambda_n Z_{n,i}Z'_{n,i}\Lambda'_n) \\
&= \frac{1}{n}\sum_{i=1}^n E[U_{n,i}U'_{n,i}] - \frac{1}{n}\Big(\sum_{i=1}^n E[U_{n,i}]Z'_{n,i}\Big)\Big(\sum_{i=1}^n Z_{n,i}Z'_{n,i}\Big)^{-1}\Big(\sum_{i=1}^n Z_{n,i}E[U'_{n,i}]\Big) \\
&= \frac{1}{n}\sum_{i=1}^n E[U_{n,i}U'_{n,i}] - \frac{1}{n}\Lambda_n Z'EU.
\end{aligned}$$

To estimate Δ^{cond}, first run a least squares regression of $X_{n,i}\hat{\epsilon}_{n,i}$ on $Z_{n,i}$ and get

$$\hat{G}_n = \Big(\sum_{i=1}^n X_{n,i}\hat{\epsilon}_{n,i}Z'_{n,i}\Big)\Big(\sum_{i=1}^n Z_{n,i}Z'_{n,i}\Big)^{-1} = ((\widehat{X\epsilon})'Z)(Z'Z)^{-1},$$

[1] This estimator can be used when $E[U_{n,i}U'_{n,i}]$ is known. If it is unknown, the estimator suggested in AAIW is $\hat{H}_n = \frac{1}{n}\sum_{i=1}^n X_{n,i}X'_{n,i}$.

where $\hat{\epsilon}_{n,i} = Y_{n,i} - X'_{n,i}\hat{\theta}_n - Z'_{n,i}\hat{\gamma}_n$ and $\widehat{X\epsilon} = (X_{n,1}\hat{\epsilon}_{n,1}, \ldots, X_{n,n}\hat{\epsilon}_{n,n})'$. The estimator for Δ^{cond} is

$$\begin{aligned}
\hat{\Delta}_n^{\text{Z}} &= \frac{1}{n}\sum_{i=1}^{n}(X_{n,i}\hat{\epsilon}_{n,i} - \hat{G}_n Z_{n,i})(X_{n,i}\hat{\epsilon}_{n,i} - \hat{G}_n Z_{n,i})' \\
&= \frac{1}{n}(\widehat{X\epsilon} - Z\hat{G}'_n)'(\widehat{X\epsilon} - Z\hat{G}'_n).
\end{aligned}$$

Finally the estimator for the asymptotic variance of $\hat{\theta}_n$ is

$$\widehat{AVar}(\hat{\theta}_n) = H_n^{-1}\hat{\Delta}_n^{\text{Z}} H_n^{-1}/n,$$

and the AAIW standard error is the square root of the diagonal elements of $\widehat{AVar}(\hat{\theta}_n)$.

The EHW standard error

If researchers assume the population is infinite ($\rho = 0$) and ingore the design-based uncertainty, by Theorem 3 (iii), the asymptotic distribution they consider is

$$\sqrt{N}(\hat{\theta}_n - \theta_n^{\text{descr}}) \xrightarrow{d} \mathcal{N}(0, H^{-1}\Delta^{\text{ehw}}H^{-1}),$$

where $\Delta^{\text{ehw}} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} E[X_{n,i}\epsilon_{n,i}^2 X'_{n,i}]$. The estimator for Δ^{ehw} is

$$\begin{aligned}
\hat{\Delta}_n^{\text{ehw}} &= \frac{1}{n}\sum_{i=1}^{n} X_{n,i}\hat{\epsilon}_{n,i}^2 X'_{n,i} \\
&= \frac{1}{n}(\widehat{X\epsilon})'(\widehat{X\epsilon}).
\end{aligned}$$

Then the estimator for the EHW asymptotic variance is

$$\widehat{AVar}^{\text{ehw}}(\hat{\theta}_n) = H_n^{-1}\hat{\Delta}_n^{\text{ehw}} H_n^{-1}/n,$$

and the EHW standard error is the square root of the diagonal elements of $\widehat{AVar}^{\text{ehw}}(\hat{\theta}_n)$. Note that $\hat{\Delta}_n^{\text{ehw}} - \hat{\Delta}_n^{\text{Z}} = \frac{1}{n}(\widehat{X\epsilon})'Z(Z'Z)^{-1}Z'(\widehat{X\epsilon})$, which is positive semi-definite.

The conventional robust standard error

Let $\tilde{W} = [U \;\; Z]$, $(\tilde{\theta}_n, \tilde{\gamma}_n)' = \arg\min_{(\theta,\gamma)'} \sum_{i=1}^{n}(Y_{n,i} - U'_{n,i}\theta - Z'_{n,i}\gamma)^2 = (\tilde{W}'\tilde{W})^{-1}(\tilde{W}'Y)$, $\tilde{\epsilon}_{n,i} = Y_{n,i} - U'_{n,i}\tilde{\theta}_n - Z'_{n,i}\tilde{\gamma}_n$ and $\widetilde{W\epsilon} = (\tilde{W}_{n,1}\tilde{\epsilon}_{n,1}, \ldots, \tilde{W}_{n,n}\tilde{\epsilon}_{n,n})'$. The robust standard error Stata outputs is the square root of the diagonal elements of

$$
\begin{aligned}
&\Big(\frac{1}{n}\sum_{i=1}^{n}\tilde{W}_{n,i}\tilde{W}'_{n,i}\Big)^{-1}\Big(\frac{1}{n-m-l}\sum_{i=1}^{n}\tilde{W}_{n,i}\tilde{\epsilon}^2_{n,i}\tilde{W}'_{n,i}\Big)\Big(\frac{1}{n}\sum_{i=1}^{n}\tilde{W}_{n,i}\tilde{W}'_{n,i}\Big)^{-1}/n\\
=&(\tilde{W}'\tilde{W})^{-1}(\widetilde{W\epsilon}'\widetilde{W\epsilon})(\tilde{W}'\tilde{W})^{-1}\Big(\frac{n}{n-m-l}\Big).
\end{aligned}
$$