Problem Set #2

Danny Edgel
Econ 715: Econometric Methods
Fall 2021

November 7, 2021

Question 1

(a) To determine $Pr\left(\sum_{i=1}^{n}(X_i^* - \overline{X}_i^*)^2 = 0|\{W_i\}\right)$, we need only find $Pr\left(\sum_{i=1}^{n} X_i^* - \overline{X}_i^* = 0|\{W_i\}\right)$. Since $X_i \in \{0,1\}$, $X_i^* \in \{0,1\}$, so:

$$Pr\left(X_i^* - \overline{X}_i^* = 0|\{W_i\}\right) = Pr\left(X_i^* = X_j^* \forall i,j|\{W_i\}\right)$$

Where $Pr(X_i^* = X_j^*) \geq \frac{1}{2}\ \forall i,j$. Then,

$$Pr\left(X_i^* = X_j^* \forall i,j|\{W_i\}\right) = \left(\frac{1}{2}\right)^n$$

Where, for finite n, $2^{-n} > 0$.

(b) If the conditions for the Lindeberg CLT are satisfied, then $\hat{\theta}_n^*$ converges to the same distribution as $\hat{\theta}_n$. Thus, we must show that

$$\sup \frac{1}{n}\sum_{i=1}^{n}|X_i(s)|^{2+\delta} < \infty$$

For some $\delta > 0$. By the domain of X, this is necessarily satisfied.

(c) Since $\Phi^{-1}(\cdot)$ is a constant,

$$\frac{F^{-1}_{\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)|\{W_i\}}(0.75|\{W_i\}) - F^{-1}_{\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)|\{W_i\}}(0.25|\{W_i\})}{\Phi^{-1}(0.75) - \Phi^{-1}(0.25)} \rightarrow_p \frac{F^{-1}_{\mathcal{N}(0,\sigma_u^2/\sigma_X^2)}(0.75) - F^{-1}_{\mathcal{N}(0,\sigma_u^2/\sigma_X^2)}(0.25)}{\Phi^{-1}(0.75) - \Phi^{-1}(0.25)}$$

Thus,

$$v_n^{iqr} \rightarrow_{a.s.} \frac{\frac{\sigma_u}{\sigma_X}\left(\Phi^{-1}(0.75) - \Phi^{-1}(0.25)\right)}{\Phi^{-1}(0.75) - \Phi^{-1}(0.25)} = \sigma_u/\sigma_X$$

(d) First, we can solve for $Var(\hat{\theta}_n^\dagger|\{W_i\})$:

$$
\begin{aligned}
\mathbb{E}\left[\hat{\theta}_n^\dagger|\{W_i\}\right] &= \hat{\theta}_n + \frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2}\mathbb{E}\left[\varepsilon_i|\{W_i\}\right] \\
&= \hat{\theta}_n + \frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2} \\
\mathbb{E}\left[\hat{\theta}_n^\dagger|\{W_i\}\right]^2 &= \hat{\theta}_n^2 + 2\hat{\theta}_n \frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2} + \left(\frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2}\right)^2 \\
(\hat{\theta}_n^\dagger)^2 &= \hat{\theta}_n^2 + 2\hat{\theta}_n \frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i\varepsilon_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2} + \left(\frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i\varepsilon_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2}\right)^2 \\
Var(\hat{\theta}_n^\dagger|\{W_i\}) &= \mathbb{E}\left[(\hat{\theta}_n^\dagger)^2|\{W_i\}\right] - \mathbb{E}\left[\hat{\theta}_n^\dagger|\{W_i\}\right]^2 \\
&= \left(\frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2}\right)^2 \mathbb{E}\left[\varepsilon_i^2|\{W_i\}\right] \\
&= \left(\frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2}\right)^2
\end{aligned}
$$

By the law of large numbers, this converges to its unconditional expectation. By assumption,

$$
\left(\frac{\sum_{i=1}^n (X_i - \overline{X}_n)\hat{u}_i}{\sum_{i=1}^n (X_i - \overline{X}_n)^2}\right)^2 \rightarrow_{a.s.} \frac{\mathbb{E}\left[(X - \mathbb{E}[X])^2\right]}{\mathbb{E}\left[(X - \mathbb{E}[X])^2\right]^2}\mathbb{E}\left[u^2\right] = \sigma_u^2/\sigma_X^2
$$

(e) The results of the simulation are below. See edgel_ps2.jl and functions.jl for the code that generates them.

	Mean	SD	Coverage
$s.e.^{asy}$	0.322	0.089	0.907
$s.e.^{bt}$	0.152	0.075	0.595
$s.e.^{IQR}$	0.366	0.071	0.964
$s.e.^{bt-r}$	0.052	0.027	0.218
$s.e.^{sbt-normal}$	0.111	0.070	0.443
$s.e.^{sbt-rad}$	0.111	0.070	0.441

Question 2

(a) The inequalities for this data generating process are:

$$
\begin{aligned}
[g(0,0) - g(1,0)](-1) &\geq 0 \\
[g(0,0) - g(1,1)](-1-\theta) &\geq 0 \\
[g(0,0) - g(0,1)](-\theta) &\geq 0 \\
[g(1,0) - g(0,0)] &\geq 0 \\
[g(1,0) - g(0,1)](1-\theta) &\geq 0 \\
[g(1,0) - g(1,1)](-\theta) &\geq 0 \\
[g(1,1) - g(1,0)](\theta) &\geq 0 \\
[g(1,1) - g(0,0)](1+\theta) &\geq 0 \\
[g(1,1) - g(0,1)] &\geq 0 \\
[g(0,1) - g(0,0)](\theta) &\geq 0 \\
[g(0,1) - g(1,1)](-1) &\geq 0
\end{aligned}
$$

Note that the trivial inequalities ($0 \geq 0$) are excluded. The unique set of inequalities, letting $g(x_1, x_2) = F(x_1 + \theta x_2)$, is:

$$
\begin{aligned}
[F(\theta) - F(0)]\theta &\geq 0 && (1) \\
[F(1) - F(0)] &\geq 0 && (2) \\
[F(1+\theta) - F(0)](1+\theta) &\geq 0 && (3) \\
[F(1) - F(\theta)](1-\theta) &\geq 0 && (4) \\
[F(1+\theta) - F(1)]\theta &\geq 0 && (5) \\
[F(1+\theta) - F(\theta)] &\geq 0 && (6)
\end{aligned}
$$

Note that this set of inequalities holds for all values of θ.

(b) If $\theta = 1$, then

$$
\begin{aligned}
&g(1,0) = g(0,1) = F(1) \\
&g(0,0) = F(0) \\
&g(1,1) = F(2)
\end{aligned}
$$

Thus, the only useful moment inequalities for identification yield:

$$\hat{\Theta}_n = \{\theta : \theta \geq 0\}$$

(c) In this case, any variation in Y is due entirely to variation in ε. Then, for each (x_1, x_2), $Var(Y) = Var(\varepsilon)$ and $\mathbb{E}\left[Y|(x_1, x_2)\right] = F(x_1 + \theta x_2)$. Thus, the standard error of $\hat{\theta}_n$ is $\frac{1}{n}\widehat{Var(Y)}$, where $\widehat{Var(Y)}$ is asymptotically normal. Then, letting c^*_α be the critical value of the normal distribution for α, the confidence interval for α is:

$$\left[F_\varepsilon^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} Y(0,1)\right) - c^*_\alpha \widehat{Var(Y)}, F_\varepsilon^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} Y(0,1)\right) + c^*_\alpha \widehat{Var(Y)}\right]$$