Econ241b: PS 6

Casey O'Hara

2018-03-17

1. Asymptotic Behavior of Hypothesis Tests

   Your fieldwork requires that you test $H_0 : \beta_k = \overline{\beta}_k$ vs. $H_1 : \beta_k \neq \overline{\beta}_k$

   (a) You know that under H_0,

   $$\sqrt{n}(B_k - \overline{\beta}_k) \xrightarrow{d} N(0, Avar(B_k))$$
   $$\widehat{Avar(B_k)} \xrightarrow{p} Avar(B_k)$$

   where $\widehat{Avar(B_k)} = S_{xx}^{-1}\hat{S}S_{xx}^{-1}$. To perform your test, you wish to use a t-statistic and you want to make sure that the limit distribution is standard normal. Show that (1) implies that under H_0,

   $$\frac{B_k - \overline{\beta}_k}{SE^*(B_k)} \xrightarrow{d} N(0, 1),$$

   where $SE^*(B_k) = \sqrt{\frac{1}{n}\widehat{Avar(B_k)}}$.

   Assume S_{xx} and $\hat{S}$ are invertible ($\frac{1}{n}\sum x_t x_t'$ is non-singular, by Large Sample Dist class notes, assumption 2.4) thus $\widehat{Avar(B_k)}^{-1}$ exists.

   $$\begin{aligned}
   &\text{Given } \widehat{Avar(B_k)} \xrightarrow{p} Avar(B_k) \\
   &\Rightarrow \widehat{Avar(B_k)}^{-1} \xrightarrow{p} Avar(B_k)^{-1} && \text{(assuming } Avar(B_k)^{-1} \text{ exists)} \\
   &\Rightarrow \widehat{Avar(B_k)}^{-1/2} \xrightarrow{p} Avar(B_k)^{-1/2} && \text{(continuous xform on both sides)}
   \end{aligned}$$

   $$\begin{aligned}
   \frac{B_k - \overline{\beta}_k}{SE^*(B_k)} &= (\sqrt{n}(B_k - \overline{\beta}_k))(\widehat{Avar(B_k)}^{-1/2}) && \text{(algebra)} \\
   \frac{B_k - \overline{\beta}_k}{SE^*(B_k)} &\xrightarrow{d} N(0, Avar(B_k))Avar(B_k)^{-1/2} && \text{(Slutsky's theorem)} \\
   \Rightarrow \frac{B_k - \overline{\beta}_k}{SE^*(B_k)} &\xrightarrow{d} N(0, 1) && \text{(algebra)} \blacksquare
   \end{aligned}$$

   (b) When performing the test, under what circumstances would you select the critical values from a normal distribution versus a t distribution?

   Typically t distribution would be used when the population error variance is not known, and under small sample sizes. Normal distribution is appropriate when the population variance is known. For large n, the t distribution approaches the normal distribution (t distribution with infinite degrees of freedom = normal). Here, though we don't know the population error variance, since we are looking at large samples and asymptotic distributions, we should be able to use the normal distribution.

   (c) Show what $SE^*(B_k)$ converges in probability to.

   We are given that $SE^*(B_k) = \sqrt{\frac{1}{n}\widehat{Avar(B_k)}}$. We know that $\widehat{Avar(B_k)} \xrightarrow{p} Avar(B_k)$, so $\sqrt{\widehat{Avar(B_k)}} \xrightarrow{p} \sqrt{Avar(B_k)}$ (shown in part a. We also know that $\lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0$.

So $SE^*(B_k) = \sqrt{\frac{1}{n}\widehat{Avar(B_k)}}$ converges to 0 as n increases.

(d) With your answer from part c, intuitively explain the convergence in (2).

We've shown that $\frac{B_k - \overline{\beta}_k}{SE^*(B_k)}$ converges to $N(0,1)$ distribution. But according to c, the denominator converges to 0! So the numberator must converge to zero as well: B_k approaches $\overline{\beta}_k$).

Furthermore, they must both converge at the same rate: if the numerator converged toward zero faster, the fraction would converge to zero, and if it converged more slowly, the fraction would not converge.

2. Consistency and Conditional Mean Independence

You want to estimate the following scalar equation (in deviation-from-means form):

$$y_t = \beta x_t + u_t.$$

There are two potentially important assumptions you can make:
(A1): $\mathbb{E}[x_t u_t] = 0$ for all t, and $\mathbb{E}[x_t^2] = \sigma_{xx} < \infty$
(A2): $\mathbb{E}[u_t|X] = 0$ for all t, and $\mathbb{E}[x_t^2] = \sigma_{xx} < \infty$

(a) Show that, under (A1), the OLS estimator is a consistent estimator for β.

OLS estimator:

$$\hat{\beta} = (\sum_{t=1}^{T} x_t x_t')^{-1} \sum_{t=1}^{T} x_t y_t = \beta + (\sum_{t=1}^{T} x_t x_t')^{-1} \sum_{t=1}^{T} x_t u_t$$

By assumptions (A1), we know our variables are stationary and ergodic. By the ergodic theorem,

$$\overline{Z}_n = \frac{1}{n}\sum_{t=1}^{n} Z_t \xrightarrow{as} \mathbb{E}[Z_t]$$

If we rewrite our OLS estimator as:

$$\hat{\beta} = \beta + \left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\frac{1}{T}\sum_{t=1}^{T} x_t u_t\right)$$

then

$$\frac{1}{T}\sum_{t=1}^{T} x_t x_t' \xrightarrow{as} \mathbb{E}[x_t^2] = \sigma_{xx} < \infty$$

and

$$\frac{1}{T}\sum_{t=1}^{T} x_t u_t \xrightarrow{as} \mathbb{E}[x_t u_t] = 0$$

Therefore,

$$\hat{\beta} = \beta + \left(\frac{1}{T}\sum_{t=1}^{T} x_t x_t'\right)^{-1} \left(\frac{1}{T}\sum_{t=1}^{T} x_t u_t\right) \xrightarrow{as \Rightarrow p} \beta + \frac{0}{\sigma_{xx}} = \beta$$

So under assumption (A1), OLS estimator $\hat{\beta}$ is a consistent estimator for β.

(b) Your coauthor suggests you run a first-difference model in which your outcome of interest is no longer the level of y in period t, but rather the change in y from time period $t-1$ to t.

Denote $\Delta y_t = y_t - y_{t-1}$ and $\Delta x_t = x_t - x_{t-1}$. Will an OLS regression of Δy_t on Δx_t yield a consistent estimator of β under (A1)? Will it yield a consistent estimator under (A2)?

In this case, our OLS estimator becomes

$$\hat{\beta} = \left(\sum_{t=1}^{T} \Delta x_t^2\right)^{-1} \sum_{t=1}^{T} \Delta x_t \Delta y_t = \beta + \left(\sum_{t=1}^{T} \Delta x_t^2\right)^{-1} \sum_{t=1}^{T} \Delta x_t \Delta u_t$$

Expanding the delta terms back into t and $t-1$ terms, we want to show (as in part a) that

$$\frac{\sum_{t=1}^{T} \Delta x_t \Delta u_t}{\sum_{t=1}^{T} \Delta x_t^2} \to 0$$

$$\begin{aligned}
\frac{\sum_{t=1}^{T} \Delta x_t \Delta u_t}{\sum_{t=1}^{T} \Delta x_t^2} &= \frac{\sum_{t=1}^{T} (x_t - x_{t-1})(u_t - u_{t-1})}{\sum_{t=1}^{T} \Delta x_t^2} \\
&= \frac{\sum_{t=1}^{T} (x_t u_t - x_{t-1} u_t - x_t u_{t-1} + x_{t-1} u_{t-1})}{\sum_{t=1}^{T} \Delta x_t^2} \\
&= \frac{\frac{1}{n}\sum_{t=1}^{T} x_t u_t - \frac{1}{n}\sum_{t=1}^{T} x_{t-1} u_t - \frac{1}{n}\sum_{t=1}^{T} x_t u_{t-1} + \frac{1}{n}\sum_{t=1}^{T} x_{t-1} u_{t-1}}{\sum_{t=1}^{T} \Delta x_t^2}
\end{aligned}$$

Expand out the terms on the bottom, by ergodic theorem:

$$\sum_{t=1}^{T} (x_t - x_{t-1})^2 \to \mathbb{E}[(x_t - x_{t-1})^2]$$

which we will assume exists ($< \infty$), and is not zero.

By the ergodic theorem, the terms on top converge almost surely:

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{T} x_t u_t &\to \mathbb{E}[x_t u_t] = 0 \\
\frac{1}{n}\sum_{t=1}^{T} x_{t-1} u_t &\to \mathbb{E}[x_{t-1} u_t] \\
\frac{1}{n}\sum_{t=1}^{T} x_t u_{t-1} &\to \mathbb{E}[x_t u_{t-1}] \\
\frac{1}{n}\sum_{t=1}^{T} x_{t-1} u_{t-1} &\to \mathbb{E}[x_{t-1} u_{t-1} = 0]
\end{aligned}$$

Based on assumption (A1) we don't know the values of $\mathbb{E}[x_{t-1}u_t]$ and $\mathbb{E}[x_t u_{t-1}]$, so we can't tell whether the numerator converges to zero.

Therefore, by assumption (A1), $\hat{\beta}$ is not necessarily a consistent estimator for β.

Under assumption (A2), $\mathbb{E}[u_t|X] = 0$, and using LIE:

$$\begin{aligned}
\mathbb{E}[x_{t-1}u_t] &= \mathbb{E}[\mathbb{E}[x_{t-1}u_t|X]] && \text{(law of iter exp.)} \\
&= \mathbb{E}[x_{t-1}\mathbb{E}[u_t|X]] && \text{(conditioning thm)} \\
&= \mathbb{E}[x_{t-1}0] = 0 && \text{(substitute)}
\end{aligned}$$

The same process would work to show that $\mathbb{E}[x_t u_{t-1}] = 0$. So under assumption (A2), $\hat{\beta}$ is consistent for β.

(c) Imagine you and your coauthor are working on a project in which you observe individuals (indexed by i) in two time-periods, $t = 1$ and $t = 2$. You observe outcomes ($y_{i,t}$) and a mean-zero scalar regressor ($x_{i,t}$) for each individual in both time periods. There also exists a characteristic specific to each individual that does not change over time (α_i) and that you cannot observe, yet it influences the outcome of interest. You and your coauthor want to estimate the following scalar equation

$$y_{i,t} = \alpha_i + x_{i,t} + u_{i,t}.$$

The assumptions that correspond to A1 and A2 are, for each value of i:
(A1'): $\mathbb{E}[x_{i,t}u_{i,t}] = 0$ for all t, and $\mathbb{E}[x_{i,t}^2] = \sigma_{xx} < \infty$
(A2'): $\mathbb{E}[u_{i,t}|X] = 0$ for all t, and $\mathbb{E}[x_{i,t}^2] = \sigma_{xx} < \infty$
If α_i is uncorrelated with $x_{i,t}$ and $x_{i,t-1}$, will the OLS estimator be a consistent estimator for β under (A1')?

Set up the OLS estimator: how does α_i fit in?

$$\begin{aligned}\hat{\beta} &= \frac{\sum_{i=1}^n x_{i,t}y_{i,t}}{\sum_{i=1}^n x_{i,t}^2} \\ &= \beta + \frac{\sum_{i=1}^n x_{i,t}(u_{i,t} + \alpha_i)}{\sum_{i=1}^n x_{i,t}^2}\end{aligned}$$

And each of these sums would be additionally summed for periods $t = 1$ and $t = 2$.

For consistency, we need:

$$\frac{\sum_{t=1}^2 \left(\sum_{i=1}^n x_{i,t}u_{i,t}\right) + \sum_{t=1}^2 \left(\sum_{i=1}^n x_{i,t}\alpha_i\right)}{\sum_{t=1}^2 \sum_{i=1}^n x_{i,t}^2} = 0$$

Using the same ideas as prior problems, and multiplying both top and bottom of this fraction by $\frac{1}{n}$, the ergodic theorem tells us that the bottom will be positive and finite, and the numerator terms:

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{2}\sum_{i=1}^{n} x_{i,t}u_{i,t} &\xrightarrow{p} \sum_{t=1}^{2}\mathbb{E}[x_{i,t}u_{i,t}] = 0 \\ \frac{1}{n}\sum_{t=1}^{2}\sum_{i=1}^{n} x_{i,t}\alpha_i &\xrightarrow{p} \sum_{t=1}^{2}\mathbb{E}[x_{i,t}\alpha_i] \\ &= \sum_{t=1}^{2}\mathbb{E}[x_{i,t}]\mathbb{E}[\alpha_i] && (x_{i,t}, \alpha_i \text{ uncorrelated}) \\ &= \sum_{t=1}^{2} 0 \times \mathbb{E}[\alpha_i] && (x_{i,t} \text{ is mean zero}) \\ &= 0 && \blacksquare\end{aligned}$$

Therefore, $\hat{\beta}$ is consistent for β, under assumption (A1').

(d) Suppose you have reason to believe that the individual characteristic may in fact be correlated with the regressor of interest. In light of your answer in part b, can you construct a consistent estimator for β under (A1')? Can you construct a consistent estimator for β under (A2')?

Under (A1'), if $x_{i,t}, \alpha_i$ are correlated, then the $\sum_{t=1}^2 \mathbb{E}[x_{i,t}\alpha_i]$ term does not necessarily equal zero, so we cannot show consistency.

But the α_i term is not time-dependent, and so if we do a first-difference model as in part b, it will appear equally in both the $t = 1$ period and $t = 2$, and thus subtract out. Due to the parallelism of the assumptions (A1) and (A1'), as in *b*, the (A1') assumption will not get us to consistency in this case, but the (A2') assumption should, by parallel logic to part *b* and parallelism between assumptions (A2) and (A2').