Economics 672: Econometric Analysis II
Winter 2018

Matias Cattaneo
Department of Economics
University of Michigan
Authored collaboratively by PhD entering class of 2017

January 25, 2018

Outline

1. Introduction
2. Asymptotic Properties of LS
3. Next Lecture: Topics and Readings

Introduction

- Last lecture:
    - GLS
    - Intro to LS Asymptotics
- This lecture:
    - More LS Asymptotics

Outline

1. Introduction
2. Asymptotic Properties of LS
3. Next Lecture: Topics and Readings

Consistency

- Our standard set-up

$$\hat{\beta}_{LS} = \beta_0 + \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} x_i \varepsilon_i\right) = \beta_0 + \left(\mathbb{E}[x_i x_i']\right)^{-1}\left(\mathbb{E}[x_i \varepsilon_i]\right) + o_{\mathbb{P}}(1)$$

- The second equality requires regularity conditions
- "Weak instrument" type problems involve the denominator converging in probability to 0

Asymptotic Distribution

- By centering and scaling, we get:

$$\sqrt{n}\left(\hat{\beta}_{LS}-\beta_0\right)=\sqrt{n}(X'X)^{-1}X'\varepsilon=\left(\frac{1}{n}\sum_{i=1}^{n}x_ix_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}x_i\varepsilon_i\right)$$

- Where

$$\frac{1}{n}\sum_{i=1}^{n}x_ix_i'\to_p\mathbb{E}[x_ix_i'],\quad\frac{1}{\sqrt{n}}\sum_{i=1}^{n}x_i\varepsilon_i\to_d\mathcal{N}(0,\mathbb{V}[x_i\varepsilon_i])$$

- Then

$$\sqrt{n}\left(\hat{\beta}_{LS}-\beta_0\right)\to_d\mathcal{N}(0,H_0^{-1}V_0H_0^{-1})$$

- Where

$$H_0=\mathbb{E}[x_ix_i'],\quad V_0=\mathbb{V}[x_i\varepsilon_i]=\mathbb{E}[x_ix_i'\mathbb{E}[\varepsilon_i^2|x_i]],\quad\mathbb{E}[\varepsilon_i^2|x_i]=\sigma^2(x_i)$$

Inference under Homoskedasticity

- If we assume homoskedasticity ($\sigma^2(x_1) = \sigma^2$), then we get $V_0 = \sigma^2 H_0$ and the asymptotic variance of $\hat{\beta}_{LS}$ simplifies.
- We can then standardize as follows:

$$\left(\frac{\sigma^2 H_0}{n}\right)^{-1/2} \left(\hat{\beta}_{LS} - \beta_0\right) \to_d \mathcal{N}(0, I_d)$$

- Which means we can use the following plug-in estimates to conduct asymptotic inference:

$$\hat{H} = \frac{1}{n}\sum_{i=1}^{n} x_i x_i', \quad \hat{\sigma}^2 \frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_i^2$$

- The problem here is that $\hat{\sigma}^2$, although consistent, is biased for σ^2, because it is not scaled by the appropriate (finite sample) degrees of freedom. This is because we derived the asymptotic distribution first (disregarding d.o.f.), and then determined the sample analogues that we needed to conduct inference.

$$H_0 = \mathbb{E}[x_i x_i'], \quad V_0 = \mathbb{V}[x_i \varepsilon_i] = \mathbb{E}[x_i x_i' \mathbb{E}[\varepsilon_i^2 | x_i]], \quad \mathbb{E}[\varepsilon_i^2 | x_i] = \sigma^2(x_i)$$

"Robust" Standard Errors

- Huber-Eicher-White: use $\hat{\Sigma} = \hat{H}^{-1}\hat{V}\hat{H}^{-1}$ to approximate $AsyVar(\hat{\beta}_{LS}) = H_0^{-1}V_0H_0^{-1}$. How you correct for heteroskedasticity is a matter of how you calculate $\hat{V}$.
- Although there is no theoretical motivation, Stata's default uses:

$$\hat{V} = \hat{\mathbb{E}}[x_i x_i'(y_i - x_i'\hat{\beta}_{LS})^2]$$

- The following estimator is unbiased under homoskedasticity ($\sigma^2(x_i) = \sigma^2 \Rightarrow \mathbb{E}[\hat{V}_{HC2}] = V_0$):

$$\hat{V}_{HC2} = \frac{1}{n}\sum_{i=1}^{n} x_i x_i' \frac{1}{p_{ii}}(y_i - x_i'\hat{\beta}_{LS})^2 \quad (p_{ii} = [P_X]_{ii})$$

However, $\hat{V}_{HC2}$ requires a lot of the computer because it has to form the projection matrix.

"Robust" Standard Errors

- The following estimator is motivated by the "jackknife" estimator of the asymptotic variance from the problem set, which equals $\hat{H}^{-1}\hat{V}_{HCE}\hat{H}^{-1}$. Estimate $\hat{V}_{HCE}$ as follows:

$$\hat{V}_{HC3} = \frac{1}{n}\sum_{i=1}^{n} x_i x_i' \frac{1}{p_{ii}^2}(y_i - x_i'\hat{\beta}_{LS})^2 \Rightarrow \hat{\Sigma}_{HC3} = \frac{n-1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_{(i)} - \hat{\beta}_{LS}\right)\left(\hat{\beta}_{(i)} - \hat{\beta}_{LS}\right)'$$

- The following estimator is biased upward (so will produce standard errors that are always conservative, i.e. "too large")

$$\hat{\Sigma}_{JACK} = \frac{n-1}{n}\sum_{i=1}^{n}\left(\hat{\beta}_{(i)} - \hat{\beta}_{(\cdot)}\right)\left(\hat{\beta}_{(i)} - \hat{\beta}_{(\cdot)}\right)'$$

- All of these estimators of V_0 are heteroskedasticity-consistent so long as we're talking about *conditional* heteroskedasticity, i.e. $\mathbb{V}(\varepsilon|x_i) = \sigma^2(x_i)$. Unconditional heteroskedasticity takes us away from iid data and into the world of clustering (dependence across observations).

Outline

1 Introduction

2 Asymptotic Properties of LS

3 Next Lecture: Topics and Readings

Next Class Topics and Readings

- Problem Set #
- Next topic: blah blah blah
- **Readings:**
  1. books
  2. more books