

Economics 672: Econometric Analysis II

Winter 2018

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

January 18, 2018

Outline

- 1 Introduction
- 2 Review of SLRM
- 3 Finite Sample Distributional Results
- 4 Next Lecture: Topics and Readings

Introduction

- Last lecture:
 - Intro to Least Squares
 - Gauss-Markov Theorem
- This lecture:
 - Finite sample results
 - Basic inference

Outline

- 1 Introduction
- 2 Review of SLRM
- 3 Finite Sample Distributional Results
- 4 Next Lecture: Topics and Readings

Standard Linear Regression Model (SLRM)

- Consider $X \in \mathbb{R}^{n \times d}$ and $y \in \mathbb{R}^n$
- The standard assumptions can be written as:

$$\mathbb{E}[y|X] = X\beta_0, \mathbb{V}[y|X] = \sigma_0^2 \mathbf{I}_n$$

- The least squares estimator of β_0 is then:

$$\hat{\beta} = \arg \min_{\beta} \|y - X\beta\|^2 = (X'X)^{-} X'y$$

- With associated predicted values $\hat{y} = P_X y$ and $\hat{\varepsilon} = (I - P_X)y = M_X y$
 - Where $P_X = X(X'X)^{-} X'$, and $(\cdot)^{-}$ denotes the generalized inverse.

Estimating the Variance

- Note that we can write the sum of squared residuals as:

$$\begin{aligned}
 \|\hat{\varepsilon}\|^2 &= \left[\sqrt{\sum_{i=1}^n \hat{\varepsilon}_i^2} \right]^2 = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \hat{\varepsilon}'\hat{\varepsilon} = (M_X y)'(M_X y) \\
 &= \left(M_X(X\hat{\beta} + \varepsilon) \right)' \left(M_X(X\hat{\beta} + \varepsilon) \right) = (M_X \varepsilon)'(M_X \varepsilon) \\
 &= \varepsilon' M_X' M_X \varepsilon = \varepsilon' M_X \varepsilon = \varepsilon'(I - P_X)\varepsilon
 \end{aligned}$$

- Here we used the fact that M_X is the orthogonal projection matrix (hence $M_X X = \mathbf{0}$) and is symmetric and idempotent.

Estimating the Variance

- Taking the expectation gives a useful result:

$$\begin{aligned}\mathbb{E}[\|\hat{\varepsilon}\|^2 | X] &= \mathbb{E}[\varepsilon'(I - P_X)\varepsilon | X] = \mathbb{E}[\text{tr}(\varepsilon'(I - P_X)\varepsilon) | X] \\ &= \mathbb{E}[\text{tr}((I - P_X)\varepsilon\varepsilon') | X] = \text{tr}((I - P_X)\mathbb{E}[\varepsilon\varepsilon' | X]) \\ &= \text{tr}(I - P_X)\sigma_0^2 = (n - d)\sigma_0^2\end{aligned}$$

- Here we used the fact that the trace operator is order invariant and linear.
- This result means that an unbiased estimator of σ_0^2 is given by:

$$s^2 = \frac{\|\hat{\varepsilon}\|^2}{n - d} = \frac{1}{n - d} \sum_{i=1}^n \hat{\varepsilon}_i^2$$

Outline

- 1 Introduction
- 2 Review of SLRM
- 3 Finite Sample Distributional Results**
- 4 Next Lecture: Topics and Readings

Adding Normality

- Adding the assumption that $y|X \sim \mathcal{N}(X\beta_0, \sigma_0^2 \mathbf{I}_n)$ allows us to derive several useful distributional results.
- Using that $\mathbb{E}[\hat{y}|X] = X\beta_0$ and $\mathbb{V}[\hat{y}|X] = \sigma_0^2 P_X$, we can use the fact that \hat{y} is a linear combination of normally distributed random variables to deduce that:

$$\hat{y}|X \sim \mathcal{N}(X\beta_0, \sigma_0^2 P_X)$$

- Similarly, we have:

$$\hat{\varepsilon}|X \sim \mathcal{N}(\mathbf{0}, \sigma_0^2(I - P_X))$$

Adding Normality

- We can write the variance estimator as:

$$s^2 = \frac{\sigma_0^2}{n-d} \frac{1}{\sigma_0} \varepsilon' (I - P_X) \varepsilon \frac{1}{\sigma_0}$$

- Since $I - P_X$ is symmetric, we can use a singular value decomposition to write it as $I - P_X = U \Lambda U'$, where Λ is an $n \times n$ diagonal matrix of eigenvalues ($n-d$ of which are 1 and d of which are 0) while U is an orthogonal matrix.
- This brings us to the following result:

$$s^2 = \frac{\sigma_0^2}{n-d} \left[\frac{1}{\sigma_0} U' \varepsilon \right]' \Lambda \left[\frac{1}{\sigma_0} U' \varepsilon \right] \implies s^2 | X \sim \frac{\sigma_0^2}{n-d} \chi_{n-d}^2$$

- Here, $\frac{1}{\sigma_0} U' \varepsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, and the specific form of Λ means that we are summing $n-d$ squares of standard normal random variables, yielding the Chi-square distribution (with corresponding d.o.f.)

Inference

- Consider the following hypothesis:

$$H_0 : A'\beta = a, A \in \mathbb{R}^{d \times r}$$

- And the following statistics:

$$T_1 = \frac{[A'\hat{\beta} - a]'(A'(X'X)^{-1}A)^{-1}[A'\hat{\beta} - a]}{\sigma_0^2} \sim \chi_r^2, T_2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{\sigma_0^2} \sim \chi_{n-d}^2$$

- Normality allows us to conclude that $T_1 \perp\!\!\!\perp T_2$
- Then, taking the ratio of independent chi-square random variables gives another well known distribution:

$$F = \frac{T_1/r}{T_2/(n-d)} \sim \mathcal{F}_{r,n-d}$$

- In the case $r = 1$, we have $F = (T_{n-d})^2$

Outline

- 1 Introduction
- 2 Review of SLRM
- 3 Finite Sample Distributional Results
- 4 Next Lecture: Topics and Readings

Next Class Topics and Readings

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