EC2020 Session 5

tomglinnan.github.io/EC2020

t.m.glinnan@lse.ac.uk

$$y = \beta_0 + \beta_1 x_1 \dots + \beta_k x_k + u$$

$\mathbb{E}(u|X) = 0$

MLR assumptions

Gauss-Markov Assumptions

MLR 1 ✓

MLR 2 ✓

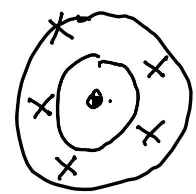
MLR 3 ✓ → no multicollinearity

MLR 4 ✓

MLR 5 ✓ → homoscedasticity

$\hat{\beta}$ is unbiased → "CORRECT on average"

$\mathbb{E}(\hat{\beta}) = \beta$



$\hat{\beta}$ is consistent

↳ As n (amount of data) → ∞

$\hat{\beta} \xrightarrow{p} \beta \Leftrightarrow \text{plim } \hat{\beta} = \beta$

1) $y_i = \beta_1 x_i + u_i$ and assume MLR.1 – MLR.4 all hold

DGP

↳ OLS unbiased

↳ consistent

$\tilde{\beta}_1$ is the OLS estimator

Prove that $\tilde{\beta}_1$ is consistent

1. Write down the estimator $\tilde{\beta}_1$
2. plug in the DGP "data generating process"
3. If a consistency proof, use LLN.

In general the "remainder" part is $\simeq 0$

$$\tilde{\beta}_1 \overset{1}{=} \frac{\sum x_i y_i}{\sum x_i^2}$$

$$\overset{2}{=} \frac{\sum x_i(\beta_1 x_i + u_i)}{\sum x_i^2}$$

$$= \frac{\beta_1 \sum x_i^2 + \sum x_i u_i}{\sum x_i^2}$$

$$= \beta_1 + \frac{\frac{1}{n}\sum x_i u_i}{\frac{1}{n}\sum x_i^2}$$

$$\mathbb{E}\left(\frac{A}{B}\right) \neq \frac{\mathbb{E}(A)}{\mathbb{E}(B)}$$

$$\boxed{\frac{1}{n}\sum f(x_i) \xrightarrow{p} \mathbb{E}[f(x_i)]}$$

the LLN

$$\text{plim}\,\tilde{\beta}_1 = \beta_1 + \frac{\text{plim}\,\frac{1}{n}\sum x_i u_i}{\text{plim}\,\frac{1}{n}\sum x_i^2}$$

Slutsky's theorem

$$\text{plim}\,\tilde{\beta}_1 = \beta_1 + \frac{\text{plim}\,\frac{1}{n}\sum x_i u_i}{\text{plim}\,\frac{1}{n}\sum x_i^2} \overset{LLN}{=} \beta_1 + \frac{\mathbb{E}\,x_i u_i}{\mathbb{E}\,x_i^2} = \beta_1 + \frac{0}{\mathbb{E}\,x_i^2} = \beta_1 \quad \square$$

MLR.3: $\mathbb{E}\,x_i^2 \neq 0$

MLR.4: $\mathbb{E}(u_i | x_i) = 0$

$\Rightarrow \mathbb{E}(u_i x_i) = 0$ Law of Iterated Expectations (LIE)

In math:

$X_n \to X$

g continuous

then $g(X_n) \to g(X)$

. . . .

In econometrics

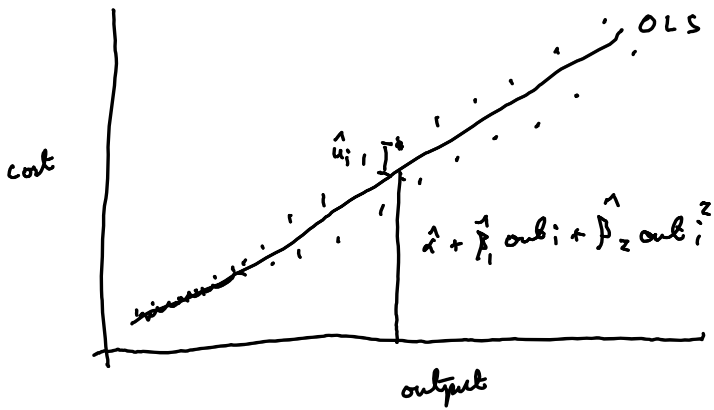
$X_n \xrightarrow{p} X$

g continuous

then $g(X_n) \xrightarrow{p} g(X)$

2a) Out_i

$Cost_i$



$$Cost_i = \underline{\alpha + \beta_1 out_i + \beta_2 out_i^2} + \underline{u_i}$$

1-sided $H_0: \beta_1 = 0$

$H_1: \beta_1 > 0$

2-sided $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

"T test"

i) $Var(u_i)$ not a constant

is a function of output

ii) consequences $se(\hat{\beta})$ no longer valid

$$\frac{\hat{\beta} - \beta}{se(\hat{\beta})}$$

$$F = \frac{RSS - URSS/\ldots}{URSS/\ldots}$$

"F test"

H_0: all $\beta_i = 0$

H_1: not H_0

H_0: $\underbrace{\text{Homo}}_{\text{same}}\underbrace{\text{scedasticity}}_{\text{Var}}$

Testing for heteroscedasticity

$Var(u_i) = \sigma^2 \quad \forall i$

H_1:

$\boxed{\text{B.P.}}$ $\boxed{\text{white}}$

$H_0: Var(u_i) = \sigma^2$

<u>BP</u>: $H_1: Var(u_i) = \delta_0 + \delta_1 \, out_i + \delta_2 \, out_i^2$

for <u>some</u> (unknown) $(\delta_0, \delta_1, \delta_2)$

<u>white</u>: H_0: $Var(u_i) = \sigma^2$

H_1: $Var(u_i) \neq \sigma^2$

$Var(u_i) = \mathbb{E}(u_i^2) \cancel{- (\mathbb{E} u_i)^2}$ as $\mathbb{E}(u_i) = 0$

<u>BP</u>: H_0: $\mathbb{E}(u_i^2) = \sigma^2$

H_1: $\mathbb{E}(u_i^2) = \delta_0 + \delta_1 \, out_i + \delta_2 \, out_i^2$

$u_i = y_i - x_i \beta$

$\hat{u}_i = y_i - x_i \hat{\beta}$

Imagine I could do the regression in H_1

If δ_1 and $\delta_2 = 0$ then homoscedastic

$$\hat{u}_i^2 = \delta_0 + \delta_1 \, out_i + \delta_2 \, out_i^2 + \varepsilon_i$$

(1) get $\hat{u}_i$: run $cost_i = \alpha + \beta_1 out_i + \beta_2 out_i^2 + u_i$

↳ gives $\hat{\alpha}, \hat{\beta}_1, \hat{\beta}_2$

↳ therefore can calculate $\hat{u}_i$

(2) run $\hat{u}_i^2 = \delta_0 + \delta_1 cost_i + \delta_2 cost_i^2 + \varepsilon_i$

↳ Test if $\delta_1 = \delta_2 = 0$

↳ This is the F-test

white: let $C_i = cost_i$

same stage (1)

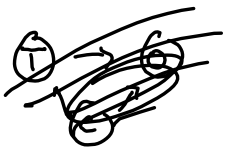
stage (2): $\hat{u}_i^2 = \delta_0 + \delta_1 C_i + \delta_2 C_i^2 + \delta_3 C_i^3 + \delta_4 C_i^4 + \varepsilon_i$

↳ Then F-test of

$\delta_1 = \delta_2 = \delta_3 = \delta_4 = 0$

3) $wage_i = \alpha + \beta_1 female_i + u_i$

↳ $female_i = \begin{cases} 1 & \text{if female} \\ 0 & \text{if not} \end{cases}$



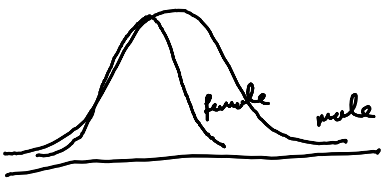
$\hat{\alpha} = 7.10$

$\hat{\beta} = -2.51$

Suppose that $Var(wage \mid males) \neq Var(wage \mid females)$

$\Leftrightarrow$

$Var(u_i \mid female_i = 1) \neq Var(u_i \mid female_i = 0)$



NO MLR.5

$se(\hat{\beta})$ is an estimator of $\sqrt{Var(\hat{\beta})}$

only good under MLR.5

$T = \frac{\hat{\beta} - \beta}{se(\hat{\beta})}$

"use robust SEs"

3b) Suppose no MLR.5

GLS : a special case is WLS

$\mathrm{cov}(u_i, u_j) = 0$

$$\mathrm{Var}(u \mid female) = \cancel{\sigma^2}(1 + 0.2\,female)$$

$$\frac{wage_i}{\sqrt{1+0.2f_i}} = \frac{\beta_0}{\sqrt{1+0.2f_i}} + \beta_1 \frac{female_i}{\sqrt{1+0.2f_i}} + \frac{u_i}{\sqrt{1+0.2f_i}}$$