ECON 61001: *Lecture 3*

Alastair R. Hall

The University of Manchester

- Testing hypothesis about β_0 in the CLR model based on OLS estimators
  - Testing hypotheses about individual coefficients
  - Testing hypotheses about linear combinations of the coefficients
- Restricted Least Squares
- Variable selection

- asset returns: $R - R_f = \beta_0(R_m - R_f) + error$

  - $\beta_0 < 0 \Rightarrow$ stock is inversely related to market index;
  - $\beta_0 = 1 \Rightarrow$ stock moves in line with market index.

- returns to education (dropping exp to simplify and reparameterizing):

$$ln(w) = \beta_{0,1} + \beta_{0,2} * ed + \beta_{0,3} * D + \beta_{0,4} * (D * ed) + error$$

  - $D = 1$ is female and zero else
  - $\beta_{0,3} = 0$, $\beta_{0,4} = 0 \Rightarrow$ no difference between men and women.

- aggregate production function:

$$lnQ = \beta_{0,1} + \beta_{0,2} * ln(L) + \beta_{0,3} * ln(K) + error$$

- $\beta_{0,2} + \beta_{0,3} \left\{ \begin{array}{c} < \\ = \\ > \end{array} \right\} 1 \Rightarrow \left\{ \begin{array}{c} \text{diminishing} \\ \text{constant} \\ \text{increasing} \end{array} \right\}$ returns to scale

Testing hypotheses about $\beta_{0,i}$

Consider inference about $\beta_{0,i}$ based on $\hat{\beta}_{T,i}$.

Recall that we showed last lecture that

$$\frac{\hat{\beta}_{T,i} - \beta_{0,i}}{\sigma_0 \sqrt{m_{i,i}}} \sim N(0, 1),$$

where $m_{i,i}$ is the i^{th} main diagonal element of $(X'X)^{-1}$,

and noted that if we replace σ_0 by $\hat{\sigma}_T$ then,

$$\frac{\hat{\beta}_{T,i} - \beta_{0,i}}{\hat{\sigma}_T \sqrt{m_{i,i}}} \sim \text{Student's t distribution with T-k df}$$

Inference about $\beta_{0,i}$

Consider the two-sided test: $H_0 : \beta_{0,i} = \beta_{*,i}$ vs. $H_1 : \beta_{0,i} \neq \beta_{*,i}$.

Natural to base test statistic on:

$$\hat{\tau}_{T,i}(\beta_{*,i}) = \frac{\hat{\beta}_{T,i} - \beta_{*,i}}{\hat{\sigma}_T \sqrt{m_{i,i}}}$$

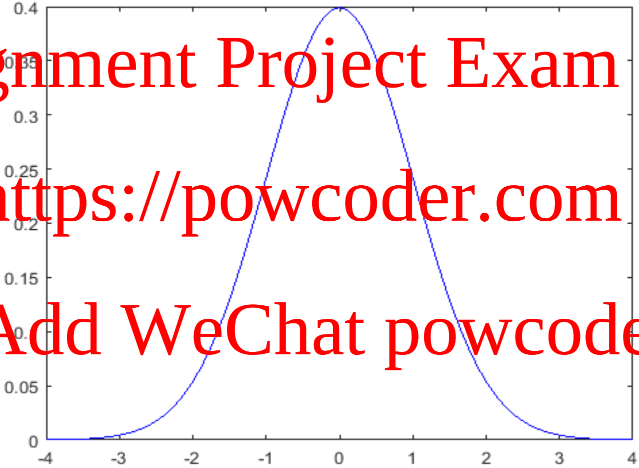
because under H_0:

$$\hat{\tau}_{T,i}(\beta_{*,i}) \sim \text{Student's t distribution with T-k df}$$

Decision rule: reject H_0 at $100\alpha\%$ significance level if

$$|\hat{\tau}_{T,i}(\beta_{*,i})| > \tau_{T-k}(1-\alpha/2)$$

Note: significance level is $100 \times P(\text{Type I error})$.


0.4
0.35
0.3
0.25
0.2
0.15
0.05
0
-4
-3
-2
-1
0
1
2
3
4

Suppose H_0 is false. How does our test statistic behave?

$$\hat{\tau}_{T,i}(\beta_{*,i}) = \frac{\hat{\beta}_{T,i} - \beta_{*,i}}{\hat{\sigma}_T \sqrt{m_{i,i}}} = \frac{\hat{\beta}_{T,i} - \beta_{0,i}}{\hat{\sigma}_T \sqrt{m_{i,i}}} + \frac{\beta_{0,i} - \beta_{*,i}}{\hat{\sigma}_T \sqrt{m_{i,i}}}$$

So under Assumptions *CA1-CA6*:

$$\hat{\tau}_{T,i}(\beta_{*,i}) \sim \text{ Student's t distribution with } T-k \text{ df and ncp } \nu$$

where the non-centrality parameter (ncp) is:

$$\nu = \frac{\beta_{0,i} - \beta_{*,i}}{\sigma_0 \sqrt{m_{i,i}}}$$

As a result:

- $P(\text{reject } H_0 \mid H_1 \text{ true}) > \alpha \Rightarrow$ unbiased test.

- power $\uparrow$ as $|\nu| \uparrow$.
- power depends on $|\nu|$ and not $|\beta_{0,i} - \beta_{*,i}|$ *per se*.

From Lecture 1:

$$\hat{y}_t = \text{controls} - 0.030 * belt_t + 0.0671 * mph_t$$

Did passage of seat belt law affect % of accidents with fatalities?

- $H_0 : \beta_{belt,0} = 0$ (no effect) vs $H_1 : \beta_{belt,0} \neq 0$ (has effect)
- test statistic:

$$|\hat{\tau}_{belt}| = \left| \frac{\hat{\beta}_{belt} - 0}{s.e.(\hat{\beta}_{belt})} \right| = \left| \frac{-0.030}{0.023} \right| = 1.304$$

p-value is 0.195 and so fail to reject at all conventional significance levels.

Example: traffic fatalities

Did passage of seat belt law reduce % of accidents with fatalities?

- $H_0 : \beta_{belt,0} \geq 0$ (no) vs $H_1 : \beta_{belt,0} < 0$ (yes)

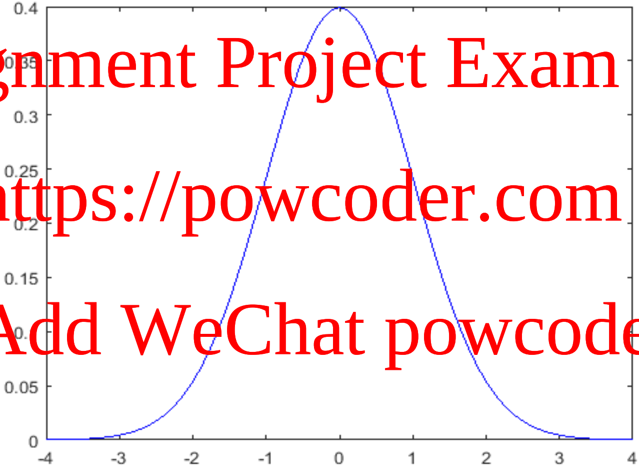
- Example of one-sided test: $H_0 : \beta_{0,i} \geq \beta_{*,i}$ vs $H_1 : \beta_{0,i} < \beta_{*,i}$
- Test statistic is now:

$$\hat{\tau}_{T,i}(\beta_{*,i}) = \frac{\hat{\beta}_{T,i} - \beta_{*,i}}{\hat{\sigma}_T \sqrt{m_{i,i}}}$$

- Decision rule is to reject H_0 in favour of H_1 at the $100\alpha\%$ significance level if

$$\hat{\tau}_{T,i}(\beta_{*,i}) < \tau_{T-k}(\alpha)$$

- In our example, the critical value is -1.291 (-1.662) for the 10% (5%) significance level test and so marginal evidence against H_0.


0.4
0.35
0.3
0.25
0.2
0.15
0.05
0
-4
-3
-2
-1
0
1
2
3
4

Inference about $R\beta_0 = r$

Consider testing: $H_0 : R\beta_0 = r$ vs $H_1 : R\beta_0 \neq r$ where R, r are $n_r \times k$ and $n_r \times 1$ are specified constants.

We need $rank(R) = n_r$ to rule out redundancies.

Natural to base inference on: $R\hat{\beta}_T - r$.

Given sampling distribution of $\hat{\beta}_T$, we have:

$$R\hat{\beta}_T - R\beta_0 \sim N(0, \sigma_0^2 R(X'X)^{-1}R')$$

and so under H_0

$$R\hat{\beta}_T - r \sim N(0, \sigma_0^2 R(X'X)^{-1}R')$$

Test statistic:

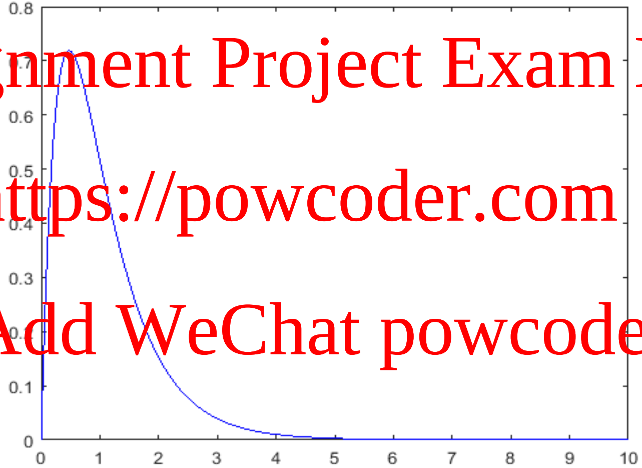
$$F = \frac{(R\hat{\beta}_T - r)'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r)}{n_r\hat{\sigma}_T^2}$$

Under H_0, $F \sim F_{n_r, T-k}$, the F distribution with $(n_r, T-k)$ df.

Decision rule: reject $H_0 : R\beta_0 = r$ at the $100\alpha\%$ significance level if:

$$F > F_{n_r, T-k}(1-\alpha)$$

where $F_{n_r, T-k}(1-\alpha)$ is the $100(1-\alpha)^{th}$ percentile of the F distribution with $(n_r, T-k)$ df.


0.8
0.6
0.5
0.4
0.3
0.1
0
0
1
2
3
4
5
6
7
8
9
10

Alternative representation

$$F = \left(\frac{RSS_R - RSS_U}{RSS_U}\right)\left(\frac{T-k}{n_r}\right)$$

where

- RSS_U is RSS from regression without imposing $R\beta = r$
- RSS_R is RSS from regression imposing $R\beta = r$

Did two traffic laws offset each other?

- H_0 : $\beta_{belt,0} + \beta_{mph,0} = 0$ vs H_1 : $\beta_{belt,0} + \beta_{mph,0} \neq 0$.

- Reject H_0 at the $100\alpha\%$ significance level if $F > F_{1,91}(1-\alpha)$.
- $F = 3.126$, $F_{1,91}(.95) = 3.946 \Rightarrow$ Fail to reject at 5% level.

- p-value $= 0.080$ so reject at 10% significance level.

Do regressors collectively help to explain y?

- $H_0 : R\beta_0 = 0_{k-1}$ (no) vs $H_1 : R\beta_0 \neq 0_{k-1}$ (yes) for

$$R = [0_{k-1} \;\; I_{k-1}]$$

where 0_{k-1} is $(k-1) \times 1$ null vector.

- Decision rule: reject at $100\alpha\%$ significance level if $F > F_{k-1,T-k}(1-\alpha)$.

- In this case, F-test statistic given by:

$$F = \left(\frac{R^2}{1-R^2}\right)\left(\frac{T-k}{k-1}\right)$$

Do regressors (monthly dummies, time trend, unem, wkends, belt, mph) collectively help to explain y (% of traffic accidents involving fatalities)?

- $R^2 = 0.72$
- $F = 14.625$
- $F = F_{16,91}(1-\delta) \Rightarrow$ p-value ≈ 0 and so reject H_0 at all conventional sig. levels.

Restricted Least Squares

Suppose we wish to impose linear restrictions on estimated coefficients. Can do this via method of Restricted Least Squares.

Recall that OLS is: $\hat{\beta}_T \;=\; argmin_{\beta \in \mathcal{B}}\, Q_T(\beta)$.

Define: $\mathcal{B}_R = \{\beta; R\beta = r, \beta \in \mathcal{B}\}$.

Restricted Least Squares (RLS) estimator of β_0 is:

$$\hat{\beta}_{R,T} \;=\; argmin_{\beta \in \mathcal{B}_R}\, Q_T(\beta).$$

Note: $R\hat{\beta}_{R,T} = r$ by construction.

Obtain estimator via Lagrange's method with Lagrangean:

$$\mathcal{L}(\beta, \lambda) \;=\; Q_T(\beta) + 2\lambda'(R\beta \;-\; r)$$

It can be shown that (see Lecture Notes)

$$\hat{\beta}_{R,T} \;=\; \hat{\beta}_T - (X'X)^{-1}R'\{R(X'X)^{-1}R'\}^{-1}(R\hat{\beta}_T \;-\; r).$$

If Assumptions *CA1*- *CA6*, $rank(R) = n_r$ and $R\beta_0 = r$ then

$$\hat{\beta}_{R,T} \sim N\left(\beta_0, \sigma_0^2 D\right)$$

where

$$D = (X'X)^{-1} - (X'X)^{-1}R'\{R(X'X)^{-1}R'\}^{-1}R(X'X)^{-1}.$$

Under these assumptions, RLS is at least as efficient as OLS because:

$$Var[\hat{\beta}_T] - Var[\hat{\beta}_{R,T}] = \sigma_0^2 (X'X)^{-1}R'\{R(X'X)^{-1}R'\}^{-1}R(X'X)^{-1} = psd$$

However, if $R\beta_0 \neq r$ then: $E[\hat{\beta}_{R,T}] \neq \beta_0$.

Variable selection

So far have taken X as given but in practice need to choose regressors

- Choice may come from economic theory.

- Data-based methods:
  - Maximize R^2? Not a good idea.
  - Maximize $\bar{R}^2 = 1 - \frac{RSS/T-k}{TSS/T-1}$? Ok, but equivalent to including all variables with $|tstat| > 1$.

For further discussion please read Notes Section 2.9

- Notes: Sections 2.8 - 2.10 and Section 2.13 (Appendix on Statistical Distributions)

- Greene:
  - Classical hypothesis testing framework, Section C.7
  - Inference based on OLS estimators - Sections 5.1-5.5