Erasmus School of Economics

MOOC Econometrics

Lecture 2.4.2 on Multiple Regression: Evaluation - Statistical Tests

Christiaan Heij

Erasmus University Rotterdam

Test for a single restriction: t-test

- Under assumptions A1-A6:

  $E(b) = \beta$ and $\text{var}(b) = \sigma^2(X'X)^{-1}$.

- A7: ε is normally distributed.

Test

Check that A1-A7 imply $b \sim N(\beta, \sigma^2(X'X)^{-1})$.

- Answer: $b = (X'X)^{-1}X'(X\beta + \varepsilon) = \beta + (X'X)^{-1}X'\varepsilon$

  is linear function of $\varepsilon \sim N(0, \sigma^2 I)$.

t-test

- Test for relevance of single explanatory factor j:

  Test $H_0 : \beta_j = 0$ against $H_1 : \beta_j \neq 0$.

- A1-A7: $b_j \sim N(\beta_j, \sigma^2 a_{jj})$, a_{jj} is element (j, j) of $(X'X)^{-1}$.

  If $H_0 : \beta_j = 0$ holds, then $z_j = \frac{b_j - \beta_j}{\sigma\sqrt{a_{jj}}} = \frac{b_j}{\sigma\sqrt{a_{jj}}} \sim N(0, 1)$.

- Replace unknown σ by s, square root of $s^2 = e'e/(n-k)$.

  Test statistic: $t_j = \frac{b_j}{s\sqrt{a_{jj}}} = \boxed{\frac{b_j}{\text{SE}(b_j)}}$, with $\text{SE}(b_j) = s\sqrt{a_{jj}}$.

- A1-A7: $t_j \sim t(n-k)$ (close to normal unless $n-k$ small).

Test for multiple restrictions: F-test

- Test for multiple linear restrictions:

  Test $H_0 : R\beta = r$ against $H_1 : R\beta \neq r$.

  $\rightarrow$ R is given $(g \times k)$ matrix with $\text{rank}(R) = g$
  $\rightarrow$ r is given $(g \times 1)$ vector

- A1-A7 imply $b \sim N(\beta, \sigma^2(X'X)^{-1})$.

Test

Under H_0: $Rb \sim N(m, \sigma^2 V)$. Compute m and $\sigma^2 V$.

- Answer: $m = E(Rb) = RE(b) = R\beta = r$.

  $$\sigma^2 V = \text{var}(Rb) = R\text{var}(b)R' = \sigma^2 R(X'X)^{-1}R'.$$

F-test

- Then $(1/\sigma)(Rb - r) \sim N(0, V)$.
- Facts: $(1/\sigma^2)(Rb - r)'V^{-1}(Rb - r) \sim \chi^2(g)$.

$$F = (1/s^2)(Rb - r)'V^{-1}(Rb - r)/g \sim F(g, n - k).$$

- F can be computed from residual sums of squares:

$$\boxed{F = \frac{(e_0'e_0 - e_1'e_1)/g}{e_1'e_1/(n-k)}}$$

$\rightarrow$ $e_0'e_0$: sum of squared residuals of restricted model (H_0)

$\rightarrow$ $e_1'e_1$: sum of squared residuals of unrestricted model (H_1)

Test for removing a set of explanatory factors

- Restricted model: remove set of g explanatory factors.
- Re-order k factors so that last g are removed:

  Re-order $X = (X_1 \ X_2)$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, and $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

  X_2: last g columns of X (factors removed in restricted model)

  β_2: last g elements of β

  b_2: last g elements of b

- Then $y = X_1\beta_1 + X_2\beta_2 + \varepsilon = X_1 b_1 + X_2 b_2 + e$.

F-test

- $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$.
- Test $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$.
- If H_0 holds, then $F = \frac{(e_0'e_0 - e_1'e_1)/g}{e_1'e_1/(n-k)} \sim F(g, n - k)$

$\rightarrow$ $e_0'e_0$: sum of squared residuals of restricted model

(OLS in model $y = X_1\beta_1 + \varepsilon$)

$\rightarrow$ $e_1'e_1$: sum of squared residuals of unrestricted model

(OLS in model $y = X_1\beta_1 + X_2\beta_2 + \varepsilon$)

TRAINING EXERCISE 2.4.2

- Train yourself by making the training exercise (see the website).
- After making this exercise, check your answers by studying the webcast solution (also available on the website).