Coursera MOOC Econometrics - Test Exercise 3

Thijs Meijerink
2022-12-18

a) Consider the usual linear model, where $y = X\beta + \varepsilon$. We now compare two regressions, which differ in how many variables are included in the matrix X. In the full (unrestricted) model p_1 regressors are included. In the restricted model only a subset of $p_0 < p_1$ regressors are included Show that the smallest model is preferred according to the AIC if:

$$\frac{s_0^2}{s_1^2} < e^{\frac{2}{n}(p_1 - p_0)}$$

Answer:

$$AIC = \log\left(s_i^2\right) + \frac{2p_i}{n}$$

$$\log(s_0^2) + \frac{2p_0}{n} < \log(s_1^2) + \frac{2p_1}{n}$$

$$\log(s_0^2) - \log(s_1^2) < \frac{2}{n}(p_1 - p_0)$$

$$\log\left(\frac{s_0^2}{s_1^2}\right) < \frac{2(p_1 - p_0)}{n}$$

$$\frac{s_0^2}{s_1^2} < e^{\frac{2}{n}(p_1 - p_0)}$$

b) Argue that for very large values of n the inequality of (a) is equal to the condition

$$\frac{s_0^2 - s_1^2}{s_1^2} < \frac{2}{n}(p_1 - p_0)$$

Use that $e^x \approx 1 + x$ for small values of x.

Answer:

For small x:

$$e^{\frac{2}{n}(p_1 - p_0)} \approx 1 + \frac{2}{n}(p_1 - p_0)$$

$$\frac{S_0^2}{S_1^2} < 1 + \frac{2}{n}(p_1 - p_0)$$

$$\frac{S_0^2}{S_1^2} - 1 < \frac{2}{n}(p_1 - p_0)$$

$$\frac{s_0^2 - s_1^2}{s_1^2} < \frac{2}{n}(p_1 - p_0)$$

c) Show that for very large values of *n* the condition in (b) is approximately equal to

$$\frac{e_R' e_R - e_U' e_U}{e_U' e_U} < \frac{2}{n}\,(p_1 - p_0)$$

where e_R is the vector of residuals for the restricted model with p_0 parameters and e_U the vector of residuals for the full unrestricted model with p_1 parameters.

Answer:

$S_0^2 = \frac{e_R' e_R}{(n-p_0)}$ and $S_1^2 = \frac{e_R' e_R}{(n-p_0)}$

so:

$$\frac{s_0^2 - s_1^2}{s_1^2} = \frac{\frac{e_R' e_R}{(n-p_0)} - \frac{e_R' e_R}{(n-p_0)}}{\frac{e_R' e_R}{(n-p_0)}}$$

For large n :

$n - p_0 \approx n$ and $n - p_1 \approx n$

Therefore:

$$\frac{e_R' e_R - e_U' e_U}{e_U' e_U} < \frac{2}{n}\,(p_1 - p_0)$$

d) Finally, show that the inequality from (c) is approximately equivalent to an F-test with critical value 2, for large sample sizes.

Answer:

I did not figure this one..