Erasmus School of Economics

MOOC Econometrics

Test Exercise 3

Notes:

- See website for how to submit your answers and how feedback is organized.

Goals and skills being used:

- Apply the Akaike Information Criterion (AIC) for model selection.
- Examine large-sample behavior of AIC.
- Link AIC to F-test discussed in multiple regression lectures.

Questions

This test exercise is of a theoretic nature. The exercise is based on Exercise 5.2c of 'Econometric Methods with Applications in Business and Economics'. The question of interest is how the decision whether or not to include a group of variables differs based on AIC from that based on the F-test. We will stepwise show that for large samples selection based on AIC corresponds to an F-test with a critical value of approximately 2.

(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

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

(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.

(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.

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