HG
Nov. 19

ECON 4130 19H

Note that, in the multinomial case, testing a sub-model against the full multinomial model, the LR test and the Pearson chi-square test are asymptotically equivalent (see Rice page 342 in section 9.5), and they use the same chi-square distribution as an approximation to the distribution of the test statistics under H_0.

Exercises for no-seminar week 47

(The solution set will be put on the net on Thursday 21 Nov.)

I)

Rice chapter 9: No. 12, 33 (**Hint:** note that there are 0 parameters under H_0 here, so the DF for the Chi-square test must be equal to the number of free parameters in the full model.)

No. 40 (Remember that $Z \sim N(0,1) \quad \Rightarrow \quad Z^2 \sim \chi_1^2$ - distributed.) (See, e.g., Rice, example C, sec 2.3, p.61)

No. 41

II) An introductory exercise on F-testing

Note. An F-test is a test for several linear restrictions, tested jointly, in a regression problem. The F-test may be looked upon as a generalization of the T-test that is a test for just a single linear restriction. Note also that the F-test may be interpreted as a likelihood ratio test (LR-test). This is justified in the appendix (optional reading) of the lecture note on F-testing. (**End of note.**)

An econometric model contains a response, Y, and 6 (exogenous) explanatory variables, $X, Z_1, Z_2, U_1, U_2, U_3$. The data are observations of $n = 22$ *iid*[1] corresponding random vectors, $(Y_i, X_i, Z_{i1}, Z_{i2}, U_{i1}, U_{i2}, U_{i3})$, and the (full) regression model is (using the observed values of the explanatory variables as fixed[2])

(1) $$Y_i = \alpha + \beta x_i + \delta_1 z_{i1} + \delta_2 z_{i2} + \gamma_1 u_{i1} + \gamma_2 u_{i2} + \gamma_3 u_{i3} + e_i \quad \text{for } i = 1, 2, \ldots, 22$$

---

[1] i.e., the joint distribution for the seven variables in one vector is the same for all *i*, and two different vectors are stochastically independent.

[2] See appendix 1 in the lecture note on prediction and the iid model for a justification of this – i.e., that we may consider the explanatory variables in a regression model as fixed numbers without loosing information. The justification is based on the maximum likelihood principle.

Where, $e_1, e_2, \ldots, e_n$ are iid and normal, $e_i \sim N(0, \sigma^2)$.

A. Estimating (1) by OLS gives the following table of sums of squares (using Stata terminology)

Table 1 (for full model)

Source	SS	df
Model	7817	?
Residual	3743	?
Total	11560	?

Fill in the degrees of freedom (df's) in the table. Estimate the error term variance, σ^2 , using an unbiased estimator.

B. A submodel of interest is assuming both $\delta_1 = \delta_2$ and $\gamma_1 = \gamma_2 = \gamma_3$. We want to check if there is evidence in the data against this submodel using an appropriate F-test. We then need to re-estimate the model assuming the submodel (that we call the "reduced model") to be true. Using OLS for the reduced model implies that we must regress the response Y on a modified set of explanatory variables.

Write up the corresponding (to (1)) regression model in the reduced case.

[**Hint:** Introduce two new parameters, δ for the common value of δ_1, δ_2 , and γ for the common value of $\gamma_1, \gamma_2, \gamma_3$, and substitute in (1). Define new regressor (i.e., explanatory) variables whenever necessary.]

C. Estimating the reduced model by OLS gives the following table of sums of squares (using Stata terminology)

Table 2 (for the reduced model)

Source	SS	df
Model	5332	?
Residual	6228	?
Total	11560	?

Use this information to perform an F-test for testing the sub-model against the more general model in (1).

Calculate the P-value, either approximately using the quantile table 5 in the back of Rice's book, or exactly using (e.g.) the "F.dist" function in Excel, or the F(df1,df2,f) – function (or Ftail(df1,df2,f)-function) in STATA.

III) You may benefit from looking at problem 2 of regular exam 2014, and problem 2 of regular exam 2015 (with guidelines on the course web page).