Week 2: Multiple Regression, Training Exercises

Coursera/Erasumus U., Econometric Methods and Applications

Anthony Nguyen

Training Exercise 2.5

Notes:

- This exercise can be made without a computer.
- If you wish, you can use dataset `TrainExer25` that is available on the website.

Questions

Let e_i be the residuals of the model at the beginning of Lecture 2.5, where log-wage was regressed on a constant, and the variables *Female*, *Age*, *Educ*, and *Parttime*. If these resideuals are regressed on a constant and the three education dummies, then the result with the coefficients rounded to two decimals is:

$$e_i = 0.03 - 0.06DE2_i - 0.09DE3_i + 0.06DE4_i + res_i$$

(with $R^2 = 0.04$). Here res_i denote the residuals of this regression, which have the property that the sample mean is zero for eac of the four education levels.

(a) Give an interpretation of the four regression coefficients.

(b) Test if the three dummy coefficients are jointly significant, by means of the F-test:

$$F = \frac{(R_1^2 - R_0^2)/g}{(1 - R_1^2)/(n-k)}$$

Hint: First prove that $R_0^2 = 0$ *under* $H_0 : \beta_2 = \beta_3 = \beta_4 = 0$.
Note: The relevant 5% critical value is 2.6.

(c) Give an economic interpretation of the result in part (b).

(d) Above, it was stated that the residuals res_i have sample mean zero for each of the four education levels. Can you prove this result?

Hint: Use the fact that $X'e = 0$ *for OLS in* $y = Xb + e$. *Which* y *and* X *are relevant here?*

Answers

Let e_i be the residuals of the model at the beginning of Lecture 2.5, where log-wage was regressed on a constant, and the variables *Female*, *Age*, *Educ*, and *Parttime*. If these resideuals are regressed on a constant and the three education dummies, then the result with the coefficients rounded to two decimals is:

$$e_i = 0.03 - 0.06DE2_i - 0.09DE3_i + 0.06DE4_i + res_i$$

(with $R^2 = 0.04$). Here res_i denote the residuals of this regression, which have the property that the sample mean is zero for eac of the four education levels.

(a) Give an interpretation of the four regression coefficients.

(b) Test if the three dummy coefficients are jointly significant, by means of the F-test:

$$F = \frac{(R_1^2 - R_0^2)/g}{(1 - R_1^2)/(n-k)}$$

Hint: First prove that $R_0^2 = 0$ *under* $H_0 : \beta_2 = \beta_3 = \beta_4 = 0$.
Note: The relevant 5% critical value is 2.6.

(c) Give an economic interpretation of the result in part (b).

(d) Above, it was stated that the residuals res_i have sample mean zero for each of the four education levels. Can you prove this result?

Hint: Use the fact that $X'e = 0$ *for OLS in* $y = Xb + e$. *Which* y *and* X *are relevant here?*