Problem set 4

July 29th, 2011
due August 3rd, 2011

Question 1

Suppose that you use OLS to estimate the following model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i \qquad \text{where } i = 1, 2, ..., 95$$

The following parameters are obtained (standard errors in parentheses):

$$\begin{aligned} \hat{\beta}_1 &= 0.5 \quad (0.04) \\ \hat{\beta}_2 &= 2.3 \quad (1.1) \\ \hat{\beta}_3 &= 0.03 \quad (0.025) \end{aligned}$$

- perform the following tests on each parameter at the 10 percent and 1 percent significance levels.

$$\begin{aligned} H_0 : \beta_j = 0 \quad H_1 : \beta_j \neq 0 \\ H_0 : \beta_j = 0 \quad H_1 : \beta_j > 0 \end{aligned}$$

- construct a 95 percent confidence interval for β_2

Question 2

Using Stata and the data set crime1.dta to run an OLS regression on the following model:

$$avgsen_i = \beta_0 + \beta_1 durat_i + \beta_2 inc86_i + \beta_3 inc86sq_i + \beta_4 narr86_i + \epsilon_i$$

Descriptions of the variables are provided in the data set.

- report coefficients estimates, standard errors, and the t-statistics and p-values for the exclusion test for each coefficient
- perform the following tests on each parameter at the 5 percent significance levels reporting the t-statistic, critical value, and decision reached

$$\begin{aligned} H_0 : \beta_1 = 0 \quad H_1 : \beta_1 \neq 0 \\ H_0 : \beta_1 = -0.1 \quad H_1 : \beta_1 \neq -0.1 \\ H_0 : \beta_2 = -0.1 \quad H_1 : \beta_2 \neq -0.1 \\ H_0 : \beta_3 = 0.02 \quad H_1 : \beta_3 \neq 0.02 \\ H_0 : \beta_4 = 0 \quad H_1 : \beta_4 \neq 0 \\ H_0 : \beta_4 = 0 \quad H_1 : \beta_4 > 0 \end{aligned}$$

Briefly comment on each result.

- how would you state the null hypothesis that income in 1986 has no impact on the average sentence? what would be the alternative hypothesis for this test?

Question 3

Using Stata and the data set wage2.dta to run OLS regression on the following model:

$$ln(wage)_i = \beta_0 + \beta_1 educ_i + \beta_2 exper_i + \beta_3 exper_i^2 + \beta_4 age_i + \beta_5 tenure_i + \epsilon_i$$

Descriptions of the variables are provided in the data set. Notice that you first need to generate experience squared.

- provide coefficient estimates, standard errors, and the t-statistics and p-values for the exclusion test on each coefficient
- construct and carry out a t-test to test the null hypothesis that tenure and education have the same effect on the individual's wage, carrying out further regressions that may be necessary

Question 4 - more difficult, optional

Suppose you regress y on five explanatory variables and get the residuals e. Now suppose you obtain a sixth explanatory variable w. What relationship, if any, is there between the coefficient on w from regressing y on all six explanatory variables and from regressing e on all six explanatory variables? (Hint: you might want to look at the Ballentain graph (the circles I draw on the blackboard to illustrate some points) to answer this question)