Problem Set 1

October 22, 2020

Rules

- Solutions to problem sets must be completed and submitted individually.
- The deadline for submission is **14.00 on 30 October 2020**. Submissions after the deadline will not be accepted.
- The datasets for the STATA exercises can be found at SUCourse.
- **Submissions must include two files: 1) one word file including your answers 2) STATA log file.** Results in the log file and interpreted results must match. Otherwise, you will get zero points for those exercises.
- You are expected to answer all questions very clearly. Therefore, please be careful about use of language and writing. **For instance; if you are asked to interpret a coefficient, you need to interpret it in a way that someone who cannot see the data and does not know econometrics can understand what you mean.**
- Double check that you write your name/surname and student ID number.
- Failure in fulfilling any of these will result in a FAIL grade for that homework.

Question 1 (25 points)

You are hired by the governor to study whether a tax on liquor has decreased average liquor consumption in your state. You are able to obtain, for a sample of individuals selected at random, the difference in liquor consumption (in ounces) for the years before and after the tax. For person i who is sampled randomly from the population, Y_i denotes the change in liquor consumption. Treat these as a random sample from a Normal(μ, σ^2) distribution.

1) The null hypothesis is that there was no change in average liquor consumption. State this formally in terms of μ. The alternative is that there was a decline in liquor consumption; state the alternative in terms of μ. (5 points)
2) Now, suppose your sample size is n=900 and you obtain the estimates $\bar{Y}$=-32.8 and s=466.4. Calculate the t statistic for testing H_0 against H_1. (Because of the large sample size, just use the standard normal distribution tabulated in Table G.1.) Do you reject H_0 at the 5% level? At the 1% level? (10 points)
3) What has been implicitly assumed in your analysis about other determinants of liquor consumption over the two-year period in order to infer causality from the tax change to liquor consumption? (10 points)

Question 2 (15 points)

Using data from 1988 for houses sold in Andover, Massachusetts, from Kiel and McClain (1995), the following equation relates housing price (*price*) to the distance from a recently built garbage incinerator (*dist*):

$$\widehat{log(price)} = 9.40 + 0.312log(dist)$$

$$n == 135 \text{ and } R^2 = 0.162$$

1)Interpret the coefficient on log(*dist*). Is the sign of this estimate what you expect it to be? (5 points)
2) Do you think simple regression provides an unbiased estimator of the ceteris paribus elasticity of price with respect to *dist*? (Think about the city's decision on where to put the incinerator.) (5 points)
3) What other factors about a house affect its price? Might these be correlated with distance from the incinerator? (5 points)

Question 3 (15 points)

The data set in CEOSAL2.RAW contains information on chief executive officers for U.S. corporations. The variable *salary* is annual compensation, in thousands of dollars, and *ceoten* is prior number of years as company CEO.

1) Find the average salary and the average tenure in the sample. (5 points)

2) How many CEOs are in their first year as CEO (that is, $ceoten = 0$)? What is the longest tenure as a CEO? (5 points)

3) Estimate the simple regression model

$$log(salary) = \beta_0 + \beta_1 ceoten + u,$$

What is the (approximate) predicted percentage increase in salary given one more year as a CEO? (5 points)

Question 4 (20 points)

We used the data in MEAP93.RAW for Example 2.12 in the book. Now we want to explore the relationship between the math pass rate ($math10$) and spending per student ($expend$).

1) Do you think each additional dollar spent has the same effect on the pass rate, or does a diminishing effect seem more appropriate? Explain. (10 points)

2) Use the data in MEAP93.RAW to estimate the model:

$$math10 = \beta_0 + \beta_1 log(expend) + u,$$

Report the estimated equation in the usual way, including the sample size and R-squared. What is the explanatory power of the model? Explain. (5 points)

3) How big is the estimated spending effect? Namely, if spending increases by 10%, what is the estimated percentage point increase in $math10$? (5 points)

Question 5 (25 points)

Use the data in WAGE2.RAW to estimate a simple regression explaining monthly salary ($wage$) in terms of IQ score (IQ).

1) Find the average salary and average IQ in the sample. What is the sample standard deviation of IQ? (IQ scores are standardized so that the average in the population is 100 with a standard deviation equal to 15.) (5 points)

2) Estimate a simple regression model where a one-point increase in IQ changes $wage$ by a constant dollar amount. Use this model to find the predicted increase in wage for an increase in IQ of 15 points. Does IQ explain most of the variation in $wage$? (10 points)

3) Now, estimate a model where each one-point increase in IQ has the same percentage effect on $wage$. If IQ increases by 15 points, what is the approximate percentage increase in predicted $wage$? (10 points)