Practice Problems 4

Econometrics, Spring 2019
Belk College of Business, UNCC
Ercan Karadas

[1] Suppose a researcher wishes to examine the relationship between the income of an individual and her food expenditure. As data, a random sample of 5 people chosen and their income and expenditure data are as follows:

- Dependent variable (Y): food expenditure ($)
- Independent variable (X): monthly income ($100)

Individual (i)	Y_i	X_i
1	115	15
2	136	17.5
3	120	19.2
4	50	10
5	165	34

a) Obtain a scatterplot of the data by hand.

b) Using a ruler, draw a line that fits through the data. Measure the slope and intercept of the line you have drawn.

c) Write the population regression model for this example.

d) Write the sample regression line for this example.

e) Compute $\widehat{\beta}_0$ and $\widehat{\beta}_1$.

f) Compute predicted values ($\widehat{Y}_i$'s) and OLS residuals ($\widehat{u}_i$'s).

g) Compute R^2 and SER.

h) Compute the predicted food expenditure for someone with monthly income $2500 ($\widehat{Y}$) and compute the OLS residual ($\widehat{u}$).

[2] A soda vendor at UNC Charlotte football games observes that more sodas are sold the warmer the temperature at game time is. Based on 32 home games covering five years, the vendor estimates the relationship between soda sales and temperature to be

$$\widehat{Y} = -240 + 8\,X$$

where Y is the number of sodas she sells and X is temperature in degrees Fahrenheit.

a) Interpret the estimated slope and intercept. Do the estimates make sense? Why, or why not?

b) On a day when the temperature at game time is forecast to be 80 degrees Fahrenheit, predict how many sodas the vendor expects to sell?

c) Below what temperature are the predicted sales zero?

d) Sketch a graph of the estimated regression line.

[3] From monthly observations from 1985 to 2015, the following regression results were obtained, where Y = exchange rate of the Canadian dollar to the U.S. dollar (CAD/USD) and X = ratio of the U.S. consumer price index to the Canadian consumer price index; that is, X represents the relative prices in the two countries:

$$\widehat{Y} = -0.902 + 2.430\,X, \qquad R^2 = 0.440$$

a) Interpret the estimated intercept and slope coefficient. How would you interpret R^2?

b) Does the positive value of the estimated slope coefficient make economic sense? What is the underlying economic theory?

c) Suppose we were to redefine X as the ratio of the Canadian CPI to the U.S. CPI. Would that change the sign of the coefficient of X? Why?

[4] [**SW 4.1**] Suppose that a researcher, using data on class size (CS) and average test scores from 50 third-grade classes, estimates the OLS regression:

$$\widehat{\text{Test Score}} = 640.3 - 4.93 \times \text{CS}, \qquad R^2 = 0.11, \qquad SER = 8.7$$

a) A classroom has 25 students. What is the regressions prediction for that classrooms average test score?

b) Last year a classroom had 21 students, and this year it has 24 students. What is the regression's prediction for the change in the classroom average test score?

c) The sample average class size across the 50 classrooms is 22.8. What is the sample average of the test scores across the 50 classrooms? (Hint: Review the formulas for the OLS estimators.)

d) What is the sample standard deviation of test scores across the 50 classrooms? (Hint: Review the formulas for the R^2 and SER.)

[5] [**SW 4.2**] Suppose a random sample of 100 20-year-old men is selected from a population and that these mens height and weight are recorded. A regression of weight on height yields

$$\widehat{\text{Weight}} = -79.24 + 4.16 \times \text{Height}, \qquad R^2 = 0.72, \qquad SER = 12.6$$

where `Weight` is measured in pounds and `Height` is measured in inches.

a) What is the regressions weight prediction for someone who is (i) 64 in. tall, (ii) 68 in. tall, (iii) 72 in. tall?

b) A man has a late growth spurt and grows 2 in. over the course of a year. What is the regression's prediction for the increase in this man's weight?

c) Suppose, that instead of measuring `Weight` and `Height` pounds and inches, these variables are measured in centimeters and kilograms. What are the regression estimates from this new centimeter-kilogram regression, i.e. compute $\widehat{\beta}_0$ and $\widehat{\beta}_1$. (Hint: Use the formulas for these estimators and figure out how X and Y changes when you change the units of measurement)

[6] Consider the regression model

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

a) Suppose you know that $\beta_0 = 0$. Derive a formula for the least squares estimator of β_1, i.e. derive $\widehat{\beta}_1$.

b) Suppose you know that $\beta_0 = 4$. Derive a formula for the least squares estimator of β_1, i.e. derive $\widehat{\beta}_1$.

[7] A large class of students just sat an exam consisting of 30 true-false questions. The TA recorded the number of correct answers (Y) and the number of incorrect answers (X) for each student. Suppose we then regressed Y on X *(Hint: This is quite an unusual regression example. Just pay special attention to the relationship between X and Y and think about whether there should be an error term in the population regression model).*

a) Write the population regression function

b) What will be the value of the R^2? Explain

c) What will be the value of the standard error of the regression? Explain

[8] A researcher wishes to evaluate the inflation rate as a predictor of the spot rate in German treasury bills market. For a sample of 89 quarterly observations the following sample linear regression was obtained

$$\widehat{Y} = 0.0027 + 0.79X, \qquad R^2 = 0.097$$

where Y is the spot rate, and X is the inflation rate (in %). Both variables were measured annually.

a) Interpret the slope term of the estimated regression line?

b) Interpret the regression R^2?

c) Suppose that the inflation rate was 3% last year, and 5% this year. What is the regression's prediction for the change in the spot rate?

d) Suppose that the sample variance of spot rate is given as 0.3, i.e. $s_Y^2 = 0.3$. Find the standard error of the regression (SER)?

[9] Short-answer questions:

a) For a population with $\mu_Y = 110$ and $\sigma_Y^2 = 49$, use the Central Limit Theorem to find $P(\overline{Y} > 98)$.

b) When Y_i's are i.i.d. from a population with mean μ_Y and variance σ_Y^2 the following two results always hold: $E(\overline{Y}) = \mu_Y$ and $var(\overline{Y}) = \sigma_Y^2/n$. True or False? Explain.

c) If the mean and the variance of a random variable are given we can characterize the whole distribution of the random variable? True or False? Explain.