Homework Set 2

Due on April 1st

All questions are from the end-of-chapter exercises. The question numbers refer to those in the book. I highly recommend you reading the textbook and lecture notes before completing the homework questions. When reading the textbook, please pay attention to the sections on how to interpret the estimated coefficients.

Exercises

4.2 Suppose that a random sample of 200 twenty-year-old mean is selected from a population and that these men's height and weight are recorded. A regression of weight on height yields

$$\widehat{Weight} = -99.41 + 3.94 \times Height, R^2 = 0.81, SER = 10.2$$

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

a. What is the regression's weight prediction for someone who is 70 in. tall? 65 in. tall? 74 in. tall?

b. A man has a late growth spurt and grows 1.5 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 in pounds and inches these variables are measured in centimeters and kilograms. What are the regression estimates from this new centimeter-kilogram regression? (Give all results, estimated coefficients, R^2, and SER)

4.3 A regression of average weekly earning (AWE, measured in dollars) on age (measured in years) using a random sample of college-educated full-time worker aged 25-65 yields the following:

$$\widehat{AWE} = 696.7 + 9.6 \times Age, R^2 = 0.023, SER = 624.1$$

a. Explain what the coefficient values 696.7 and 9.6 mean.

b. The standard error of the regression (SER) is 624.1. What are the units of measurement for the SER? (Dollars? Years? Or is SER unit-free?)

c. The regression R^2 is 0.023. What are the units of measurement for the R^2? (Dollars? Years? Or is R^2 unit-free)

d. What is the regression's predicted earnings for a 25-year-old worker? A 45-year-old worker?

e. Will the regression give reliable predictions for a 99-year-old worker? Why or why not?

f. Give what you know about the distribution of earnings, do you think it is plausible that the distribution of errors in the regression is normal? (*Hint*: Do you think that the distribution is symmetric or skewed? What is the smallest value of earnings, and is it consistent with a normal distribution?)

g. The average age in this sample is 41.6 years. What is the average value of AWE in the sample? (*Hint*: Review Key Concept 4.2)

4.5 A professor decides to run an experiment to measure the effect of time pressure on final exam scores. He gives each of the 400 students in his course the same final exam, but some students have 90 minutes to complete the exam while others have 120 minutes. Each students is randomly assigned one of the examination times based on the flip of a coin. Let Y_i denote the number of points scored on the exam by the i$^{\text{th}}$ student ($0 \leq Y_i \leq 100$), let X_i denote the amount of time that the student has to complete the exam ($X_i = 90 or 120$), and consider the regression model $Y_i = \beta_0 + \beta_1 X_i + u_i$.

a. Explain what the term u_i represents. Why will different students have different values of u_i?

b. Explain why $E(u_i|X_i) = 0$ for this regression model.

c. Are the other assumptions in Key Concept 4.3 satisfied? Explain.

d. The estimated regression is $\hat{Y}_i = 49 + 0.24X_i$.

i. Compute the estimated regression's prediction for the average score of students given 90 minutes to complete the exam. Repeat for 120 minutes and 150 minutes.

ii. Compute the estimated gain in score for a student who is given an additional 10 minutes on the exam.

4.10 Suppose that $Y_i = \beta_0 + \beta_1 X_i + u_i$, where (X_i, u_i) are i.i.d., and X_i is a Bernoulli random variable with $\Pr(X = 1) = 0.20$. When $X = 1, u_i$ is $N(0, 4)$; when $X = 0, u_i$ is $N(0, 1)$.

a. Show that the regression assumptions in Key Concept 4.3 are satisfied.

b. Derive an expression for the large-sample variance of $\hat{\beta}_1$. (*Hint*: Evaluate the terms in Equation (4.21))

4.12 **a.** Show that the regression R^2 in the regression of Y on X is the squared value of the sample correlation between X and Y. That is show that $R^2 = r^2_{XY}$.

b. Show that the R^2 from the regression of Y on X is the same as the R^2 from the regression of X on Y.

c. Show that $\hat{\beta}_1 = r_{XY}(s_Y/s_X)$, where r_{XY} is the sample correlation between X and Y, and s_Y and s_Y are the sample standard deviations of X and Y.

Empirical Exercise

For the empirical exercise, you need to include the table for regression results, the graphs, like the scatterplot, and the R or STATA codes. The program codes should be appended at the end of all answers.

E4.2 On the text Web site `http:://www.pearsonhighered.com/stock_watson/`, you will find a data file **TeachingRatings** that contains data on course evaluations, course characteristics, and professor characteristics for 463 courses at the University of Texas at Austin. A detailed description is given in **TeachingRatings_Description**, also available on the Web site. One of the characteristics is an index of the professor's "beauty" as rated by a panel of six judges. In this exercise, you will investigate how course evaluations are related to the professor's beauty.

a. Construct a scatterplot of average course evaluations (*Course_Eval*) on the professor's beauty (*Beauty*). Does there appear to be a relationship between the variables?

b. Run a regression of average course evaluations (*Course_Eval*) on the professor's beauty (*Beauty*). What is the estimated intercept? What is the estimated slope? Explain why the estimated intercept is equal to the sample mean of *Course_Eval*. (*Hint*: What is the sample mean of *Beauty*?)

c. Professor Watson has an average value of *Beauty*, while Professor Stock's value of *Beauty* is one standard deviation above the average. Predict Professor Stock's and Professor Watson's course evaluations.

d. Comment on the size of the regression's slope. Is the estimated effect of *Beauty* on *Course_Eval* large or small? Explain what you mean by "large" and "small".

e. Does *Beauty* explain a large fraction of the variance in evaluations across courses? Explain.