Homework Set 3

Due on April 17th

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

5.5 In the 1980s Tennessee conducted an experiment in which kindergarten students were randomly assigned to "regular" and "small" classes, and given standardized tests at the end of the year. (Regular classes contained approximately 24 students, and small classes contained approximately 15 students.) Suppose that, in the population, the standardized tests have a mean score of 925 points and a standard deviation of 75 points. Let *SmallClass* denote a binary variable equal to 1 if the student is assigned to a small class and equal to 0 otherwise. A regression of *TestScore* on *SmallClass* yields

$$\widehat{TestScore} = \underset{(1.6)}{918.0} + \underset{(2.5)}{13.9} \times SmallClass,\ R^2 = 0.01,\ SER = 74.6.$$

a. Do small classes improve test scores? By how much? Is the effect large? Explain.

b. Is the estimated effect of class size on test scores statistically significant? Carry out a test at the 5% level.

c. Construct a 99% confidence interval for the effect of *SmallClass* on test score.

5.6 Refer to the regression described in Exercise 5.5.

a. Do you think that the regression errors plausibly are homoskedastic? Explain.

b. $SE(\hat{\beta}_1)$ was computed using Equation (5.3). Suppose that the regression errors were homoskedastic: Would this affect the validity of the confidence interval constructed in Exercise 5.5(c)? Explain.

5.8 Suppose that (Y_i, X_i) satisfy the assumptions in Key Concept 4.3 and, in addition, $u_i \sim N(0, \sigma_u^2)$ and is independent of X_i. A sample of size $n = 30$ yields

$$\hat{Y} = \underset{(10.2)}{43.2} + \underset{(7.4)}{61.5}X,\ R^2 = 0.54,\ SER = 1.52$$

where the numbers in parentheses are the homoskedastic-only standard errors for the regression coefficients.

a. Construct a 95% confidence interval for β_0.

b. Test $H_0 : \beta_1 = 55$ vs. $H_1 : \beta_1 \neq 55$ at the 55% level.

c. Test $H_0 : \beta_1 = 55$ vs. $H_1 : \beta_1 > 55$ at the 5% level. (*Hint*: this problem involves small samples with normally distributed errors. So what distribution does the t-statistic follow?)

5.10 Let X_i denote a binary variable and consider the regression $Y_i = \beta_0 + \beta_1 X_i + u_i$. Let $\bar{Y}_0$ denote the sample mean for observations with $X = 0$ and $\bar{Y}_1$ denote the sample mean for observations with $X = 1$. Show that $\hat{\beta}_0 = \bar{Y}_0$, $\hat{\beta}_0 + \hat{\beta}_1 = \bar{Y}_1$, and $\hat{\beta}_1 = \bar{Y}_1 - \hat{Y}_0$.

5.14 Suppose that $Y_i = \beta X_i + u_i$, where (u_i, X_i) satisfy the Gauss-Markov conditions given in Equation (5.31).

a. Derive the least squares estimator of β and show that it is a linear function of $Y_1, \ldots, Y_n$.

b. Show that the estimator is conditionally unbiased.

c. Derive the conditional variance of the estimator.

d. (Optional)[1] Prove that the estimator is BLUE.

Empirical Exercise

For the empirical exercise, you need to explain your results and 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.

E5.2 Using the data set `TeachingRatings` described in Empirical Exercise E4.2, run a regression of `Course_Eval` on `Beauty`. Is the estimated regression slope coefficient statistically significant? That is, can you reject the null hypothesis $H_0 : \beta_1 = 0$ versus a two-sided alternative at the 10%, 5%, or 1% significance level? What is the p-value associated with coefficient's t-statistic?

[1] That this question is optional means you do not need to complete this question. However, if you do, you will earn an extra point in your homework.