ES465 - Econometrics

Dr. Todd R. Yarbrough

Problem Set 2

1. Explain the difference between $\hat{\beta}_1$ and β_1; between the residual $\hat{u}_i$ and the regression error u_i; and the OLS predicted value $\hat{Y}_i$ and $E(Y_i|X_i)$.

2. A regression of average weekly earnings (AWE), measure in dollars) on age (measured in years) using a random sample of college-educated full-time workers aged 25-65 yields the following:

$$A\hat{W}E = 696.7 + 9.6 \times Age, R^2 = 0.023, SER = 624.1$$

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

   b. The SER is 624.1. What are the units of measurement for SER?

   c. The regression R^2 is 0.023. What are the units of measurement for the R^2?

   d. What does the regression predict will be the earnings for a 25 year-old worker? For a 45-year-old worker?

   e. Will the regression give reliable predictions for a 99-year-old worker? For a 45-year-old worker?

   f. Given what you know about the distribution of earnings, is it plausible that the distribution of errors ins symmetric or skewed?

   g. The average age in this sample is 41.6 years. What is the average value of AWE in the sample?

3. Recall the regression from problem set 1 on class-size and average test-scores:

$$Test\hat{S}core = 520.4 + 5.82 \times Age, R^2 = 0.08, SER = 11.5$$

   with: $\sigma_{\beta_0} = 20.4$ and $\sigma_{\beta_1} = 2.21$

   a. Construct a 95% confidence interval for β_1, the regression slope coefficient.

   b. Calculate the p-value for the two-sided test of the null hypothesis $H_0 : \beta_1 = 0$. Do you reject the null at the 5% level? At the 1% level?

   c. Calculate the p-value for the two-sided test of the null hypothesis $H_0 : \beta_1 = -5.6$. W/o doing any additional calculations, determine whether -5.6 is contained in the 95% confidence interval for β_1

   d. Construct a 99% confidence interval for β_0.

4. Use "BirthweightSmoking" data for the following questions. Be sure to read its description provided by the PDF in the data folder.

   a. In the sample:

   i. What is the average value of "Birthweight" for all mothers?

   ii. For mothers who smoke?

   iii. For mothers who do not smoke?

   b. i. Use the data in the sample to estimate the difference between average birth weight for smoking and non-smoking mothers.

   ii. What is the standard error for the estimated different in (i)?

   iii. Construct a 95% confidence interval for the difference in the average birth weight for smoking and nonsmoking mothers.

   c. Run a regression of "Birthweight" on the binary variable "Smoker".

   i. Explain how the estimated slope and the intercept are related to your answers in parts (a) and (b).

   ii. Explain how the $SE(\hat{\beta}_1)$ is related to your answer in b(ii).

   iii. Construct a 95% confidence interval for the effect of smoking on birth weight.