PROBLEM SET 4: MORE ON BIVARIATE REGRESSION

PAUL SCHRIMPF

DUE: FEBRUARY 3, 2015 AT THE START OF LECTURE

UNIVERSITY OF BRITISH COLUMBIA

ECONOMICS 326

Problem 1 (Based on Stock and Watson (2009) question 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 student is randomly assigned one of the examination times based on the flip of a coin. Let y_i denote the score of student i and x_i denote the time the student had to complete the exam ($x_i = 90$ or 120). Consider the regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i.$$

(i) Explain what ϵ_i represents. Why will different students have different values of ϵ_i?

(ii) Is exogeneity, $\mathrm{E}[\epsilon_i | x_i] = 0$ a good assumption for this model? Explain why or why not.

(iii) Are the other six simple linear regression assumptions from lecture 3 likely to hold? Explain.

(iv) The estimated regression is

$$\hat{y}_i = 41 + 0.25 x_i$$

  (a) What is the predicted average test score for students given 90 minutes to complete the exam? 120 minutes? 150 minutes?

  (b) Compute the estimated gain in test score for a student who is given an additional 10 minutes on the exam.

Problem 2: Consider the standard simple regression model

$$y = \beta_0 + \beta_1 x + \epsilon$$

with assumptions SR.1-SR.6. Define the estimator

$$\bar{\beta}_1 = \frac{\frac{\sum_{i=1}^n y_i 1\{x_i \geq 0\}}{\sum_{i=1}^n 1\{x_i \geq 0\}} - \frac{\sum_{i=1}^n y_i 1\{x_i < 0\}}{\sum_{i=1}^n 1\{x_i < 0\}}}{\frac{\sum_{i=1}^n x_i 1\{x_i \geq 0\}}{\sum_{i=1}^n 1\{x_i \geq 0\}} - \frac{\sum_{i=1}^n x_i 1\{x_i < 0\}}{\sum_{i=1}^n 1\{x_i < 0\}}}.$$

where $1\{x_i \geq 0\}$ is 1 if x_i is greater than 0 and 0 if x_i is less than 0. In other words, $\bar{\beta}_1$ is the difference between the average y conditional on x positive and the average y conditional on x negative divided by the difference between the average x conditional on x positive and the average x conditional on x negative. Assume that $\frac{\sum_{i=1}^n x_i 1\{x_i \geq 0\}}{\sum_{i=1}^n 1\{x_i \geq 0\}} \neq \frac{\sum_{i=1}^n x_i 1\{x_i < 0\}}{\sum_{i=1}^n 1\{x_i < 0\}}$.

(i) Show that $\mathrm{E}[\bar{\beta}_1] = \beta_1$

(ii) Is the variance of $\bar{\beta}_1$ less than $\mathrm{Var}(\hat{\beta}_1 | \{x_i\}_{i=1}^n) = \frac{\sigma_\epsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$? Why? [Hint: you do not need to actually calculate $\mathrm{Var}(\bar{\beta}_1)$ to answer this question.]

REFERENCES

Stock, J.H. and M.W. Watson. 2009. "Introduction to Econometrics, 2/E." *Instructor* :10.