Econometrics – Problem Set #4

Junwoo Yang

April 10, 2021

1 Exercises

Exercise 4.2. A random sample of 100 20-year-old men is selected from a population and these men's height and weight are recorded. A regression of weight on height yields

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

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

(a) What is the regression's weight prediction for someone who is 64 inches tall? 68 inches tall? 72 inches tall?

(b) A man has a late growth spurt and grows 2 inches 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*.)

Solution. (a)

(b)

(c)

□

Exercise 4.9. (a) A linear regression yields $\hat{\beta}_1 = 0$. Show that $R^2 = 0$.

(b) A linear regression yields $R^2 = 0$. Does this imply that $\hat{\beta}_1 = 0$?

Solution. (a)

(b)

□

Exercise 4.11. 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.

(b) Suppose you know that $\beta_0 = 4$. Derive a formula for the least squares estimator of β_1.

Solution. (a)

(b)

□

Exercise 4.12 (optional). (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 $R^2 = r_{XY}^2$.

(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_X and s_Y are the sample standard deviations of X and Y.

Solution. (a)

(b)

(c)

□

Exercise 4.14. Show that the sample regression line passes through the point $(\overline{X}, \overline{Y})$.

Solution. □

2 Additional questions

Exercise 4.5 (3rd edition). 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 number of points scored on the exam by the i^{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.

Solution. (a)

(b)

(c)

(d) (i)

(ii)

□

Appendix 4.2. SW Appendix 4.2 provides the derivation of the OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ in the simple regression model, $Y_i = \beta_0 + \beta_1 X_i + u_i$, $i = 1, \cdots, n$. To get the final expression for $\hat{\beta}_1$ in

$$\begin{aligned}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}) &= \sum_{i=1}^{n}(X_i - \overline{X})[\beta_1(X_i - \overline{X}) + (u_i - \overline{u})] \\ &= \beta_1 \sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{i=1}^{n}(X_i - \overline{X})(u_i - \overline{u}),\end{aligned} \tag{4.27}$$

show the following equalities:

(a) $\displaystyle\sum_{i=1}^{n} X_i^2 - n\overline{X}^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2$

(b) $\displaystyle\sum_{i=1}^{n} X_i Y_i - n\overline{X}\,\overline{Y} = \sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})$

Solution. (a)

(b)

□

Appendix 4.3. SW Appendix 4.3 shows that $\hat{\beta}_1 = \frac{s_{XY}}{s_X^2}$ is an unbiased estimator of β_1 in the simple linear regression model, $Y_i = \beta_0 + \beta_1 X_i + u_i$, $i = 1, \cdots, n$, given the three least squares assumptions (LSA) (in Key Concept 4.3). Take a close look at

$$\hat{\beta}_1 = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{4.28}$$

and

$$\begin{aligned}E(\hat{\beta}_1 | X_1, \cdots, X_n) &= \beta_1 + E\left[\left.\frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})u_i}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2}\right| X_1, \cdots, X_n\right] \\ &= \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})E(u_i | X_1, \cdots, X_n)}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2},\end{aligned} \tag{4.29}$$

and make sure you understand where in the proof LSA #1 and #2 are used.

Solution.

□