SOLUTIONS TO EXERCISE 6

FENG LI

NOVEMBER 25, 2013

13.2

Given the true model $Y_i = \beta_1 X_i + u_i$, we can have $y_i = \beta_1 x_i + (u_i - \bar{u})$, where $y_i = Y_i - \bar{Y}$ and $x_i = X_i - \bar{X}$. To estimate the incorrect model, we obtain

$$\begin{aligned}\hat{\alpha}_1 &= \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i \left(\beta_1 x_i + (u_i - \bar{u})\right)}{\sum x_i^2} \\ &= \frac{\beta_1 \sum x_i^2 + \sum x_i (u_i - \bar{u})}{\sum x_i^2} = \beta_1 + \frac{\sum x_i (u_i - \bar{u})}{\sum x_i^2}\end{aligned}$$

To check if α_1 is unbiased, we calculate the expectation,

$$E(\hat{\alpha}_1) = E\left(\beta_1 + \frac{\sum x_i (u_i - \bar{u})}{\sum x_i^2}\right) = \beta_1 + \frac{\sum x_i (E(u_i) - \bar{u})}{\sum x_i^2} = \beta_1.$$

That is, even if we introduce the unneeded intercept in the second model, the slope coefficient remains unbiased. But they don't have the same variance,

$$\begin{aligned}Var\left(\hat{\beta}_1\right) &= \frac{\sigma_u^2}{\sum X_j^2} \qquad \text{See 6A for a proof} \\ Var(\hat{\alpha}_1) &= \frac{\sigma_v^2}{\sum x_j^2} = \frac{\sigma_u^2}{\sum \left(X_i - \bar{X}\right)^2}\end{aligned}$$

13.3

At this time, we assume the true model is $Y_i = \alpha_0 + \alpha_1 X_i + v_i$. By E.q 6.1.6,

$$\hat{\beta}_1 = \frac{\sum X_i Y_i}{\sum X_i^2} = \frac{\sum X_i (\alpha_0 + \alpha_1 X_i + v_i)}{\sum X_i^2} = \frac{\alpha_0 \sum X_i}{\sum X_i^2} + \alpha_1 + \frac{\sum v_i X_i}{\sum X_i^2}$$

so, $E\left(\hat{\beta}_1\right) \neq \alpha_1$. The variances are as given in Exercise 13.2.

13.11

This is the third time we do this kind of exercise, the proof is pretty easy, please refer to our Ex. 3.14 to prove it.

(a).

$$\hat{\beta}_{1(TRUE)} = \hat{\beta}_1 + 5\hat{\beta}_2, \ \hat{\beta}_{2(TRUE)} = \hat{\beta}_2$$

(b).

$$\hat{\beta}_{1(TRUE)} = \hat{\beta}_1, \ \hat{\beta}_{2(TRUE)} = 3\hat{\beta}_2$$

School of Statistics and Mathematics, Central University of Finance and Economics. E-mail: m@feng.li.

(c). The intercept coefficient will be unbiased but the slope coefficient will be biased and inconsistent.

13.19

This is a general case, we can fix the problem without any knowledge of RESET test. Since $\hat{Y}_i = \hat{\beta}_i + \hat{\beta}_2 X_i$, we square this and obtain the following result

$$\hat{Y}_i^2 = \left(\hat{\beta}_1 + \hat{\beta}_2 X_i\right)^2 = \hat{\beta}_1^2 + \hat{\beta}_2^2 X_i^2 + 2\hat{\beta}_1\hat{\beta}_2 X_i$$

We substitute it in the RESET equation, we get

$$\begin{aligned} Y_i &= \alpha_1 + \alpha_2 X_i + \alpha_3 \hat{Y}_i^2 + v_i \\ &= \alpha_1 + \alpha_2 X_i + \alpha_3 \left(\hat{\beta}_1^2 + \hat{\beta}_2^2 X_i^2 + 2\hat{\beta}_1\hat{\beta}_2 X_i\right) + v_i \\ &= \left(\alpha_1 + \alpha_3 \hat{\beta}_1^2\right) + \left(\alpha_2 + 2\alpha_3 \hat{\beta}_1 \hat{\beta}_2\right) X_i + \left(\alpha_3 \hat{\beta}_2^2\right) X_i^2 + v_i \\ &= \gamma_1 + \gamma_2 X_i + \gamma_3 X_i^2 + v_i \end{aligned}$$

which is exactly the same thing as estimating the following model directly: $Y_i = \beta_1 + \beta_2 X_i + \beta_3 X_i^2 + u_i$.

13.20

We take Figure 13.4 (p. 497) as an example. the solid line gives the OLS line for all the data and the broken line gives the OLS line with the outlier omitted.

a. TRUE. In (a), the outlier is near the mean value of X and has low leverage and little influence on the regression coefficients.

b. TRUE. In (b), the outlier is near the mean value of X and has low leverage and little influence on the regression coefficients.

c. TRUE. In (c) the outlier has high leverage but low influence on the regression coefficients because it is in line with the rest of the observations.

d. TRUE. This is very easy to understand: without X, how comes X^2 !

e. This is because once we have the model (model 1)

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$$

We can obtain this model(model 2) also (by averaging the previous model)

$$\bar{Y} = \beta_1 + \beta_2 \bar{X}_2 + \beta_3 \bar{X}_3 + \bar{u}$$

If we let (model 1 - model 2), finally we have

$$Y_i - \bar{Y} = \beta_2 \left(X_{2i} + \bar{X}_2\right) + \beta_3 \left(X_{3i} + \bar{X}_3\right) + \left(u_i - \bar{u}\right).$$