Econ 103 Chapter 6 Discussion Questions

Ryan Martin

February 13, 2018

Note, all questions are 6.3, 6.4 and 6.14. 6.3 is a hand problem

6.3

Consider the model

$$y = \beta_1 + x_2\beta_2 + x_3\beta_3 + e$$

and suppose that application of least squares to 20 observations on these variables yields the following results

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} .96587 \\ .69914 \\ 1.7769 \end{bmatrix}$$

and

$$\widehat{cov(b)} = \begin{bmatrix} .21812 & .019195 & -.050301 \\ .019195 & .048526 & -.031223 \\ -.050301 & -.031223 & .037120 \end{bmatrix}$$

with

$$\hat{\sigma}^2 = 2.5193, \qquad R^2 = .9466$$

a

Find the total variation, unexplained variation and explained variation for this model

b

Find 95% interval estimates for β_2 and β_3.

c

Use a t-test to test the hypothesis $H_0 : \beta_2 \geq 1$ against the alternative $H_1 : \beta_2 < 1$

d

Use your answers in part (a) to test the joint hypothesis $H_0 : \beta_2 = 0, \beta_3 = 0$.

e

Test the hypothesis $H_0 : 2\beta_2 = \beta_3$

6.4

```
Consider the wage equation
\begin{align*}
\log(WAGE) & = \beta_1 + \beta_2 EDUC + \beta_3 EDUC^2 + \beta_4 EXPER + \beta_5 EXPER^2\\
& \hspace*{1cm} + \beta_6 (EDUC \times EXPER) + \beta_7 HRSWK + e
\end{align*}
 where the explanatory variables are years of education, years of experience and hours worked per week.
```

a Using an approximate 5% critical value of $t_c = 2$, what coefficient estimates are not significantly different form zero?

b

What restriction on the coefficients of Eqn (A) gives Eqn(B)? Use an F-test to test this restriction. Show how the same eresult can be obtained using a t-test.

c

What restrictions on the coefficients of Eqn (A) gives Eqn(C)? Use an F-test to test these restrictions. What question would you be trying to answer by performing this test?

d

What restrictions on the coefficients of Eqn (B) give Eqn (D)? Use an F-test to test these restrictions. What question would you be trying to answer by performing this test?

e

What restrictions on the coefficients of Eqn (A) give Eqn (E)? Use an F-test to test these restrictions. What question would you be trying to answer by performing this test?

f

Based on your answers to parts (a) to (e), which model would you prefer? Why?

g

Compute the missing AIC value for Eqn (D) and the missing SC value for Eqn (A). Which model is favored by the AIC? Which model is favored by the SC?

6.14

Following on from the example in section 6.3, the file hwage.dat contains another subset of the data used by labor economist Tom Mroz. The variables with which we are concerned are

- HW - Husband's wage in 2006 dollars

- HE - Husband's education attainment in years
- HA - Husband's age
- CIT - a variable equal to one if living in a large city, otherwise zero.

a

Estimate the model

$$HW = \beta_1 + \beta_2 HE + \beta_3 HA + e$$

What effects do changes in the level of education and age have on wages?

b

Does RESET suggest that the model in part (a) is adequate?

c

Add the variables HE^2 and HA^2 to the original equation and re-estimate it. Describe the effect that education and age have on wages in this newly estimated model.

d

Does RESET suggest that the model in part (c) is adequate?

e

Reestimate the model in part(c) with the variable CIT included. What can you say about the level of wages in large cities relative to outside those cities?

f

Do you think CIT should be included in the equation?

g

For both the model estimated in part (c) and the model estimated in part (e) evaluate the following four derivatives:

```
\begin{itemize}
    \item $\frac{\partial HW}{\partial HE} $ for $HE = 6$ and $HE = 15$
    \item $\frac{\partial HW}{\partial HA} $ for $HA = 35$ and $HA = 50$
\end{itemize}
```

Does the omission of CIT lead to omitted-variable bias? Can you suggest why?

```
my_wd <- "C:/Users/ryanj/Dropbox/TA/Econ 103/Winter 2018/Data/s4poe_statadata"
my_file <- paste(my_wd, "cps4_small.dta", sep = "/")
library(haven)
dat <- read_stata(my_file)
```

```r
reg_out <- lm(data = dat, I(log(wage)) ~ educ + exper + hrswk)
summary(reg_out)
```

```
##
## Call:
## lm(formula = I(log(wage)) ~ educ + exper + hrswk, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.65551 -0.36387 -0.01818  0.34827  1.53364
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.100540   0.109548  10.046  < 2e-16 ***
## educ        0.090306   0.006078  14.858  < 2e-16 ***
## exper       0.005776   0.001275   4.531 6.58e-06 ***
## hrswk       0.008941   0.001581   5.654 2.05e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5137 on 996 degrees of freedom
## Multiple R-squared:  0.2197, Adjusted R-squared:  0.2173
## F-statistic: 93.46 on 3 and 996 DF,  p-value: < 2.2e-16
```