6209690624 Thanapoj Visutaporn

EE325 Introductory Econometrics (Section 1 semester 1/2020)

Assignment 4

Instruction: Write your answer in either paper or digital paper. However, if you write on paper, please scan it and save as a PDF file. Submission is via BE-Moodle as a PDF file for both cases. (Please keep the file below 10MB as that is the maximum per file capacity for student.)

Due date: Friday, November 6,2020 (Before 10 P.M.)

1. From the data for 46 states in the United States for 1992, results of the regression are displayed as follows.

$$\begin{aligned} \ln C_i &= 4.30 - 1.34 \ln P_i + 0.17 \ln Y_i \\ se &= (0.91) \quad (0.32) \qquad (0.20) \\ \bar{R}^2 &= 0.27 \end{aligned}$$

where C_i = cigarette consumption, packs per year

P_i = real price per pack, \$ per pack

Y_i = real disposable income per capita, \$ per week

1.1) Do the estimation results follow the law of demand?

1.2) What is the elasticity of demand for cigarettes with respect to price? Is it statistically significant? If so, is it statistically different from 1?

1.3) What is the income elasticity of demand for cigarettes? Is it statistically significant? If not, what might be the reasons for it?

2. From estimating the regression equation on net financial wealth (nettfa), age of the survey respondent (age), and annual family income (inc) for people in the United States. The wealth and income variables are both recorded in thousands of dollars. The OLS estimation results for the model are given by

$$nettfa_i = \beta_1 + \beta_2 inc_i + \beta_3 age_i + u_i$$

```
reg nettfa inc age

      Source |       SS           df       MS      Number of obs   =     9,275
-------------+----------------------------------   F(2, 9272)      =    943.21
       Model |   6414618.8         2   3207309.4   Prob > F        =    0.0000
    Residual |  31528770.7     9,272  3400.42825   R-squared       =    0.1691
-------------+----------------------------------   Adj R-squared   =    0.1689
       Total |  37943389.5     9,274  4091.3726    Root MSE        =    58.313

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .9533566   .0252775    37.72   0.000     .9038072    1.002906
         age |   1.030777   .0591226    17.43   0.000     .9148838     1.14667
       _cons |  -60.69654   2.596333   -23.38   0.000    -65.78592   -55.60715
------------------------------------------------------------------------------
```

```
reg nettfa inc age agesq

      Source |       SS           df       MS      Number of obs   =     9,275
-------------+----------------------------------   F(3, 9271)      =    646.80
       Model |  6567017.15         3  2189005.72   Prob > F        =    0.0000
    Residual |  31376372.3     9,271  3384.35685   R-squared       =    0.1731
-------------+----------------------------------   Adj R-squared   =    0.1728
       Total |  37943389.5     9,274   4091.3726   Root MSE        =    58.175

------------------------------------------------------------------------------
      nettfa |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         inc |   .9782522   .0254891    38.38   0.000      .928288    1.028216
         age |  -2.231489   .4897118    -4.56   0.000    -3.191432   -1.271547
       agesq |   .0377221   .0056214     6.71   0.000      .026703    .0487413
       _cons |   4.680388   10.08099     0.46   0.642    -15.08056    24.44134
------------------------------------------------------------------------------
```

2.1) Test the coefficient, in the first model, $\beta_3 < 1$ in the first model or not?

2.2) Due to estimation result by adding the age^2 variable or agesq. Perform the test whether we should include the quadratic term of the age variable or not? (Test for both t-test and F-test.) Also, interpret the meaning of this coefficient.

3. You are conducting an empirical investigation into the median prices of houses in 506 communities of a large metropolitan area. The sample data consist of 506 observations on the following observable variables:

P_i: the median house price in community i, in dollars;

NOX_i: the level of nitrous oxide in the air of community i, in parts per 100 million;

$DIST_i$: the weighted distance of community i from municipal area, in miles;

$ROOM_i$: the average number of rooms per house in community i;

$STRAT_i$: the average student-teacher ratio of schools in community i.

Researcher estimates the following model of median house price. The OLS estimation results for the model are given by

$$\begin{aligned} ln(P_i) &= 11.08 - 0.9535\, ln(NOX_i) - 0.1343\, ln(DIST_i) + 0.2545 ROOM_i - 0.05245 STRAT_i \\ se &= (0.3181)\ (0.1167) \qquad (0.04310) \qquad (0.01853) \qquad (0.005897) \\ RSS &= 35.1835 \quad TSS = 84.5822 \end{aligned}$$

3.1) Interpret each of the coefficient estimates in regression equation.

3.2) Test the individual significance of each of the slope coefficient estimates for $ln(NOX_i)$ and $ROOM_i$.

3.3) Find the R-squared, adjusted R-squared, and test the joint significance of all the slope coefficient estimates.

3.4) If researcher would like to test the proposition that the marginal effect of $ln(NOX_i)$ on $ln(P_i)$ equals the marginal effect of $ln(DIST_i)$ on $ln(P_i)$, write the restricted model and perform the test comparing restricted and unrestricted model, given that OLS estimation of this restricted regression equation yields a Residual Sum of Squares value = 41.9532.

4. Production function (Y) of the industrial sector in Thailand. It depends on the capital factor (K) and labor factor (L) in the years 1980-2010 with the following estimation.

<u>Model 1</u>:

$$ln\,Y_t = 18.27 + 0.536\,ln\,L_t + 0.024\,ln\,K_t$$

$$R^2 = 0.9389, RSS = 0.0124$$

<u>Model 2</u>:

$$ln\left(\frac{Y}{L}\right)_t = 2.13 + 1.12\,ln\left(\frac{K}{L}\right)_t$$

$$R^2 = 0.8087, RSS = 0.0153$$

4.1) Interpret the coefficients of the independent variables in models 1 and 2.

4.2) Test the hypothesis. Is the industrial production function characterized by constant return to scale? (Hint: you can perform any type of test that you see fit.)

4.3) Can we compare the R^2 value between the two regression models? Why?

6209690629

Thanapong Visutaporn

1.1) According to law of demand, if the price of product is increased, the demand will fall accordingly.

As a result, the model follows the law of demand because the price has negative effect to the consumption : $\ln C_i = 4.30 - 1.34 \ln P_i + 0.17 \ln Y_i$

1.2) So, we have to test β_2 which in the model $\ln C_i = 4.30 \underbrace{- 1.34}_{\beta_2} \ln P_i + 0.17 \ln Y_i$

Test statistically significant

$H_0 : \beta_2 = 0$

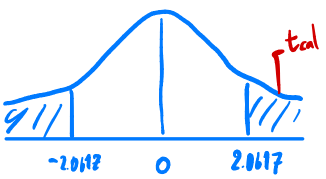
$H_a : \beta_2 \neq 0$

$$t_{cal}(\beta_2) = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \frac{-1.34 - 0}{0.32} = 4.1875$$

Supposed that we pick $\alpha = 0.05$

$t_{lower} = -2.0617$

$t_{upper} = 2.0617$



So, we can reject H_0. we can make sure that 95% of the times β_2 will not equal to 0

$H_0 : \beta_2 = 1$

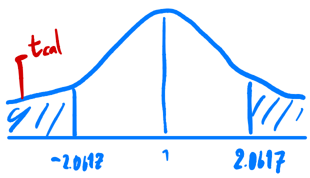
$H_a : \beta_2 \neq 1$

$$t_{cal}(\beta_2) = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \frac{-1.34 - 1}{0.32} = -7.3125 \sim t_{46-3}$$

Supposed that we pick $\alpha = 0.05$

$t_{lower} = -2.0617$

$t_{upper} = 2.0617$



So, we can reject H_0. we can make sure that 95% of the times β_2 will not equal to 1

1.3) We tested β_3 which in this model $\beta_3 = 0.17$

$H_0 : \beta_3 = 0$

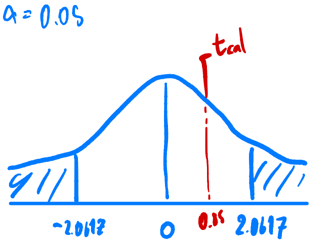
$H_a : \beta_3 \neq 0$

$$t_{cal}(\beta_3) = \frac{\hat{\beta}_3 - \beta_3}{\sigma_{\hat{\beta}_3}} = \frac{0.17 - 0}{0.20} = 0.85$$

Supposed that we pick $\alpha = 0.05$

$t_{lower} = -2.0617$

$t_{upper} = 2.0617$



- Maybe the reason why it is not statistically significant are coming from the real disposable income is a data per week which means the income per week per capita can not affect that much to whole cigarette consumption per year.

So, we can not reject H_0. we can make sure that 95% of the times β_2 will equal to 0 which means it is not statistically significant

6209690629
Thanapoj Visitsampan.

2.1) Test the coefficient, in the first model, $\beta_3 < 1$ or not

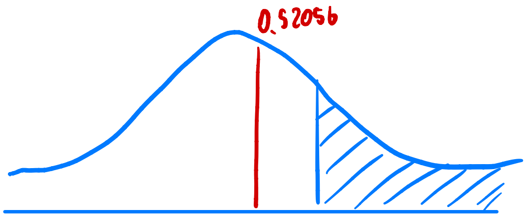
$H_0: \beta_3 \geq 1$

$H_a: \beta_3 < 1$

$t_{cal} = \frac{1.030777 - 1}{0.0591226} \approx 0.52056$

95% Conf. Interval

(0.9148838, 1.14667)



We cannot reject H_0, therefore, β_3 is not less than 1 in the first model

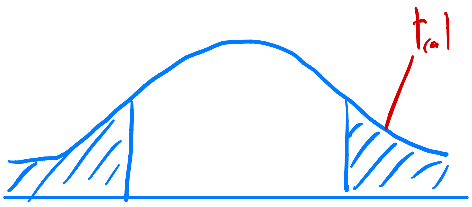
2.2) t-test

$H_0: \beta_4 = 0$

$H_a: \beta_4 \neq 0$

$t_{cal} = \frac{0.0377221 - 0}{0.0056214} \approx 6.710446$

(95% Conf. Interval)

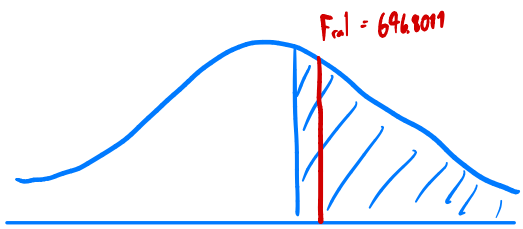


we can reject H_0, so, we can make sure that 95% of the times $\beta_4 \neq 0$

F-test

$F_{cal} = \frac{ESS/d.f.}{RSS/d.f.} = \frac{6567017.15/3}{31376372.3/9271} = \frac{2189005.72}{3384.35685}$

≈ 646.8011



$F_{cal} > F_{(3,9271)}$

we can reject H_0

β_4 is statically significant

6209690629
Thanapoj Visitanporn.

3.1) $\ln(P_i) = 11.08 - 0.9535 \ln(NOX_i) - 0.1343 \ln(DIST_i) + 0.2545\, ROOM_i - 0.052955\, STRAT_i$

① Expected/Average house price

② increased in 1 NOX_i, P_i decreases 0.9535

③ increased in 1 $DIST_i$, P_i decreases 0.1343

④ increased in 1 $ROOM_i$, P_i increases 0.2545

⑤ increased in 1 $STRAT_i$, P_i decreases 0.052955

3.2) Test $\ln(NOX)$

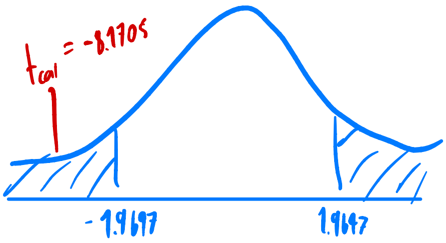
$H_0: \beta_2 = 0$

$H_a: \beta_2 \neq 0$

$$t_{cal} = \frac{-0.9535 - 0}{0.1167} = -8.1705 \sim t_{506-5}$$

$t_{upper} = 1.9647$

$t_{lower} = -1.9647$



We reject H_0

we can make sure that 95% of the times, $\beta_2 \neq 0$

β_2 is statiscally significant

Test $\ln(ROOM)$

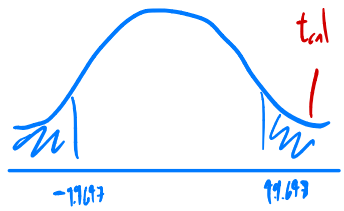
$H_0: \beta_5 = 0$

$H_a: \beta_5 \neq 0$

$$t_{cal} = \frac{0.2545 - 0}{0.01853} = 13.73 \sim t_{501}$$

$t_{upper} = 1.9647$

$t_{lower} = -1.9647$



We reject H_0

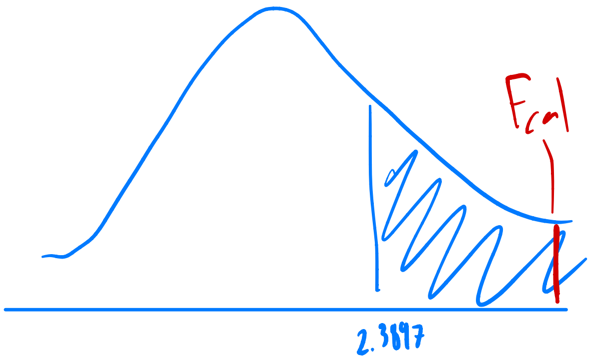
we can make sure that 95% of the times, $\beta_5 \neq 0$

β_5 is statiscally significant

6204690629
Thanapoj Visitanporn.

3.3) • $R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{35.1835}{84.5822} = 0.5840$

• Adjusted $R^2 / \bar{R}^2 = 1 - (1-R^2)\frac{n-1}{n-k} = 1\left[(1-0.5840)\frac{505}{501}\right] \Rightarrow 0.5807$



The joint significance of all the slope

$H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$
$H_a:$ otherwise

$F_{cal} = \frac{ESS/(k-1)}{RSS/(n-k)} = \frac{(49.3987)/4}{35.1835/501} = 175.8$

$F_{upper}(4, 501) = 2.3897$

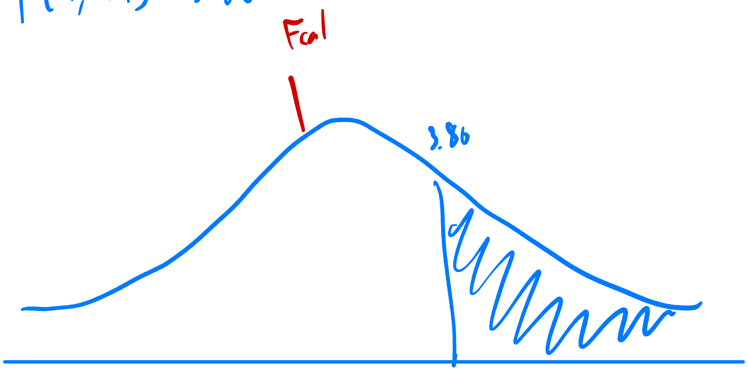
$F_{cal} > F_{(5,501)}$
– We reject H_0
all slope is not equal zero.

3.4) $H_0: \beta_2 = \beta_3$ or $\beta \ln(NOX) - \beta \ln(DIST) = 0$ (restrict is valid)
$H_a: \beta_2 \neq \beta_3$ (otherwise)

UR $= \ln(P_i) = 11.08 - 0.9535 \ln(NOX) + 0.2595(ROOM) - 0.0529(STAT)$
R $= \ln(P_i) = 11.08 - 6.7893 \ln(NOX \cdot DIST) + 0.2595(ROOM) - 0.0525(START)$

$F_{cal} = \frac{RSS_R - RSS_{UR}/m}{RSS_{UR}(n-k_{UR})} = \frac{41.9532 - 35.1835/1}{35.1835(506-5)} = 0.0003841$

$F(1, 501) = 3.86$



We cannot reject H_0
The restriction is not valid

6209690629

Thanapol Visitanporn.

Model 1

4.1) • 18.27 ⇒ constant

• increased 1 labor factor, Y increases 0.536

• increased 1 capital factor, Y increases 0.024

Model 2

• 2.13 ⇒ constant

• increased 1 $\frac{K}{L}$, Y increases 1.12

4.2) F-test

H_0 $\beta_2+\beta_3=1$

H_a : $\beta_2+\beta_3\neq 1$

$$F_{cal} = \frac{(RSS_R - RSS_{UR})/k_{UR}-k_R)}{RSS_{UR}/(n-k_{UR})} = \frac{(0.0153-0.0129)/(3-2)}{0.0129/(31-3)} \approx 6.598$$

$F_{table} = 4.23$

$F_{cal} > F_{table}$

We reject H_0, $\beta_2+\beta_3=1$

4.3) We cannot compare the R^2 value because the variable y did not same both models

Model 1 $\ln Y_t$

Model 2 $\ln\left(\frac{Y}{L}\right)_t$