ECO 5350
Intro. Econometrics

T. Fomby
Fall 2016

EXERCISE 2
KEY

Purpose: To learn some of the basic concepts of bivariate regression analysis. **This exercise is due on Tuesday, September 13.**

Consider the following data:

Y	X
8	1
7	2
13	3
11	4
16	5
15	6

Using the above data, compute the following quantities:

1) $\bar{y} = \sum y_i / n = (8+7+13+11+16+15)/6$
$= 70/6 = 11.67$

2) $\bar{x} = \sum x_i / n = (1+2+3+4+5+6)/6$
$= 21/6 = 3.5$

3) $\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{30}{17.5} = 1.714$

$\underline{(x_i - \bar{x})(y_i - \bar{y})}$

$(1-3.5)(8-11.67) = 9.175$
$(2-3.5)(7-11.67) = 7.005$
$(3-3.5)(13-11.67) = -0.665$
$(4-3.5)(11-11.67) = -0.335$
$(5-3.5)(16-11.67) = 6.495$
$(6-3.5)(15-11.67) = 8.325$

$\therefore \sum(x_i - \bar{x})(y_i - \bar{y}) = 30$

$\underline{(x_i - \bar{x})^2}$

$(-2.5)^2 = 6.25$
$(-1.5)^2 = 2.25$
$(-0.5)^2 = 0.25$
$(0.5)^2 = 0.25$
$(1.5)^2 = 2.25$
$(2.5)^2 = 6.25$

$\therefore \sum(x_i - \bar{x})^2 = 17.5$

4) $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 11.67 - (1.714)(3.5) = 5.671$

5) $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i = 5.671 + 1.714 X_i$

$\hat{y}_1 = 5.671 + 1.714(1) = 7.385$

$\hat{y}_2 = 5.671 + 1.714(2) = 9.099$

$\hat{y}_3 = 5.671 + 1.714(3) = 10.813$

$\hat{y}_4 = 5.671 + 1.714(4) = 12.527$

$\hat{y}_5 = 5.671 + 1.714(5) = 14.241$

$\hat{y}_6 = 5.671 + 1.714(6) = 15.955$

6) $\hat{u}_i = y_i - \hat{y}_i$

$\hat{u}_1 = 8 - 7.385 = 0.615$

$\hat{u}_2 = 7 - 9.099 = -2.099$

$\hat{u}_3 = 13 - 10.813 = 2.187$

$\hat{u}_4 = 11 - 12.527 = -1.527$

$\hat{u}_5 = 16 - 14.241 = 1.759$

$\hat{u}_6 = 15 - 15.982 = -0.982$

Notice that $\sum \hat{u}_i \doteq 0$

7) SST = $\sum(y_i - \bar{y})^2$ (Total Sum of Squares)

$y_i - \bar{y}$	$(y_i - \bar{y})^2$
$y_1 - \bar{y} = 8 - 11.67 = -3.67$	$(-3.67)^2 = 13.69$
$y_2 - \bar{y} = 7 - 11.67 = -4.67$	$(-4.67)^2 = 21.949$
$y_3 - \bar{y} = 13 - 11.67 = 1.33$	$(1.33)^2 = 1.7689$
$y_4 - \bar{y} = 11 - 11.67 = -0.67$	$(-0.67)^2 = 0.469$
$y_5 - \bar{y} = 16 - 11.67 = 4.33$	$(4.33)^2 = 18.7489$
$y_6 - \bar{y} = 15 - 11.67 = 3.33$	$(3.33)^2 = 11.0889$
	67.7147

8) SSE = $\sum(\hat{y}_i - \bar{y})^2$ (Explained Sum of Squares)

$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$	
$7.385 - 11.67$	$(-4.285)^2$	$= 18.361225$
$9.099 - 11.67$	$(-2.571)^2$	$= 6.610041$
$10.813 - 11.67$	$(-0.857)^2$	$= 0.734449$
$12.527 - 11.67$	$(0.857)^2$	$= 0.734449$
$14.241 - 11.67$	$(2.571)^2$	$= 6.610041$
$15.955 - 11.67$	$(4.285)^2$	$= 18.361225$
		51.41143

9) SSR = $\sum \hat{u}_i^2$ (Residual Sum of Squares)

$= (0.615)^2 + (-2.099)^2 + (2.187)^2 + (-1.527)^2 + (1.759)^2$
$+ (-0.982)^2 = 0.378225 + 4.405801 + 4.782969 +$
$2.331729 + 3.094081 + 0.96434$
$= 15.957$

10) $R^2 = \frac{SSE}{TSS}$ (Coefficient of Determination)

$= \frac{51.41143}{67.7147} = 0.759$

11) Construct the following Analysis of Variance Table:

Source	**df**	**MS**	**F**	**p-value**
SSE 51.41173	k 1	$\frac{SSE}{k}$ $\frac{51.41173}{1}$	$\frac{MSE}{MSR}$ =	$\Pr(F_{k,n-k-1} > F_0) = 0.0229$
SSR 15.957	n-k-1 6-2 = 4	$\frac{SSR}{n-k-1}$ $\frac{15.957}{4}$		
SST 67.7147	n-1 6-1 = 5			

$\frac{51.41173 \cdot 4}{15.957} = 12.88$

The number of observations is denoted by n while k denotes the number of explanatory variables in your multiple regression model not counting the intercept term.

12) $\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n-k-1} = \frac{15.957}{6-1-1} = \frac{15.957}{4} = 3.98925$

13) $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$ (Standard Error of Regression) $= \sqrt{3.98925} = 1.997$

14) $se(\hat{\beta}_0) = \sqrt{\frac{\hat{\sigma}^2 \sum x_i^2}{n \sum (x_i - \bar{x})^2}}$

$\sum x_i^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2$
$= 1 + 4 + 9 + 16 + 25 + 36$
$= 91$

$= \sqrt{\frac{3.98925(91)}{6(17.5)}}$

$= \sqrt{3.45735} = 1.859395$

15) $t_{\hat{\beta}_0} = \frac{\hat{\beta}_0}{se(\hat{\beta}_0)}$ $= \frac{5.671}{1.859} = 3.05$

16) $se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{\Sigma(x_i - \bar{x})^2}}$ $= \sqrt{\frac{3.98925}{17.5}} = 0.477$

17) $t_{\hat{\beta}_1} = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)}$ $= \frac{1.714}{0.477} = 3.59$

18) $\Pr\left(\hat{\beta}_0 - t_{n-k-1,0.025} \cdot se(\hat{\beta}_0) < \beta_0 < \hat{\beta}_0 + t_{n-k-1,0.025} \cdot se(\hat{\beta}_0)\right) = 0.95$ $t_{4,.025} = 2.776$

$[5.671 - 2.776(1.859) < \beta_0 < 5.671 + 2.776(1.859)]$

$[0.510, 10.83]$

19) $\Pr\left(\hat{\beta}_1 - t_{n-k-1,0.025} \cdot se(\hat{\beta}_1) < \beta_1 < \hat{\beta}_1 + t_{n-k-1,0.025} \cdot se(\hat{\beta}_1)\right) = 0.95$

$[1.714 - 2.776(0.477) < \beta_1 < 1.714 + 2.776(0.477)]$

$[0.39, 3.038]$

20) Suppose x = 2.5. Compute $\hat{y}$. (Prediction)

$\hat{y} = 5.671 + 1.714(2.5) = 9.956$