1. $\Sigma Y_i = 1110$ $\Sigma x_i = 1700$ $\Sigma x_i y_i = 205,500$ $\Sigma x_i^2 = 322,000$
$\Sigma y_i^2 = 132,100$

After correction of data

$\Sigma Y_i = 1110 - 90 - 140 + 80 + 150 = 1110$

$\Sigma x_i = 1700 - 120 - 220 + 115 + 230 = 1705$

$\Sigma x_i y_i = 205,500 - (90 \times 120) - (140 \times 220) + (80 \times 115) + (150 \times 230)$
$= 207,600$

$\Sigma y_i^2 = 132,100 - 90^2 - 140^2 + 80^2 + 150^2 = 133,300$

$\Sigma x_i^2 = 322,000 - 120^2 - 220^2 + 115^2 + 230^2 = 325,325$

The error in data will lead to miscalculation of r

$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{n(\Sigma x^2) - (\Sigma x)^2}\sqrt{n(\Sigma y^2) - (\Sigma y)^2}}$$

$$= \frac{10(207,600) - (1110)(1705)}{\sqrt{10(325,325) - (1705)^2}\sqrt{10(133,300) - (1110)^2}}$$

$$= \frac{183,450}{(588.40887)(317.6476)}$$

$$= 0.98158593$$

∴ The correct r is 0.98150593

(1)

2.The data file: HW4.XLSX in B.E moodle gives data on gross domestic product (GDP) for the United States for the years 1959-1997.

a. Use STATA program to plot the GDP data in current and constant (i.e., 1992) dollars against time.

b. Letting Y denote GDP and X time (measured chronologically starting with 1 for 1959, 2 for 1960, through 39 for 1997), see if the following model fits the GDP data:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

Use STATA program to estimate this model for both current and constant-dollar GDP.

c. How would you interpret β_1, β_2, r^2

d. If there is a difference between β_2 estimated for current-dollar GDP and that estimated for constant-dollar GDP, what explains the difference?

(Please print your output from STATA and submit it with your HW4)

From the stata output, we can write down our simple regression as follow:

Case 1: Using Current GDP (Y)

$\hat{Y}_t = -1024.565 + 205.5316\, X_t$ $\quad r^2 = 0.9250$

se = (220.7704) (9.619953) $\quad df = 37$

t = (-4.64) (21.37) $\quad F_{1,37} = 456.47$

from the result, we can interpret B_1, B_2, and r^2 as follow:

$\hat{B}_1 = -1024.565$

If X_t or time were zero, the average nominal (current) GDP would be about -1024.565 billions of dollars.

$\hat{B}_2 = 205.5316$

On average, when Xt or time increases (decreases) 1 year, the current GDP goes up (down) by about 205.5316 billions of dollars.

$r^2 = 0.9250$

since $r^2 = \frac{ESS}{TSS} = \frac{\Sigma(\hat{Y}_i - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2} = 0.9250$, it means that 92.50 percent of the variation in the current GDP is explained by X_t or time.

Case 2: Using the Constant-dollar GDP

$\hat{Y}_t = 1956.901 + 157.4753 X_t$ $\quad r^2 = 0.9840$

se = (75.7757) (3.3019) $\quad df = 37$

t = (25.82) (47.69) $\quad F_{1,37} = 2274.58$

(You can interpret the result by yourself)

Then, we would like to see whether X_t can explain the variation in Y_t or not. We can do the hypothesis testing.

We know that if X_t cannot be used to explain Y_t the slope (B_2) should be equal to zero.

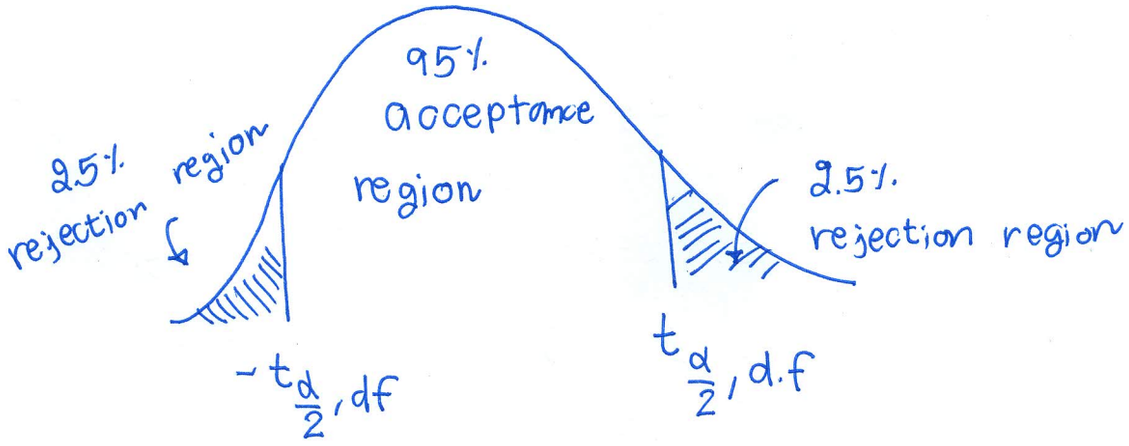
(1) $H_0: B_2 = 0$

$H_1: B_2 \neq 0$

for case 1: (using current GDP)

(2) set $\alpha = 0.05$

(3) Identify the critical values



open the t- table

$$t_{\frac{0.05}{2}, 37}$$

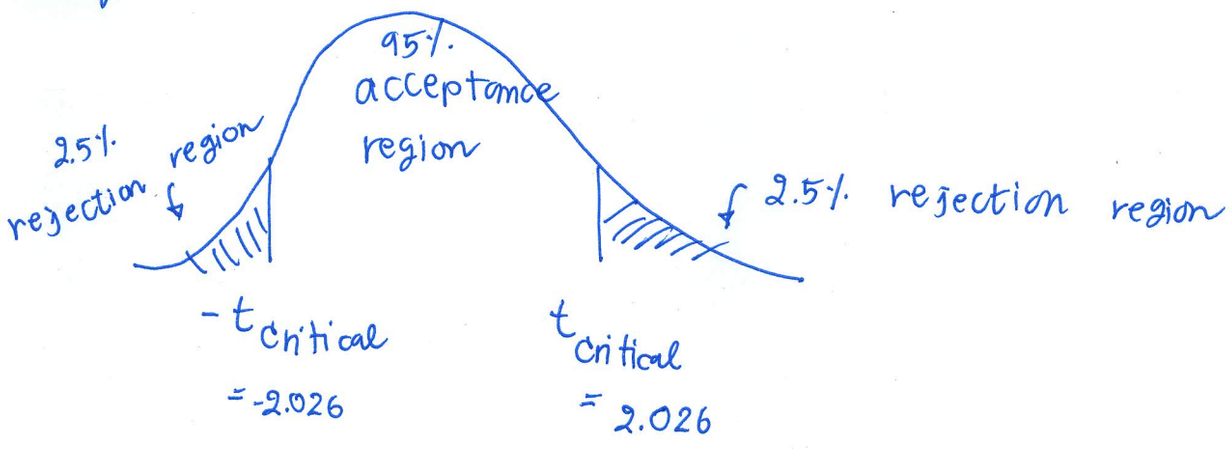
From the table

$$t_{\frac{0.05}{2}, 36} = 2.028$$

$$t_{\frac{0.05}{2}, 38} = 2.024$$

∴ We can approximate $t_{\frac{0.05}{2}, 37} = \frac{2.028 + 2.024}{2} = 2.026$

Therefore



(4) Calculate the $t_{calculation}$

we know that $= 0$ under the null hypothesis

$$t_{cal} = \frac{\hat{B}_2 - (B_2)}{se(\hat{B}_2)}$$ when $H_0: B_2 = 0$, $H_1: B_2 \neq 0$

$$= \frac{\hat{B}_2}{se(\hat{B}_2)} = \frac{205.5316}{9.619953} = 21.37$$

(which is the reported value in stata)

(5) Compare t_{cal} and $t_{critical}$

Since $t_{cal} = 21.37 > t_{critical} = 2.026$

∴ We reject the null hypothesis that $B_2 = 0$ with 95% confidence interval. In other words, X_t (time) can be used to explain the variation in Y_t. (nominal (current) GDP)

(You can do the hypothesis testing for case 2 by yourself).