EE325 Introductory Econometric (Section 1 semester 1/2020)

Assignment 3

Instructions:

(1) Write your answer in either paper or digital paper. However, if you write on paper, please scan it and save as a PDF file, **naming the file as your ID number**.

(2) Submit via BE-Moodle. (Please keep the file below 10MB as that is the maximum per file capacity for student.)

(3) Due date for this assignment is Thursday, September 24, 2020 (Before class starts at 11.00 A.M.)

Given that a random variable X is 30-year mortgage interest rate (percent) and Y is median housing price in the states (in thousands of USD), the data is given in the table below.

X	7.6	10.1	7.6	10.3	10.3	6.9	7.4	8.1	7	6.5	5.8
Y	175	175	178	183	184	188	203	230	258	310	330

Answer the following questions. Also show your process of the finding in the answer sheet. **Calculation in MS Excel is allowed**, but please paste your calculation in the answer sheet.

Calculated stats can be found in the table below.

	X	Y	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})^2$	$X_i - \bar{X}$	$Y_i - \bar{Y}$	$(X_i - \bar{X})(Y_i - \bar{Y})$	$\hat{Y}$	$Y_i - \hat{Y}$	$(Y_i - \hat{Y})^2$	X_i^2
	7.6	175	0.132	1,976.207	-0.364	-44.455	16.165	228.142	-53.142	2,824.043	57.76
	10.1	175	4.564	1,976.207	2.136	-44.455	-94.971	168.417	6.583	43.331	102.01
	7.6	178	0.132	1,718.479	-0.364	-41.455	15.074	228.142	-50.142	2,514.192	57.76
	10.3	183	5.459	1,328.934	2.336	-36.455	-85.171	163.639	19.361	374.832	106.09
	10.3	184	5.459	1,257.025	2.336	-35.455	-82.835	163.639	20.361	414.553	106.09
	6.9	188	1.131	989.388	-1.064	-31.455	33.456	244.865	-56.865	3,233.576	47.61
	7.4	203	0.318	270.752	-0.564	-16.455	9.274	232.920	-29.920	895.187	54.76
	8.1	230	0.019	111.207	0.136	10.545	1.438	216.197	13.803	190.527	65.61
	7	258	0.929	1,485.752	-0.964	38.545	-37.144	242.476	15.524	241.008	49
	6.5	310	2.142	8,198.479	-1.464	90.545	-132.526	254.420	55.580	3,089.088	42.25
	5.8	330	4.681	12,220.298	-2.164	110.545	-239.180	271.143	58.857	3,464.117	33.64
Average	7.964	219.455									
Sum			24.965	31,532.727			-596.418			17,284.453	722.58

(1) Estimate $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$ with the method of Ordinary Least Square.

- $\hat{\beta}_2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{X})^2} = \frac{-596.418}{24.965} = -23.89$

- $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} = 219.455 - (-23.89)7.964 = 409.7$

(2) Find the coefficient of determination $\boldsymbol{r^2}$.

- $r^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = 1 - \frac{17{,}284.453}{31{,}532.727} = 0.4519$

(3) If the interest rate increases by 3 percent, how much, on average, median housing price would alter?

- Since the SRF is $\hat{Y}_i = 409.7 - 23.89X_i$, replacing 3 into X_i, we get,

- $\hat{Y}_i = 409.7 - 23.89(3) = 338.03$

(4) Find the standard error of $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$.

- First, we need to estimate $\hat{\sigma}^2$ from $\frac{\sum_{i=1}^{n}(Y_i - \hat{Y})^2}{n-2} = \frac{17{,}284.453}{11-2} = 1920.4948.$

- Then, replace σ^2 with $\hat{\sigma}^2$ in both $\sigma^2_{\hat{\beta}_1}$ and $\sigma^2_{\hat{\beta}_2}$, also plug all the value into the formula,

- $\sigma^2_{\widehat{\beta}_1} = \frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2}\sigma^2 = \frac{722.58}{11(24.965)}(1920.4948) = 5053.2969$

- $\sigma^2_{\widehat{\beta}_2} = \frac{\sigma^2}{\sum(X_i - \bar{X})^2} = \frac{1920.4948}{24.965} = 76.8275$

- Lastly, find the standard error for both of them,

- $\sigma_{\widehat{\beta}_1} = \sqrt{\sigma^2_{\widehat{\beta}_1}} = 71.0865$

- $\sigma_{\widehat{\beta}_2} = \sqrt{\sigma^2_{\widehat{\beta}_2}} = 8.7708$

(5) Report your result of estimation with essential information. (P-value is not needed)

- $\hat{Y}_i = 409.7 \quad - \quad 23.89X_i$

  $(71.0865) \quad (8.7708) \quad r^2 = 0.4519$

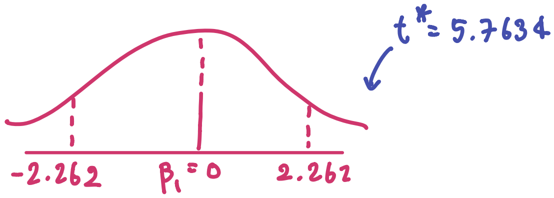
(6) Test that ~~$\hat{\beta}_1$ and $\hat{\beta}_2$~~ β_1 and β_2 are different from zero or not at the significance level of 0.05.

$H_0 : \beta_1 = 0$

$H_a : \beta_1 \neq 0$

- calculate $t^* = \dfrac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} = \dfrac{409.7}{71.0865} = 5.7634$

- Find 95% CI (d.f. = 9)



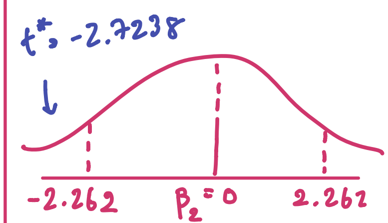
- Reject H_0 : We can make sure that 95% of the time, β_1 is significantly different from zero.

$H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

- calculate $t^* = \dfrac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \dfrac{-23.89}{8.7708} = -2.7238$

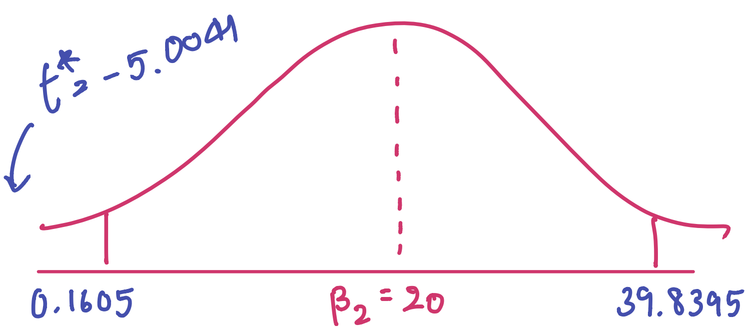
- Find 95% CI (d.f. = 9)



- Reject H_0 : We can make sure that 95% of the time, β_2 is significantly different from zero.

(7) Find the 95% confidence interval of ~~$\hat{\beta}_2$~~ $\beta_2 = 20$ and test that ~~$\hat{\beta}_2$~~ $\beta_2 = 20$ or not.

- Find 95% CI for $\beta_2 = 20$



upper bound : $\beta_2 + t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2}$

$= 20 + (2.262)\, 8.7708$

$= 39.8395$

lower bound : $\beta_2 - t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2}$

$= 20 - (2.262)\, 8.7708$

$= 0.1605$

- $H_0 : \beta_2 = 20$

  $H_a : \beta_2 \neq 20$

- calculate $t^* = \dfrac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \dfrac{-23.89 - 20}{8.7708} = -5.0041$

- Reject H_0 : We can make sure that 95% of the time, β_2 is significantly different from 20.

(8) Find the 99% confidence interval of σ^2.

- $\hat{\sigma}^2 = 1920.4948$
- construct 99% CI

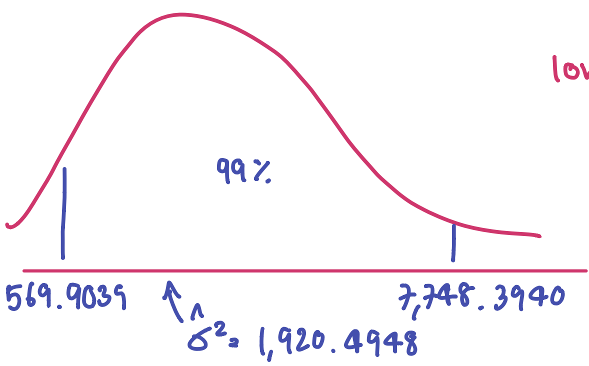
- calculate (d.f. = 9, $\frac{\alpha}{2} = 0.005$, $1 - \frac{\alpha}{2} = 0.995$)

upper bound : $(n-2)\dfrac{\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2}}} = \dfrac{7(1920.4948)}{1.735}$

$= 7{,}748.3940$

lower bound : $(n-2)\dfrac{\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}}} = \dfrac{7(1920.4948)}{23.589}$

$= 569.9039$



* $\hat{\sigma}^2$ falls within this 99% range.

(9) Elaborate the relationship between X and Y. Do you think that $\hat{\beta}_2$ is correctly estimated? Why and how 30-year mortgage interest rate correlates with median housing price in such way? Explain with economic reasoning.

(You may give your own rationale for this question)

*__NOTE:__ Please make sure that when you construct a confidence interval (for any case), the upper bound and the lower bound must be around that mean! For example, question 7 is obvious that when you are not testing against zero, you are creating a distribution around 20. Therefore, the upper bound and lower, indicating the region of acceptance H_0, are calculated from 20, NOT THE $\hat{\beta}_2$.

If anyone has the correct method of calculation, but wrongly choose a number to calculate for this part, it will be waived since this is my own fault. I was confused with the notation from the book!