Assignment #1

Instructions:

- For all questions, answer up to 4 decimal places.
- This assignment is due on **Tuesday, Mar 9, 2021 before class (11.00).**
- Write your answer in either digital or ordinary paper. For digital paper, export pages into a single PDF file. For ordinary paper, take photos of your writing and convert them into a single PDF file as well.
- There is no need to rewrite the question. Assign number item, ie. 1 a., clearly before your answer is sufficient.
- Submit your assignment into Moodle.
- Name your file as StudentID_Nickname (in Thai) such as 123456789_ไม้

1. Given this information

$n = 30$	$\sum_{i=1}^{n} X_i = 366$	$\sum_{i=1}^{n} Y_i = 631$	$\overline{X} = 12.20$	$\overline{Y} = 21.03$
$\sum_{i=1}^{n} (X_i)^2 = 5{,}564$	$\sum_{i=1}^{n} X_i Y_i = 7{,}524$	$\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 = 1098.8$	$\sum_{i=1}^{n} (Y_i - \overline{Y})^2 = 882.97$	
$\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) = -174.20$		$\sum_{i=1}^{n} \hat{u}_i^2 = 873.14$		

Answer the following questions. Show your work.

a) From regression model: $Y_i = \beta_1 + \beta_2 X_i + u_i$, $u_i \sim NIID\left(0, \sigma^2\right)$ (normally, identically and independently distributed), find the estimators of β_1 and β_2 with OLS method and explain the meaning of the model.

b) Find r^2 and explain its meaning.

c) If $X_i = 5$, estimate the value of $\hat{Y}_i$ and explain its meaning.

d) Find the estimators of $\operatorname{var}(u_i)$, $\operatorname{var}\left(\hat{\beta}_1\right)$ and $\operatorname{var}\left(\hat{\beta}_2\right)$

e) Test the hypothesis whether coefficients are different from zero at 0.05 level of significance.

f) Test the hypothesis whether coefficients are less than zero at 0.01 level of significance.

2. Given that Y is market price of a car (USD) while X is how long a car aged (years), results of the regression are as follows.

$$\hat{Y}_i = 7{,}836 - 502.4X_i$$
$$(52) \quad (411.8)$$

Given that u_i is normally, identically and independently distributed with zero mean and σ^2 variance, total number of observations is 11,

$\bar{X} = 7.45,$

$\hat{\sigma}^2 = 212{,}877,$

$\sum(X_i - \bar{X})^2 = 78.73,$

Answer the following questions. Show your work.

a) Does the sign of $\hat{\beta}_2$ make economic sense? Provide your explanation.

b) If you are a car expert and someone asks you to estimate how much his car will be **averagely** priced at when his car is 5 years old, how much is the market price range that you would estimate that you can make sure that for 95% of the time, market price will be within the specific range?

c) If you multiply all the X with 10, report the new SRF with the standard error resulted from the multiplication.

d) Calculate the elasticity of market price when a car is 10 years old.

1. Given this information

$n = 30$	$\sum_{i=1}^{n} X_i = 366$	$\sum_{i=1}^{n} Y_i = 631$	$\overline{X} = 12.20$	$\overline{Y} = 21.03$
$\sum_{i=1}^{n}(X_i)^2 = 5{,}564$	$\sum_{i=1}^{n} X_iY_i = 7{,}524$	$\sum_{i=1}^{n}(X_i-\overline{X})^2 = 1098.8$	$\sum_{i=1}^{n}(Y_i-\overline{Y})^2 = 882.97$	
$\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y}) = -174.20$		$\sum_{i=1}^{n}\hat{u}_i^2 = 873.14$		

Answer the following questions. Show your work.

a) From regression model $\boxed{Y_i = \beta_1 + \beta_2 X_i + u_i}$, $u_i \sim NIID(0, \sigma^2)$ (normally, identically and independently distributed), find the estimators of β_1 and β_2 with OLS method and explain the meaning of the model.

b) Find r^2 and explain its meaning.

c) If $X_i = 5$, estimate the value of $\hat{Y}_i$ and explain its meaning.

d) Find the estimators of $\text{var}(u_i)$, $\text{var}(\hat{\beta}_1)$ and $\text{var}(\hat{\beta}_2)$

e) Test the hypothesis whether coefficients are different from zero at 0.05 level of significance.

f) Test the hypothesis whether coefficients are less than zero at 0.01 level of significance.

a.) $\boxed{\hat{\beta}_2 = \dfrac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})}{\Sigma(X_i-\bar{X})^2}}$

$\hat{\beta}_2 = \dfrac{-174.20}{1098.8} = -\dfrac{13}{82} = -0.1585$ #

$\boxed{\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X}}$

$\hat{\beta}_1 = 21.03 - \left(-\dfrac{13}{82}\right)(12.2)$
$= 22.9641$ #

$\boxed{Y_i = \beta_1 + \beta_2 X_i + u_i} = \tilde{Y}_i = 22.9641 - 0.1585 X_i$

Meaning of the model: a linear line that minimize sum of the error tems.

$\therefore$ β_2: as X increases by 1 unit, Y will decrease by 0.1585 units, holding other thing constant.

β_1: assumed $X=0$, the average mean of Y is 22.9637. #

b.) $\boxed{r^2 = \dfrac{ESS}{TSS} = \dfrac{\Sigma(\hat{Y}_i-\bar{Y})^2}{\Sigma(Y_i-\bar{Y})^2} = 1 - \dfrac{\Sigma\hat{u}_i^2}{\Sigma(Y_i-\bar{Y})^2}}$

$r^2 = 1 - \dfrac{873.14}{882.97} = \dfrac{983}{88297} = 0.0111$ #

$\therefore$ $r^2 = 0.0111$, showing that approximately 1.11% of the observed variation can be explained by the model's inputs. #

c.) $X_i = 5 \rightarrow E(\hat{Y}_i \mid X_i = 5)$:

$\hat{Y}_i = 22.9641 - 0.1585(5) = 22.1633$ #

$\therefore$ if we're interested in out-of-sample $X_i = 5$ then $\hat{Y}_i = 22.1633$ #

d.) $\text{var}(u_i) = \hat{\sigma}^2 = \dfrac{\Sigma\hat{u}_i^2}{n-k} = \dfrac{873.14}{30-2} = 31.1836$ #

$\text{var}(\hat{\beta}_1) = \hat{\sigma}^2_{\hat{\beta}_1} = \dfrac{\Sigma X_i^2}{n\Sigma x_i^2}\hat{\sigma}^2 = \dfrac{5564}{30(1098.8)} \cdot 31.1836 = 5.2635$ #

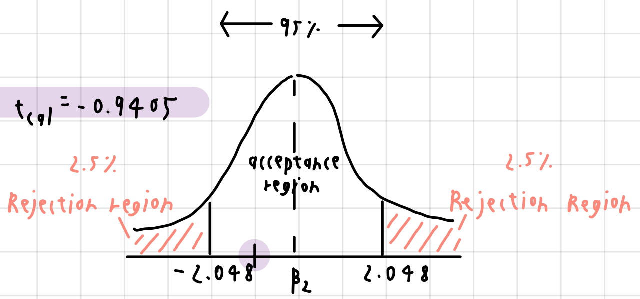
$\text{var}(\hat{\beta}_2) = \hat{\sigma}^2_{\hat{\beta}_2} = \dfrac{\hat{\sigma}^2}{\Sigma x_i^2} = \dfrac{31.1836}{1098.8} = 0.0284$ #

e.) level of significance = 0.05

β_2

1.) $H_0: \beta_2 = 0$ — Null hypothesis
$H_a: \beta_2 \neq 0$ — Alternative hypothesis

2.) $t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_{\hat{\beta}_2}} = \dfrac{-0.1585 - 0}{\sqrt{0.0284}} = -0.9405$



3.) d.f. $= n - k = 30 - 2 = 28$

The lower bound: $t_{\alpha/2} = t_{0.025} = -2.048$

The upper bound: $t_{\alpha/2} = t_{0.025} = 2.048$

confidence interval

4.) $t_{cal} = -0.9405$ lies within any boundary of CI (acceptance region), we can't reject the null hypothesis, at the significant level of 95%.

In other words, we can't say for sure that β_2 isn't zero 95 out of 100 times when we sample.

β_1

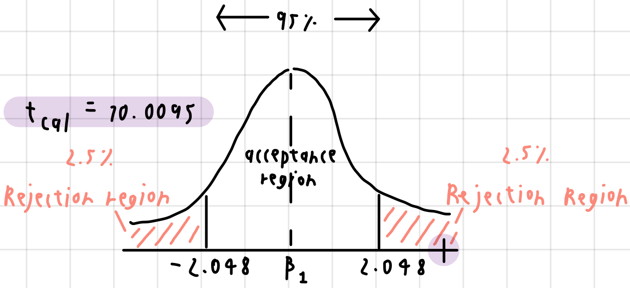
1.) $H_0: \beta_1 = 0$ — Null hypothesis
$H_a: \beta_1 \neq 0$ — Alternative hypothesis

2.) $t_{cal} = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\beta_1}} = \dfrac{22.9641 - 0}{\sqrt{5.2635}} = 10.0095$

3.) $d.f. = n - k = 28$

The lower bound: $t_{\alpha/2} = t_{0.025} = -2.048$

The upper bound: $t_{\alpha/2} = t_{0.025} = 2.048$



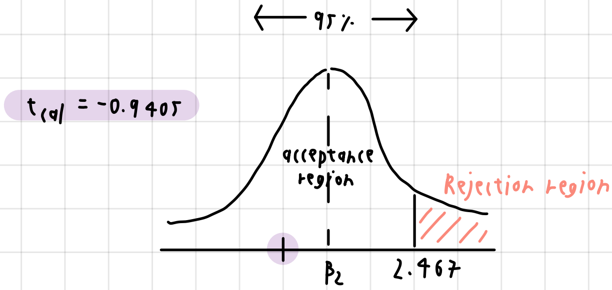
4.) $t_{cal} = 10.0095$ lies beyond the boundary of CI (rejection region), we can reject the null hypothesis, at the significance level of 95%.
In other words, we're sure that β_1 isn't zero 95 out of 100 times when we sample.

make sure that β_2 & β_1 are less than zero or not

f.) level of significance = 0.01

β_2

1.) $H_0: \beta_2 \geq 0$ — Null hypothesis
$H_a: \beta_2 < 0$ — Alternative hypothesis



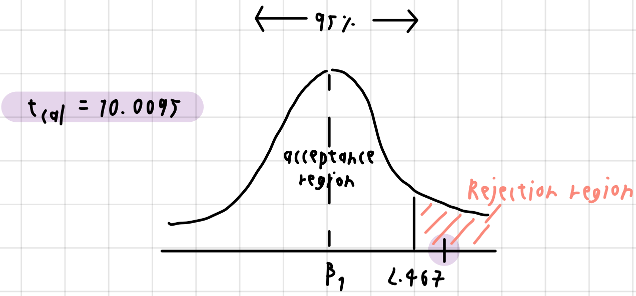
2.) $t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_{\hat{\beta}_2}} = \dfrac{-0.1585 - 0}{\sqrt{0.0284}} = -0.9405$

3.) The upper bound: $t_\alpha = 2.467$

4.) $t_{cal} = -0.9405$ lies within any boundary of CI (acceptance region), we can't reject the null hypothesis, at the significant level of 99%.
In other words, we can't say for sure that β_2 isn't less than 0 99 out of 100 times when we sample.

β_1

1.) $H_0: \beta_1 \geq 0$ — Null hypothesis
$H_a: \beta_1 < 0$ — Alternative hypothesis



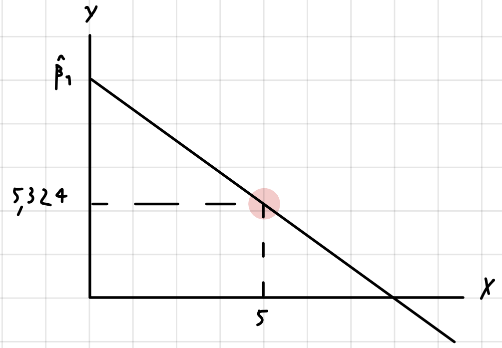
2.) $t_{cal} = \dfrac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} = \dfrac{22.9641 - 0}{\sqrt{5.2635}} = 10.0095$

3.) The upper bound: $t_\alpha = 2.467$

4.) $t_{cal} = 10.0095$ lies beyond the boundary of CI (rejection region), we can reject the null hypothesis, at the significance level of 99%.
In other words, we're sure that β_1 is less than 0 99 out of 100 times when we sample.

2. Given that Y is market price of a car (USD) while X is how long a car aged (years), results of the regression are as follows.

$$\hat{Y}_i = 7{,}836 - 502.4X_i$$
$$(52) \quad (411.8)$$



Given that u_i is normally, identically and independently distributed with zero mean and σ^2 variance,
total number of observations is 11, $n = 11$
$\bar{X} = 7.45$,
$\hat{\sigma}^2 = 212{,}877$,
$\sum(X_i - \bar{X})^2 = 78.73$,

Answer the following questions. Show your work.

a) Does the sign of $\hat{\beta}_2$ make economic sense? Provide your explanation.

b) If you are a car expert and someone asks you to estimate how much his car will be **averagely** priced at when his car is 5 years old, how much is the market price range that you would estimate that you can make sure that for 95% of the time, market price will be within the specific range?

c) If you multiply all the X with 10, report the new SRF with the standard error resulted from the multiplication.

d) Calculate the elasticity of market price when a car is 10 years old.

a.) Yes, because it shows that the more of X (how long a car aged in years), the less of Y (market price of a car in USD). This make sense because when capital is older, its value depreciates. So the price should be lower.

b.) $E(Y \mid X = 5)$:

$\hat{Y}_i = 7{,}836 - 502.4(5) = 5{,}324$

$X_0 = 5$, $\hat{Y}_0 = 5{,}324$

$n = 11$, $\bar{X} = 7.45$, $\Sigma(X_i - \bar{X})^2 = 78.73$, $\hat{\sigma}^2 = 212{,}877$

1.) $\text{var}(\hat{Y}_0) = \sigma^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\Sigma(X_i - \bar{X})^2} \right]$

$\text{var}(\hat{Y}_0 \mid X_0) = 212{,}877 \left[\frac{1}{11} + \frac{(5 - 7.45)^2}{78.73} \right] = 35582.5345$ #

2.) $\sigma_{\hat{Y}_0} = \sqrt{\text{var}(\hat{Y}_0)} = \sqrt{35582.5345} = 188.6333$ #

3.) d.f. $= n - k = 11 - 2 = 9$; $E(Y \mid X_0 = 5)$

Upper: $\hat{Y}_0 + (t_{\alpha/2} \cdot \sigma_{\hat{Y}_0}) = 5324 + (2.262 \cdot 188.6333) = 5750.6878$ #

Lower: $\hat{Y} - (t_{\alpha/2} \cdot \sigma_{\hat{Y}_0}) = 5324 - (2.262 \cdot 188.6333) = 4897.3122$ #

∴ CI over the mean value, 95 out of 100 times that CI will cover true value $E(Y \mid X_i = 5)$.
We can be 95% sure that $E(Y_0 \mid X_0 = 5)$ is within 4897.3115 USD to 5750.6885 USD. #

c.) Y in USD and X in years: if multiply X by 10 we'll get,

$\hat{Y}_i = 7{,}836 - 5024X_i$

se = (52) (4118)

∴ the slope coefficient as well as its se are 10 times their value. #

$X \uparrow$ 10 years $\rightarrow$ $Y \downarrow$ 5024 USD

∴ eventhough the unit changes, the interpretation should be the same.

d.) $X_i = 10$

$\hat{Y}_i = 7836 - 502.4(10) = 2{,}812$

$\frac{dY}{dX} \cdot \frac{X}{Y} = (-502.4)\frac{10}{2{,}812} = -\frac{1256}{703} = -1.7866$ #