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_โบ้

1. Given this information

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

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 $\text{var}(u_i)$, $\text{var}\left(\hat{\beta}_1\right)$ and $\text{var}\left(\hat{\beta}_2\right)$

α

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

↳ test parameter ว่าสามารถ imply เป็น 0 ← ไม่มีค่า

α

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

หาว่า = or ≠ 0

two tails - หา reject & acceptance ของ

One tails - หามากกว่าน้อยกว่า 0

Cheat sheet - OLS properties

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.

$$\hat{Y}_i = \overset{\hat{\beta}_1}{7{,}836} - \overset{\hat{\beta}_2}{502.4}X_i$$
$$\underbrace{(52) \quad (411.8)}_{\text{S.E.}}$$

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.

อธิบาย relationship X&Y

a) Does the sign of $\hat{\beta}_2$ make economic sense? Provide your explanation. เหมือน 1a

age

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?

→ confidence interval (หา $\hat{\beta}_1$ & $\hat{\beta}_2$)

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.

แทนค่า X

1

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

$= \frac{-174.20}{1098.8}$

$= -0.16$ *

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

$= 21.03 - (-0.16)(12.20)$

$= 21.03 - (-1.93)$

$= 21.03 + 1.93 = 22.96$ *

The meanings of $\hat{\beta}_1$ & $\hat{\beta}_2$ are that these 2 parameters show the best estimated values that are used to draw out the SRF with the least errors possible where the slope of SRF is -0.16 & the Y-intercept is 22.96 according to $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X + u_i$ where $\hat{\beta}_1$ & $\hat{\beta}_2$ refer to Y-intercept & slope, respectively. By every 1 unit change of X, Y will be decreased for 0.16, as Y is dependent variable.

(b) r^2 is used to determine how much the data points are away from the reference line which is $\bar{Y}$ in order to tell errors of SRF

$r^2 = \frac{ESS}{TSS} = \frac{\Sigma(\hat{Y}_i - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2}$

$= 1 - \frac{RSS}{TSS} = 1 - \frac{\Sigma \hat{u}_i^2}{\Sigma(Y_i - \bar{Y})^2}$

$= 1 - \frac{873.14}{882.97}$

$= 1 - 0.99 = 0.01$ *

The answer shows that the SRF has almost no unexplained part which can be referred that the estimated data point is almost be the same as the actual data point, where it is a lot far away from the mean of the sample.

(c) $\hat{Y}_i = \beta_1 + \beta_2 X_i + \hat{u}_i$

$\hat{Y}_i = 22.96 - 0.16 X_i$

$\hat{Y}_i = 22.96 - 0.16(5)$

$= 22.96 - 0.8 = 22.16$ *

By 22.16, it means when an observed variable of X_i is 5, Y will approximately be 22.16.

(d) $var(\hat{u}_1) = \hat{\sigma}^2 = \frac{\Sigma(Y_i - \hat{Y})^2}{n-k}$ ← Σu_i^2

$= \frac{873.14}{30-2} = 31.18$ *

$var(\hat{\beta}_1) = \hat{\sigma}^2_{\hat{\beta}_1} = \frac{\Sigma X_i^2}{n \Sigma x_i^2} \hat{\sigma}^2$ ← $\Sigma(X_i - \bar{X})^2$

$= \frac{5564}{30(1098.8)} \cdot 31.18 = 5.26$ *

$var(\hat{\beta}_2) = \hat{\sigma}^2_{\hat{\beta}_2} = \frac{\hat{\sigma}^2}{\Sigma x_i^2}$

$= \frac{31.18}{1098.8} = 0.03$ *

(e) Step 1 : State Hypothesis

→ $H_0 : \beta_1 = 0$ - Null Hypothesis

Step 2 : Find Tcal (test statistics)

→ $t_{cal} = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$

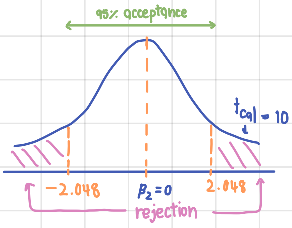
$= \frac{22.96 - 0}{\sqrt{5.26}} = 10$ *

Step 3 : State decision rule

$\alpha = 0.05$

Lower bound : $t_{\frac{\alpha}{2}} = -2.048$ *

Upper bound : $t_{\frac{\alpha}{2}} = 2.048$ *



→ Step 4 : Conclusion

Tcal lies beyond the CI boundary (Rejection Region) where we can reject the null hypothesis at the significance level of 95%

∴ We can say for sure that β_1 is not 0 for 95 out of 100 when we sample

β_1

Step 1 : State Hypothesis

→ $H_0 : \beta_2 = 0$ - Null Hypothesis

Step 2 : Find tcal (test statistics)

→ $t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_{\hat{\beta}_2}}$

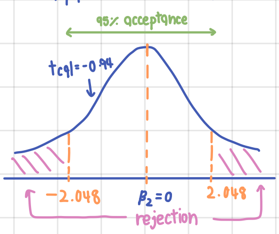
$= \frac{-0.16 - 0}{\sqrt{0.03}} = -0.94$ *

→ Step 3 : State Decision Rule

$\alpha = 0.05$ ↗ $t = 0.025$ n-k = 28

Lower bound : $t_{\frac{\alpha}{2}} = -2.048$ *

Upper bound : $t_{\frac{\alpha}{2}} = 2.048$ *



→ Step 4 : Conclusion

Tcal lies within the CI boundary (Acceptance Region) where we can't reject the null hypothesis, at the significance level of 95%

∴ We can't say for sure that β_2 is not 0 for 95 out of 100 when we sample

β_2

(f) Step 1 : State Hypothesis

→ $H_0 : \beta_1 \leq 0$ - Null Hypothesis

Step 2 : Find Tcal (test statistics)

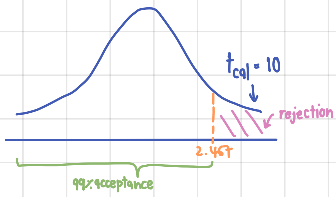
→ $t_{cal} = \frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}}$

$= \frac{22.96 - 0}{\sqrt{5.26}} = 10$ *

→ Step 3 : State Decision Rule

$\alpha = 0.01$

Upper bound : $t_{\alpha} = 2.467$



→ Step 4 : Conclusion

Tcal lies beyond the CI boundary (Rejection Region) where we can reject the null hypothesis at the significance level of 99%

∴ We can say for sure that β_1 is less than 0 for 99 out of 100% times when sampling

β_1

Step 1 : State Hypothesis

→ $H_0 : \beta_2 \leq 0$ - Null Hypothesis

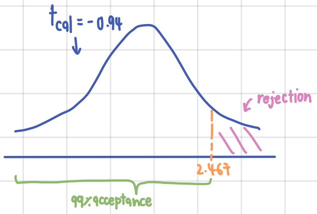
Step 2 : Find Tcal (test statistics)

→ $t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{\hat{\sigma}_{\hat{\beta}}} = \frac{-0.16 - 0}{\sqrt{0.03}} = -0.94$ *

→ Step 3 : State Decision Rule

$\alpha = 0.01$

Upper bound : $t_\alpha = 2.467$ *



→ Step 4 : Conclusion

Tcal lies within the CI boundary (Acceptance Region) where

we cannot reject the null hypothesis at the significance level of 99%

∴ we cannot say for sure that β_1 is less than 0 for 99 out of 100% times when sampling

β_2

2

(a) Yes, it is. $\hat{\beta}_1$ state the initial market price of a car at the time of purchase. Meanwhile, $\hat{\beta}_2$ states the depreciation cost of a car as each year passed by where it shows the relation of the age of a car (X_i) & the market price of a car (Y_i). By the increase in the age of a car for a year, the market price of a car will drop for 502.4 USD.

(b) Averaged price of a 5 year-old car ; $X_i = 5$

$E(Y \mid X_i = 5)$:

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

Mean Prediction : Market price range

Step 1 : Find var $(\hat{Y}_0)$

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

$\text{var}(\hat{Y}_0 \mid X_0) = 212,877 \left[\frac{1}{11} + \frac{(5 - 7.45)^2}{78.73}\right] = 35582.53$ *

Step 2: Find $\hat{\sigma}_{\hat{Y}_0}$

→ $\hat{\sigma}_{\hat{Y}_0} = \sqrt{\text{Var}(\hat{Y}_0)} = \sqrt{35582.53} = 188.63$ *

Step 3: Find 95% CI for $E(Y \mid X_0 = 5)$; d.f = 11-2 = 9

→ Upper Limit : $\hat{Y}_0 + (t_{\frac{\alpha}{2}} \cdot \hat{\sigma}_{\hat{Y}_0})$

$= 5324 + (2.262 \cdot 188.63) = 5750.69$ *

→ Lower Limit : $\hat{Y}_0 - (t_{\frac{\alpha}{2}} \cdot \hat{\sigma}_{\hat{Y}_0})$

$= 5324 - (2.262 \cdot 188.63) = 4897.31$ *

∴ CI over the mean for 95 out of 100 times, and CI will cover true value of $E(Y \mid X_i = 5)$, in which, we can be 95% sure that $E(Y \mid X_i = 5)$ is within 4897.31 USD to 5750.69

(c) Y_i is market price in USD & X_i is the age of a car in years ; when multiply X for 10 times

→ $\hat{Y}_i = 7,836 - 5024 X_i$

S.E. = (52) (4112)

∴ the slope & its S.E. will higher for 10 times of the initial value one.

where X is up to 10 years & the market price (Y) is 5024 USD less.

(d) Elasticity of Market price of a 10 year-old car :

→ $X_i = 10$

$Y_i = 7836 - 502.4(10) = 2,812$

Elasticity : $\frac{dy}{dx} \cdot \frac{X}{Y} = (-502.4) \cdot \frac{10}{2,812} = \frac{-1256}{705} = -1.79$ *

∴ it's elastic