Assignment 1
Assigned on Sep 14th, 2021. Due date Sep 27th, 2021 before midnight.

Instructions

(1) Please read the instruction carefully. Also take this habit with you into the exam room.
(2) Please read each question carefully and answer the questions straightforwardly. Always provide economic reasons at lease a paragraph for your analysis, or a graph when necessary, even when the question does not indicate so.
(3) Handing and submitting assignments are only available via BE Moodle.

Answering the questions and preparing answer sheets

(1) Answers are to be handwritten, in either digital or analog form, in a blank canvas or any clean paper. Make sure that your handwriting is clearly visible and readable.
(2) There is no need to rewrite the question. Just indicate the question number clearly for each of the answer, such as 1.a).
(3) Default decimal point is 4.
(4) Choose precise wordings, especially when you want to interpret the meaning of a test, confidence interval, or coefficients.
(5) When done, for the digital case, collage all the pages into a single PDF file. For those who write on sheets of paper, take photo of all pages then convert all of them into a single PDF file as well.
(6) Name your PDF file as StudentID_YourNickname, such as 640123456_Bo.

Submitting your answers

(1) Make sure your file does not exceed 10MB. This is the maximum file size for BE Moodle upload.
(2) Login to BE Moodle, head into the course, then the assignment topic.
(3) Choose your file to submit. Done. There will be timestamp for your upload date and time, so please make sure to not submit later than that.

1. (15 points) Given this information,

$n = 46$	$\sum_{i=1}^{n} X_i = 3{,}959.80$	$\sum_{i=1}^{n} Y_i = 3{,}180.80$
$\bar{X} = 86.0826$	$\bar{Y} = 69.1478$	
$\sum_{i=1}^{n}(X_i)^2 = 364{,}023.30$		$\sum_{i=1}^{n} X_i Y_i = 319{,}943.18$
$\sum_{i=1}^{n}(X_i - \bar{X})^2 = 23{,}153.3861$		$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = 94{,}525.1748$
$\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}) = 46{,}131.6183$		$\sum_{i=1}^{n} \hat{u}_i^{\,2} = 2{,}610.9211$

answer the following questions. Show your work.

a) (4 points) From regression model: $Y_i = \beta_1 + \beta_2 X_i + u_i$, $u_i \sim NIID\left(0, \sigma^2\right)$, find the estimators of β_1 and β_2 with OLS method and explain the meaning of the model.

b) (2 points) Find R^2 and explain its meaning.

c) (1 points) If $X_i = 60$, estimate the value of $\hat{Y}_i$ and explain its meaning.

d) (3 points) Find the estimators of $\operatorname{var}\left(u_i\right)$, $\operatorname{var}\left(\hat{\beta}_1\right)$ and $\operatorname{var}\left(\hat{\beta}_2\right)$

e) (2.5 points) What are the 95-percent confident intervals for β_2? Interpret the meaning.

f) (2.5 points) Test the hypothesis whether coefficients (both β_1 and β_2) are different from zero at 0.05 level of significance.

1

a) $\hat{\beta}_2 = \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sum(X_i-\bar{X})^2} = \frac{46{,}131.6183}{23{,}153.3861} = 1.99244$

$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} = 69.1478 - (1.99244)(86.0826) = -102.366615544$

$\hat{\beta}_1$ and $\hat{\beta}_2$ are coefficients of SRF ; $\hat{\beta}_1$ is the intercept , $\hat{\beta}_2$ is slope .

b) $r^2 = 1 - \frac{\sum \hat{u}_i^2}{\sum(Y_i-\bar{Y})^2} = 1 - \frac{2{,}610.9211}{94{,}525.1748} = 0.9724 = 97.24\%$

r^2 is the godness of fit ; how percentage of X can explain Y and fit in SRF.

c) $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i = -102.37 + (1.99 \cdot 60) = 17.03$

$\hat{Y}_i$ shows relationship in PRF ; if $X_i = 60$, $\hat{Y}_i = 17.03$

d) $var(U_i) = \frac{\sum \hat{u}_i^2}{n-k} = \frac{2{,}610.9211}{46-2} = 59.339116$

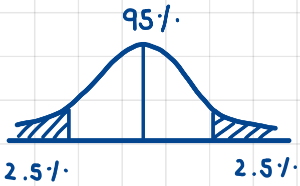
$var(\hat{\beta}_1) = \frac{\sum X_i^2}{n\sum(X_i-\bar{X})^2}\sigma^2 = \frac{364{,}023.30}{46(23{,}153.3861)}(59.339116) = 20.2814$

$var(\hat{\beta}_2) = \frac{\sigma^2}{\sum(X_i-\bar{X})^2} = \frac{(59.339116)}{23{,}153.3861} = 0.00256287 \approx 0.0026$

e) $\hat{\beta}_2 \pm t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2}$

upper : $1.99 + (2.021 \cdot \sqrt{0.0026}) = 2.09305$

lower : $1.99 - (2.021 \cdot \sqrt{0.0026}) = 1.88695$



It means that 95% would be correct mean by $(1.88695 < \hat{\beta}_2 < 2.09305)$

f) Step 1 : State hypothesis

$H_0 : \beta_2 = 0$ - Null hypothesis ← reject

$H_0 : \beta_2 \neq 0$ - Alternative hypothesis

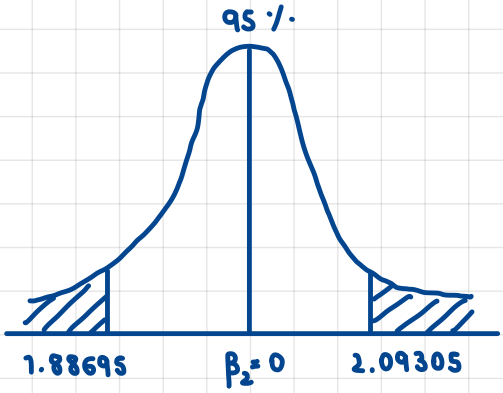
Step 2 : Calculate the statistics

$t_{cal} = \frac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \frac{1.99 - 0}{0.0026} = 765.3846$

Step 3 : Decision rule

from e) upper : 2.09305

lower : 1.88695



Step 4 : Conclude

β_2 is not 0 , 95 out of 100 times by expected.

$\therefore$ Reject $H_0 = 0$

---

2. (8 points) Answer the following question without any mathematical proof. A 3-6-line paragraph for a question is sufficient.

a) (2 points) If we have only one data point, can we create a sample regression function? Why?

b) (2 points) Does a significant β_2 sufficient for us to believe that X and Y are causally related? Provide an example to support your answer.

c) (2 points) When we test a hypothesis and find that β_2 is significantly different from zero, what does the result actually suggest? Answer with specific wording from statistical perspective.

d) (2 points) What is (are) (an) advantage(s) of an interval estimation over point estimate?

3. (7 points) Given that the dependent variable is natural log of wage (lwage) in Thai Baht and the independent variable is hours worked per week (main_hr), the result of estimation is shown in the table below here.

	Source	SS	df	MS
ESS	Model	50.060869	1	50.060869
RSS	Residual	166.152715	306	.542982728
TSS	Total	216.213584	307	.704278775

Number of obs	=	308
F(1, 306)	=	92.20
Prob > F	=	0.0000
R-squared	=	0.2315
Adj R-squared	=	0.2290
Root MSE	=	.73687

	lwage		Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
dep. Y								
indep. X	main_hr	β_2	.0318017	.003312	9.60	0.000	.0252844	.0383189
	_cons	β_1	7.658082	.1256392	60.95	0.000	7.410856	7.905308

Answer the following questions. Show your work.

a) (2 points) On average, how much is the nominal wage for a person who works 0 hour a week? (Note that this is a point estimation, not a prediction)

b) (2 points) If a person works an hour more, how much, on average, wage change do we expect?

c) (3 points) If you want to change the reading of unit from hours worked to days worked, what values in the main_hr row will differ? Calculate the changes to all the values in that row, **disregarding the constant row**. You might want to impose an assumption here. State that clearly before calculation.

****

② a) No, the SRF is the function use to estimate. It have to be more than one to create a SRF line which represent the relationship from multiple data points.

b) No, only information from β_2 is not enough. For example, positive relationship between death rate of Chinese and decreasing in Thai GDP. We cannot assume that if Chinese die, Thai GDP would fall.

c) $\beta_2 : H_1 = 0$ ← reject // $\beta_2 : H_1 \neq 0$

If β_2 is significantly different from 0 ; the null hypothesis would not be rejected.

d) Interval estimation is a range, with confident interval we can claim that it's not 100% which is safer to report the data while point estimation might be wrong and we don't know the range of the estimation.

③ a) $\ln Y = \beta_1 + \beta_2 X_i$

$$\ln \widehat{wage} = 7.658082 + 0.0318017\ (\cancel{0})$$

$$\widehat{wage} = e^{7.658082} \approx 2117.6918 \ \#$$

b) $\ln \widehat{wage} = 7.658082 + 0.0318017 \ln hour$

%Δ %Δ

$$\frac{d \ln \widehat{wage}}{d \ln hour} = 0.0318017 \qquad \text{- multiply by } \frac{100}{100}$$

$$\frac{d\,\widehat{wage}/wage}{d\,hour/hour} \cdot \frac{100}{100} = 0.031807 \cdot 100$$

$$\%\Delta \widehat{wage} = 3.1807\ d\,hour$$

∴ Increase 1 hour ; wage increase 3.1807%

c) $\ln Y = \beta_1 + \beta_2 X_i$ hour to day ; $X = 24$

$$\hat{\beta}_2 = 0.031807 \cdot 24 = 0.7632408$$

$$se_{\hat{\beta}_2} = 0.003312 \cdot 24 = 0.079488$$

CI ; upper : $\hat{\beta}_2 + t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2} = 0.7632408 + 1.984(0.079488) = 0.92095$

lower : $\hat{\beta}_2 - t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2} = 0.7632408 - 1.984(0.079488) = 0.60554$