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,

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

answer the following questions. Show your work.

a) (4 points) From regression model: $Y_i = \beta_1 + \beta_2 X_i + u_i$, $u_i \sim NIID(0, \sigma^2)$, 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 $\operatorname{var}(u_i)$, $\operatorname{var}(\hat{\beta}_1)$ and $\operatorname{var}(\hat{\beta}_2)$

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.

a) $\hat{\beta}_2 = \dfrac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sum(X_i-\bar{X})^2} = \dfrac{46{,}131.6183}{23{,}153.3861} = \underline{\underline{1.9924}}$

$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} = 69.1478 - (1.9924 \times 86.0826) = \underline{\underline{-102.3632}}$

$\underline{\underline{\hat{Y}_i = -102.3632 + 1.9924 X_i}}$

The intercept is at -102.3632 and the slope is 1.9924. Therefore, when $\hat{Y}_i$ increase by 1 unit, $\hat{Y}_i$ will also increase by 1.9924.

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

r^2 values of 0.0111 suggests that X_i about 1.11 percent of the variation in Y_i.

c) When $X_i = 60$

$$\hat{Y}_i = -102\ 3632 + 1.9924 X_i$$

$$= -102\ 3632 + 1.9924(60)$$

$$= \underline{\underline{17.1808}}$$

d) $Var(u_i)$, $Var(\hat{\beta}_1)$, $Var(\hat{\beta}_2)$

$Var(u_i)$: $\hat{\sigma}^2 = \frac{\Sigma \hat{u}_i^2}{n-k} = \frac{2610.9211}{46-2} = 59.3391$

$Var(\hat{\beta}_1)$: $\sigma^2_{\hat{\beta}_1} = \frac{\Sigma X_i^2}{n \Sigma x_i^2} \sigma^2$ $\quad \hat{\sigma}^2 = \frac{\Sigma \hat{u}_i^2}{n-k} = \frac{2610.9211}{46-2} = 59.3391$

$$= \frac{364,023.30}{46(23153.3861)} (59.3391) = 20.2814$$

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

$$= \frac{59.3391}{23153.3861} = 0.002563$$

e) $\alpha = 1-\alpha = 1-0.05 = 0.95 = 95\%$

$t_{\frac{\alpha}{2}} = t_{0.05/2} = t_{0.025}$

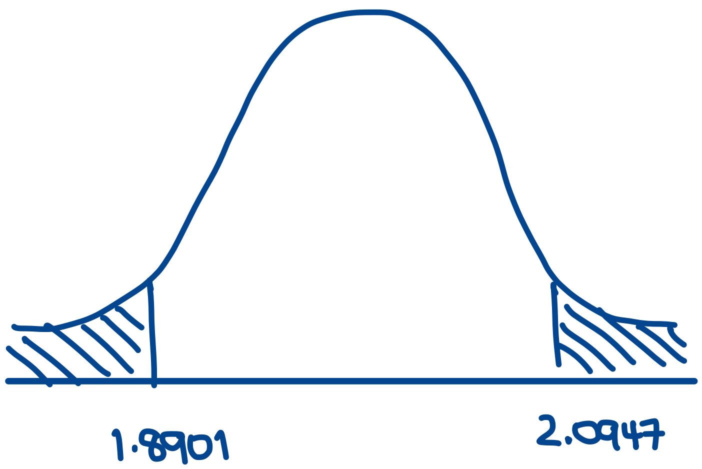
Degrees of freedom is 44

$t_{\frac{\alpha}{2}} = 2.021$

$\hat{\beta}_2 \pm t_{\frac{\alpha}{2}} \cdot \sigma\hat{\beta}_2 =$ upper : $1.9924 + (2.021 \times 0.0506) = 2.0947$

Lower : $1.9924 - (2.021 \times 0.0506) = 1.8901$

$P[1.8901 \leq \beta_2 \leq 2.0947] = 0.95$



∴ The confident interval of β_2 can be between 1.8901 to 2.0947. at the 95%.

f) $H_0 : \beta_1 = 0$ - null hypothesis

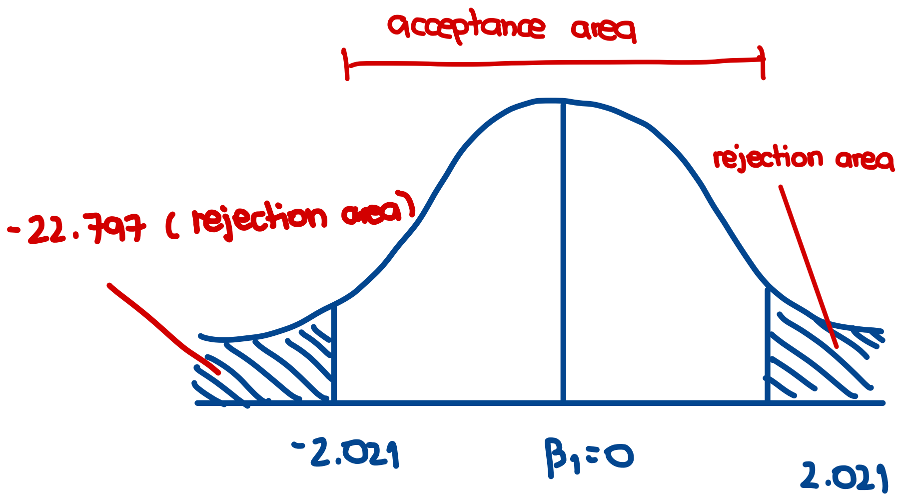
$H_a : \beta_1 \neq 0$ - Alternative hypothesis

0.05 level of significant.

$$t_{cal} = \frac{\hat{\beta}_1 - \beta_1}{\sigma \hat{\beta}_1} = \frac{-102.3632 - 0}{4.5035} = -22.7297$$

$$\sigma \hat{\beta}_1 = \sqrt{20.2814}$$

- The lower bound = -2.021
- The upper bound = 2.021



t_{cal} fall in the rejection area, we can reject the null hypothesis

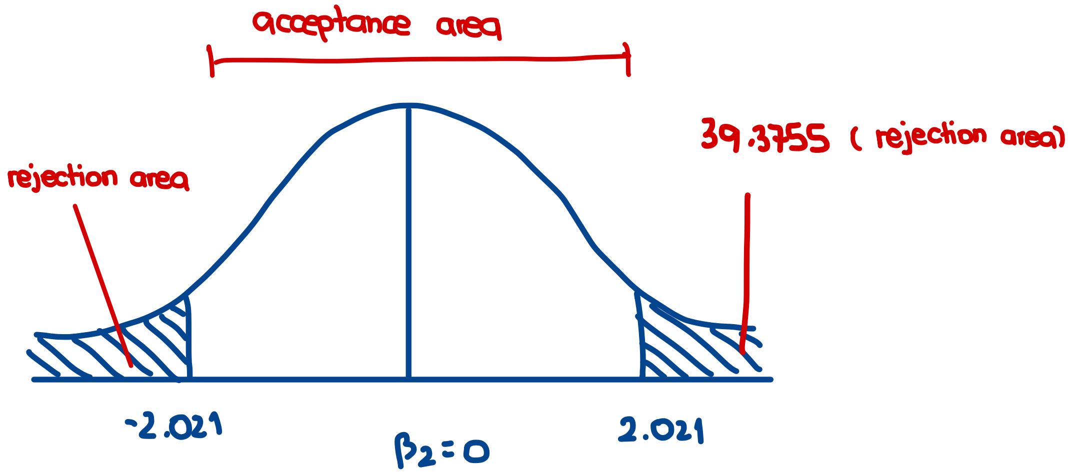
$H_0 : \beta_2 = 0$ - null hypothesis

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

$$t_{cal} = \frac{\hat{\beta}_2 - 0}{\sigma \hat{\beta}_2} = \frac{1.9924}{0.0506} = 39.3755$$

lower bound $= t_{\frac{\alpha}{2}} = -2.021$

upper bound $= t_{\frac{\alpha}{2}} = 2.021$



t_{cal} fall in the rejection area, Therefore, we can reject the null hypothesis, at the 0.05 level of significant.

---

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?

a) No, because when there is only one data, there will be only a point of line where we cannot find β_1, β_2. Need to have > 2 points.

b) Yes, since β_2 is the slope of the population regression function. This show the relationship between X and Y. For instance, when X increase by 1, Y will also increase by the amount of β_2.

c) When β_2 is significantly different from zero, X and Y are said to be related.

d) The advantages of an interval estimation is that there will be a range of values compare to the point estimation. So it can reduce the chance of having error rather than having a single data from 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
Model	50.060869	1	50.060869
Residual	166.152715	306	.542982728
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]
main_hr	.0318017	.003312	9.60	0.000	.0252844	.0383189
_cons	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.

a) $\widehat{lwage} = 7.658082 + 0.0318017\,(main\_hr)$ $main\_hr = 0$

$\widehat{lwage} = 7.658082$

$wage = 2117.6918$

The nominal wage is 2117.6918 when working at 0 hrs.

b) $\widehat{l\,wage} = 7.658082 + 0.0318\, l\,(main\_hr)$

$= \dfrac{d\, l\,\widehat{wage}}{d\ main\_hr} = 0.0318$

$= \dfrac{d\,\widehat{wage}/wage}{d\ main\_hr/main\_hr} \cdot \dfrac{100}{100} = 0.0318$

$= \%\Delta\,\widehat{wage} = 0.0318 \times 100$

$= 3.18\%$

If the person work more 1 hour, the wage expect to increase by 3.18 percent.

c) $l\,wage = Y$, $main\_hr = X$

hour $\div 24$ to day. X is divided by 24. Therefore $\hat{\beta}_2 \times 24$, $se(\hat{\beta}_2) \times 2$

$$\widehat{l\,wage} = 7.658082 + 0.0318\,(main\text{-}hr)$$

$$se = (0.1256) \quad (0.003312)$$

$\hat{\beta}_2 = 0.0318 \times 24 = 0.7632$

$se(\hat{\beta}_2) = 0.003312 \times 24 = 0.0795$

So $\hat{\beta}_2$ change from 0.0318 to 0.7632

$se(\hat{\beta}_2)$ change from 0.003312 to 0.0795.

Pinyu Kosolkarn 6304641100