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) $$\hat{\beta}_2 = \frac{\Sigma(x_i-\bar{x})(Y_i-\bar{Y})}{\Sigma(x_i-\bar{x})^2} = \frac{46{,}131.6183}{23{,}153.3861} = 1.9924 \#$$

$$\hat{\beta}_1 = \bar{Y} - \beta_2\bar{X} = 69.1478 - (1.9924)(86.0826) = -102.3632 \#$$

$$\because \hat{Y}_1 = -102.3632 + 1.9924X_i \quad (SRF)$$

$\therefore$ The intercept of this model is at -102.3632 and the slope is 1.9924 When x_i increases by 1 unit, $\hat{Y}_i$ increases by 1.9924. #

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

$\therefore$ r^2 values of 0.9724 suggests that X_i explain about 97.24% of variation in Y_i #

c) substitute $X_i = 60$ in $\hat{Y}_i = -102.3632 + 1.9924(60)$

$$\hat{Y}_i = 17.1808$$

$\therefore$ When $\hat{X}_i = 60$, $\hat{Y}_i = 17.1808$ #

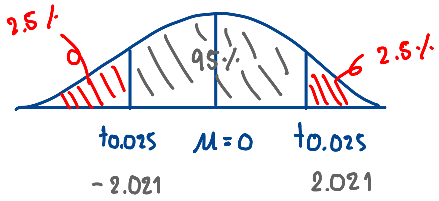
d) $$\text{var}(u_i) = \sigma^2 = \frac{\Sigma\hat{u}_i^2}{n-k} = \frac{2{,}610.9211}{46-2} = \frac{2{,}610.9211}{44} = 59.3391 \#$$

$$\text{var}(\hat{\beta}_1) = \sigma^2_{\hat{\beta}_1} = \frac{\Sigma x_i^2}{\Sigma(x_1-\bar{x})^2} = \frac{364{,}023.30}{23{,}153.3861} = 15.7222 \#$$

$$\text{var}(\hat{\beta}_2) = \sigma^2_{\hat{\beta}_2} = \frac{\sigma^2}{\Sigma x_i^2} = \frac{59.3391}{364{,}023.30} = 1.6301 \#$$

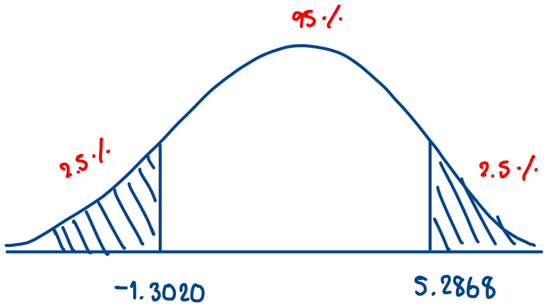
e) 1. 95% CI => $\alpha = 0.05$

2. $t_{\frac{\alpha}{2}} = t_{\frac{0.05}{2}} = t_{0.025}$ * d.f. $= n-k = 46-2 = 44$ *



3. $\hat{\beta}_2 \pm t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2}$ → upper limit . $1.9924 + (2.021 \times 1.6301) = 5.2868$ #

→ lower limit : $1.9924 - (2.021 \times 1.6301) = -1.3020$ #



$\because P[-1.3020 \leq \beta_2 \leq 5.2868] = 0.95$

∴ The 95% CI for β_2 indicates that 95% of the time that the range from -1.3020 to 5.2868 will contain true value of β_2. #

f) β_1 => 1. $H_0 : \beta_1 = 0$

$H_a : \beta_1 \neq 0$

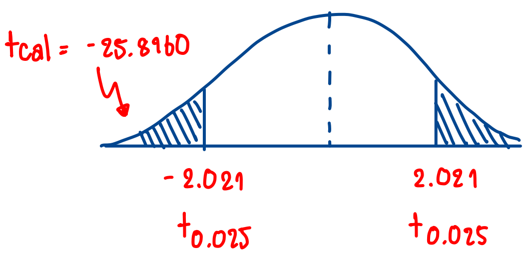
• $\sigma_{\hat{\beta}_1} = 3.9651$

2. $t_{cal} = \dfrac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} = \dfrac{-102.3632 - 0}{3.9651} = -25.8160$

3. When $\alpha = 0.05$

the lower bound : $t_{\frac{\alpha}{2}} = t_{0.025}$ d.f. $= n-k = 46-2 = 44$

the upper bound : $t_{\frac{\alpha}{2}} = t_{0.025}$



4. $\because$ $H_0 : \beta_1 = 0$ → Reject H_0

$H_a : \beta_1 \neq 0$

$\therefore$ β_1 is different from zero at 0.05 level of significance $\#$

β_2 $\Rightarrow$ 1. $H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

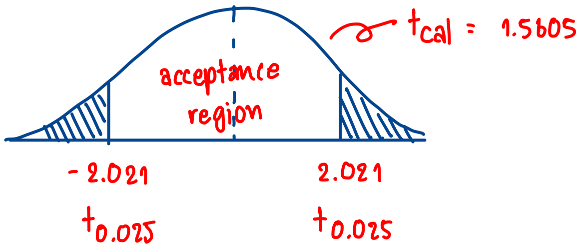
- $\sigma_{\hat{\beta}_2} = 1.2768$

2. $t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} = \dfrac{1.9924 - 0}{1.2768} = 1.5605$

3. When $\alpha = 0.05$

the lower bound : $t_{\frac{\alpha}{2}} = t_{0.025}$ $\qquad$ d.f. $= n-k = 46-2 = 44$

the upper bound : $t_{\frac{\alpha}{2}} = t_{0.025}$



4. $\therefore$ When t_{cal} lies within the acceptance region , we cannot reject H_0 Therefore,

β_2 is not different from zero at 0.05 level of significance. $\#$

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) To create a sample regression function, we need to obtain at least two data points to form a sample regression line. Therefore, with only one data point, we cannot create a sample regression function.

b) Yes, it does. This can be proved by using hypothesis testing to test against zero, to make sure that β_2 is not zero. If β_2 is statistically significant, we are sure to say that β_2 is not zero, then X and Y are casually related.

c) Given a 95% CI, when β_2 is significantly different from zero, meaning that our t_{cal} lies beyond the critical region, we can reject the null hypothesis ($\beta_2 = 0$). Then, we can be sured that β_2 is not zero 95% of times when we sample.

d) The advantage of an interval estimation is that we can expect, with the significant level of confidence, the true value of the parameter will lie within a certain range. It also helps us not to be confident that the true value is exactly equal to a given 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	Number of obs	=	308
Residual	166.152715	306	.542982728	F(1, 306)	=	92.20
				Prob > F	=	0.0000
				R-squared	=	0.2315
				Adj R-squared	=	0.2290
Total	216.213584	307	.704278775	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

a) $\hat{Y}_i = 7.658082 + 0.0318017 X_i$

$\therefore$ On average, the estimated nominal wage for a person who works 0 hour a week is about 7.6 baht. #

b) $\widehat{lnwage}_i = 7.658082 + 0.0318017\, main\_hr$

> $\frac{d\,\widehat{lnwage}_i}{d\,main\_hr} = 0.0318017 \Rightarrow \frac{d\,wage}{wage} = 0.0318017\, d\,main\_hr$

> multiply both sides by 100 ; $\boxed{\%\Delta \widehat{wage}_i = 0.0318017\ \cancel{d\,main\_hr}^{1} \cdot 100}$

$\therefore$ If a person works an hour more, we expect that wage to increase by 3.18017 % #

c) Assumption : If a person works an hour more, the nominal wage on average increases 0.0318017 baht. #

- converted hr worked to day worked by multiplied x by $\frac{1}{24}$, therefore,

> $\hat{\beta}_2 = \frac{0.0318017}{\frac{1}{24}} = 0.7632$ #

> $se(\hat{\beta}_2) = \frac{0.003312}{\frac{1}{24}} = 0.0795$ #

> 95% CI $\Rightarrow \alpha = 0.05 \Rightarrow t_{\frac{\alpha}{2}} = 2.021$

- the lower limit : $0.7632 - (2.021 \cdot 0.0795) = 0.6025$ #

- the upper limit : $0.7632 + (2.021 \cdot 0.0795) = 0.9239$ #