6204640251
Pimmada K.

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

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

Answer the following questions. Show your work.

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

Ordinary Least Square

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 $\hat{\sigma}\text{ar}(u_i)$, $\hat{\sigma}\text{ar}(\hat{\beta}_1)$ and $\hat{\sigma}\text{ar}(\hat{\beta}_2)$

two Tail Test? α

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

$1-\alpha$
$1-0.05$
0.95

α

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

$1-\alpha$
$1-0.01$
0.99

$$\mathrm{var}(U_i) = \frac{\sum \hat{U}_i^2}{n-k} = \frac{\sum(Y_i-\hat{Y})^2}{n-k}$$

$$\mathrm{var}(\hat{\beta}_1) = \frac{\sum X_i^2}{n\sum x_i^2}\sigma^2$$

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

1.a.) Ordinary Least Square (OLS)

6204640251
Pimmada K.

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i$$

Rearrange : $\hat{u}_i = Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_i)$

Set Obj. Function: $\sum \hat{u}_i^2 = \sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$

$\underset{\hat{\beta}_1, \hat{\beta}_2}{\text{Min}} \sum \hat{u}_i^2$: set them to zero

Slove for $\hat{\beta}_1$: $\frac{\partial}{\partial \hat{\beta}_1}\left(\sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2\right) = \cancel{2}\sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$ —— ①

$\frac{\partial}{\partial \hat{\beta}_2}\left(\sum (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2\right) = \cancel{2}\sum X_i (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$ —— ②

From ① ; $\sum Y_i - \sum \hat{\beta}_1 - \sum \hat{\beta}_2 X_i = 0$

$$\sum Y_i - n\hat{\beta}_1 - \hat{\beta}_2 \sum X_i = 0$$

$$n\hat{\beta}_1 = \sum Y_i - \hat{\beta}_2 \sum X_i$$

$$\hat{\beta}_1 = \frac{\sum Y_i}{n} - \hat{\beta}_2 \frac{\sum X_i}{n}$$

$\therefore \hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X}$ —— ③ #

Plug ③ into ② ; $\sum X_i (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i) = 0$

$$\sum X_i (Y_i - (\bar{Y} - \hat{\beta}_2 \bar{X}) - \hat{\beta}_2 X_i) = 0$$

$$\sum X_i (Y_i - \bar{Y} + \hat{\beta}_2 \bar{X} - \hat{\beta}_2 X_i) = 0$$

$$\sum X_i (Y_i - \bar{Y} - \hat{\beta}_2 (X_i - \bar{X})) = 0$$

$$\sum X_i (Y_i - \bar{Y}) - \sum X_i \hat{\beta}_2 (X_i - \bar{X}) = 0$$

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

$\therefore \hat{\beta}_2 = \frac{-174.2}{1098.8} = -0.1585$ #

$\hat{\beta}_1 = 21.03 - (-0.1585)(12.20) = 22.9637$ #

The OLS estimators are expressed in terms of observables.

They are point estimators instead of interval estimators,

They made SRF passes through sample mean.

Mean value of $\hat{Y}_i$ is $\bar{Y}$ and for residual value $\hat{U}_i$ is zero.

$\hat{U}_i$ are uncorrelated with X & $\hat{Y}$.

1.b.) Coefficient of determination (r^2)

$$r^2 = \frac{(\Sigma x_i y_i)^2}{\Sigma x_i^2 \Sigma y_i^2} = \frac{(\Sigma(x_i - x)(y_i - y))^2}{\Sigma(x_i - x)^2 \Sigma(y_i - y)^2}$$

Sample $$r = \frac{\Sigma x_i y_i}{\sqrt{\Sigma(x_i)^2 \Sigma(y_i)^2}}$$

r^2 is determined by how much is described SRF or the measurement of godness of fit of the fitted regression line compared to the estimator $\bar{Y}$.

$$r^2 = \frac{(\Sigma x_i y_i)^2}{\Sigma x_i^2 \Sigma y_i^2} = \frac{(\Sigma(x_i - x)(y_i - y))^2}{\Sigma(x_i - x)^2 \Sigma(y_i - y)^2}$$

$$= \frac{(-174.20)}{(1098.8)(882.97)} = 0.0313 \quad \#$$

1.c. $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

from a.) $\hat{\beta}_1 = 22.9637$, $\hat{\beta}_2 = -0.1585$

If $X_i = 5$; $\hat{Y}_i = 22.9637 - 0.1585(5)$

$\therefore \hat{Y}_i = 22.1712$ #

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

$$\text{var}(\hat{\beta}_1) = \frac{\sum X_i^2}{n \sum x_i^2}\sigma^2 = \frac{\sum X_i^2}{n\sum(X_i-\bar{X})^2}\sigma^2 = \frac{5564\ (31.1836)}{30\ (1098.8)} = 5.2635$$ #

$$\text{Var}(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2} = \frac{\sigma^2}{\sum(X_i-\bar{X})^2} = \frac{31.1836}{1098.8} = 0.0284$$ #

Different from zero → Two Tails Test

1.e. Step 1: State Hypothesis

$H_0 : \beta_2 = 0$ Null

$H_1 : \beta_2 \neq 0$ Alternative

$\alpha = 0.05$

$1 - \alpha = 1 - 0.05 = 0.95$

Step 2: $t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2^{0}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_2}} \rightarrow var} = \dfrac{\hat{\beta}_2 - \beta_2}{\sigma_{\hat{\beta}_2}} \quad \dfrac{-0.1585 - 0}{\sqrt{0.0284}} = -0.9407$

from a) $\hat{\beta}_2 = -0.1585$

from d.) $Var(\hat{\beta}_2) = \hat{\sigma}^2_{\hat{\beta}_2} = 0.0284$

$\therefore \sigma_{\hat{\beta}_2} = \sqrt{0.0284} = 0.1685$

Appendix D *Statistical Tables* 879

TABLE D.2
Percentage Points of the *t* Distribution

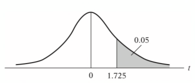
Source: From E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians,* vol. 1, 3d ed., table 12, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.

Example

$Pr(t > 2.086) = 0.025$

$Pr(t > 1.725) = 0.05$ for df = 20

$Pr(|t| > 1.725) = 0.10$



df \ Pr	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

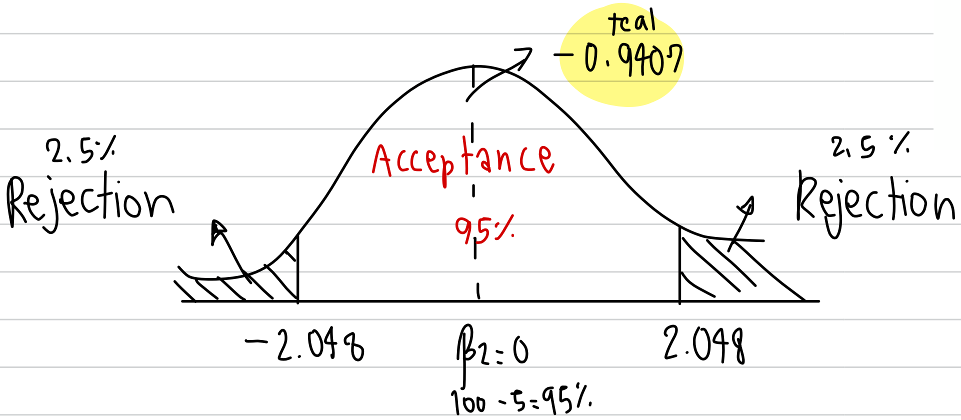
Note: The smaller probability shown at the head of each column is the area in one tail; the larger probability is the area in both tails.

step 3: upper bound $(+) t_{\frac{\alpha}{2}} = 2.048$

lower bound $(-) t_{\frac{\alpha}{2}} = -2.048$

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

$n - k = 30 - 2 = 28$



Step 4: Conclude the test

Tcal lies <
beyond reject → reject
within accept → cannot reject

∴ We "cannot reject" Ho at significant level of 95%. We cannot say for sure that β2 is not zero 95 out of 100 times when we sample.

1.F.) "Less than zero ONE TAIL TEST"

STEP1 : STATE HYPOTHESIS

H0: B2 < 0 — Null

H1: B2 ⩾ 0 — Alternative

$\alpha = 0.01$

Step 2 $\quad t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2^{\nearrow 0}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}_2}}} = \dfrac{-0.1585 - 0}{\sqrt{0.0284}} = -0.9407$

$\sqrt{\hat{\sigma}^2_{\hat{\beta}_2}} = \hat{\sigma}_{\hat{\beta}_2}$

step 3 $\quad \beta_2 = 0 \quad > \quad \alpha = 0.01$

Pr / df	0.25 / 0.50	0.10 / 0.20	0.05 / 0.10	0.025 / 0.05	0.01 / 0.02	0.005 / 0.010	0.001 / 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396

Lower Bound : $\beta_2^{\nearrow 0} - t_{\frac{\alpha}{2}} \cdot \hat{\sigma}_{\beta_2}$

& Acceptance on the RIGHT

$- t_{0.01,\, n-k} \cdot \hat{\sigma}_{\beta_2}$

$- t_{0.005,\, 30-2}^{2} \cdot \sqrt{\hat{\sigma}^2_{\beta_2}}$

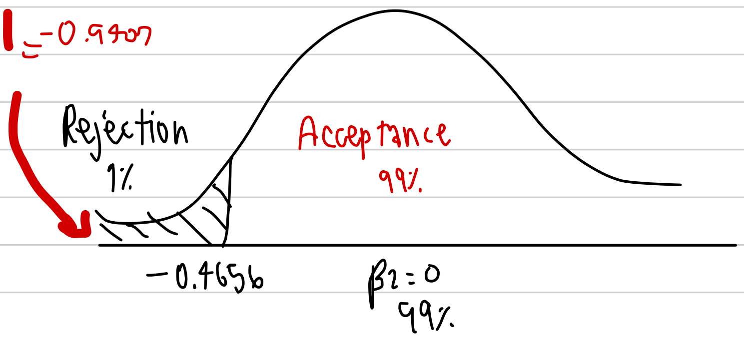
$- 2.763 \cdot \sqrt{0.0284}$

$- 2.763 \cdot 0.1685$

$- 0.4656$

* If null hypothesis is
  - LESS THAN α : acceptance region is on LEFT — หา UPPER BOUND : $\beta_2 + t_{\frac{\alpha}{2}} \cdot \hat{\sigma}_{\beta_2}$
  - MORE THAN α : acceptance region is on RIGHT — หา LOWER BOUND : $\beta_2 - t_{\frac{\alpha}{2}} \cdot \hat{\sigma}_{\beta_2}$

tcal $= -0.9407$



Step 4: Conclude the test

Tcal lies <
- beyond reject → reject
- within accept → cannot reject

∴ We "can reject" Ho at significance level of 99%. We are sure that β2 is less than zero 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

$\bar{X} = 7.45$,

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

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

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? $X_0 = X_1$

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. $x_i = 10$

$$\beta_2 = \frac{\Delta y}{\Delta x}$$

$$\varepsilon = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y} = \hat{\beta}_2 \cdot \frac{x}{y}$$

2.a.) $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i \Rightarrow \hat{Y}_i = 7{,}826 - 502.4X_i$

Negative sign (-502.4) of $\hat{\beta}_2$ makes economic sense since it reflects the slope. Every year the car was used, it depreciates its value.

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

$$= 212{,}877\left[\frac{1}{11} + \frac{(5 - 7.45)^2}{78.73}\right]$$

$$= 35{,}582.5353$$

Step 2: $\sigma_{\hat{Y}_0} = \sqrt{\sigma^2_{\hat{Y}_0}} = \sqrt{35{,}582.5353} = 188.6333$

จากโจทย์: $\hat{Y}_0 = 7{,}836 - 502.4X_i \xrightarrow{} 5 = 5{,}324$

Step 3: $t_{cal} = t_{\frac{\alpha}{2}} = t_{\frac{0.05}{2}} = t_{0.025\,9}$ $df = n-k = 11-2 = 9$

$= 2.262$

Pr / df	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144

Upper Price: $\hat{Y} = \hat{Y}_0 + (t_{\frac{\alpha}{2}, n-k})(\sigma_{\hat{Y}_0})$

$\hat{Y} = 5{,}324 + 2.262(188.6333)$

$= 5{,}750.6885$

Lower Price: $\hat{Y} = \hat{Y}_0 - (t_{\frac{\alpha}{2}, n-k})(\sigma_{\hat{Y}_0})$

$\hat{Y} = 5{,}324 - 2.262(188.6333)$

$= 4{,}897.3115$

$$1-\alpha = \text{Prediction}\left[\underbrace{\hat{Y}_0 - (t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{Y}_0})}_{\text{Lower Price}} \leq Y_0 \leq \underbrace{\hat{Y}_0 + (t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{Y}_0})}_{\text{Upper price}}\right]$$

∴ So market price of 5 years car is between \$ 4,897.3115 and \$ 5,750.6885 at 95% Confidence Interval (CI). #

2.c.) multiply all x with 10

$$\bar{X}_{new} = 10\,\bar{X}_{old} = 10 \cdot 7.45 = 74.5$$

$$\hat{\beta}_{2\,new} = \frac{\hat{\beta}_{2\,old}}{10} = \frac{-502.4}{10} = -50.24$$

$$\sigma_{\hat{\beta}_2\,new} = \frac{\sigma_{\hat{\beta}_2\,old}}{10} = \frac{411.8}{10} = 41.18$$

$$\therefore \hat{Y}_0 = 7{,}836 - 50.24\,X_{i\,new} \quad \#$$

2.d.)

$$\epsilon = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y} = \hat{\beta}_2 \cdot \frac{x}{y}$$

Find $\hat{y}$ at $X_i = 10$

$$\hat{y} = 7{,}836 - (502.4)(10) = 2{,}812$$

$$\therefore \varepsilon = -502.4 \left(\frac{10}{2{,}812}\right) = -1.7866$$

$$|\varepsilon| = 1.7866 > 1 \Rightarrow \text{Elastic} \quad \#$$