IE 522 HW08

Q1

1 Point

Suppose you want to predict the price of an asset using the following model:
$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \epsilon$,
where X_2 is the value of a factor, X_3 is the value of another factor and X_4 is the average value of the two factors. What assumption is violated for this linear regression model? Why?

It violates the "No muliticollinearity assumption", since X4 is the lineart combination of X2 and X3.

Choose Files No file chosen Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2

9 Points

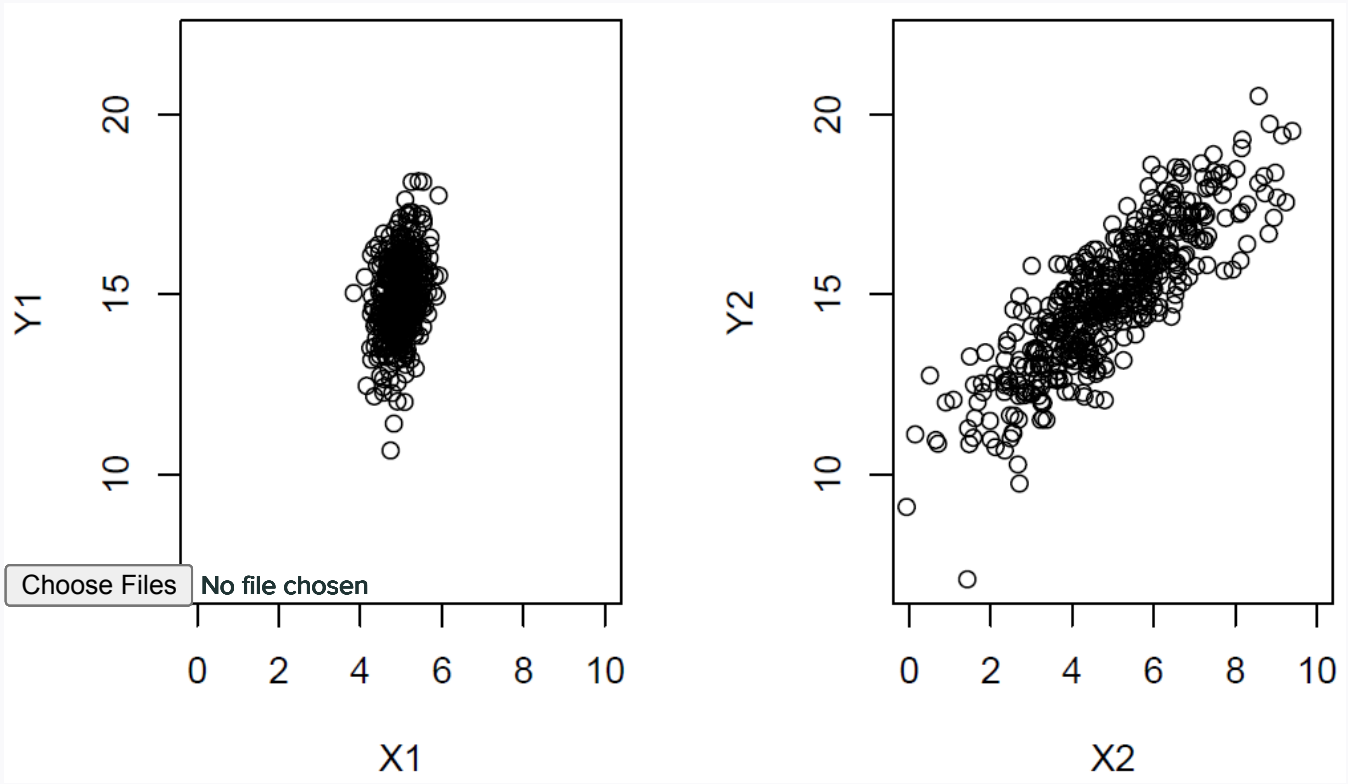
Consider the simple linear regression model $Y = \beta_1 + \beta_2 X + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$. There are two sets of data of the same size $n = 500$ from this model: $\{(X_{1i}, Y_{1i}), 1 \leq i \leq$

$n\}$ and $\{(X_{2i}, Y_{2i}), 1 \leq i \leq n\}$.

Q2.1

1 Point

According to the following scatter plots, which data set will give a more precise estimate for β_2? Why?



Data set 2 (the one on the right hand side) has more precise estimation.

CURRENTLY UPLOADED FILES

Q2-1.png

Download ✖ Remove

2-1

$$se(\hat{\beta}_1 | X) = \sqrt{\frac{\sigma^2}{S_{XX}}} = \frac{\hat{\sigma}}{\sqrt{S_{XX}}}$$

⇒ as X_i's more spreading out, $\hat{\beta}_1$ has better estimation with β_1.

Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Choose Files No file chosen

Q2.2

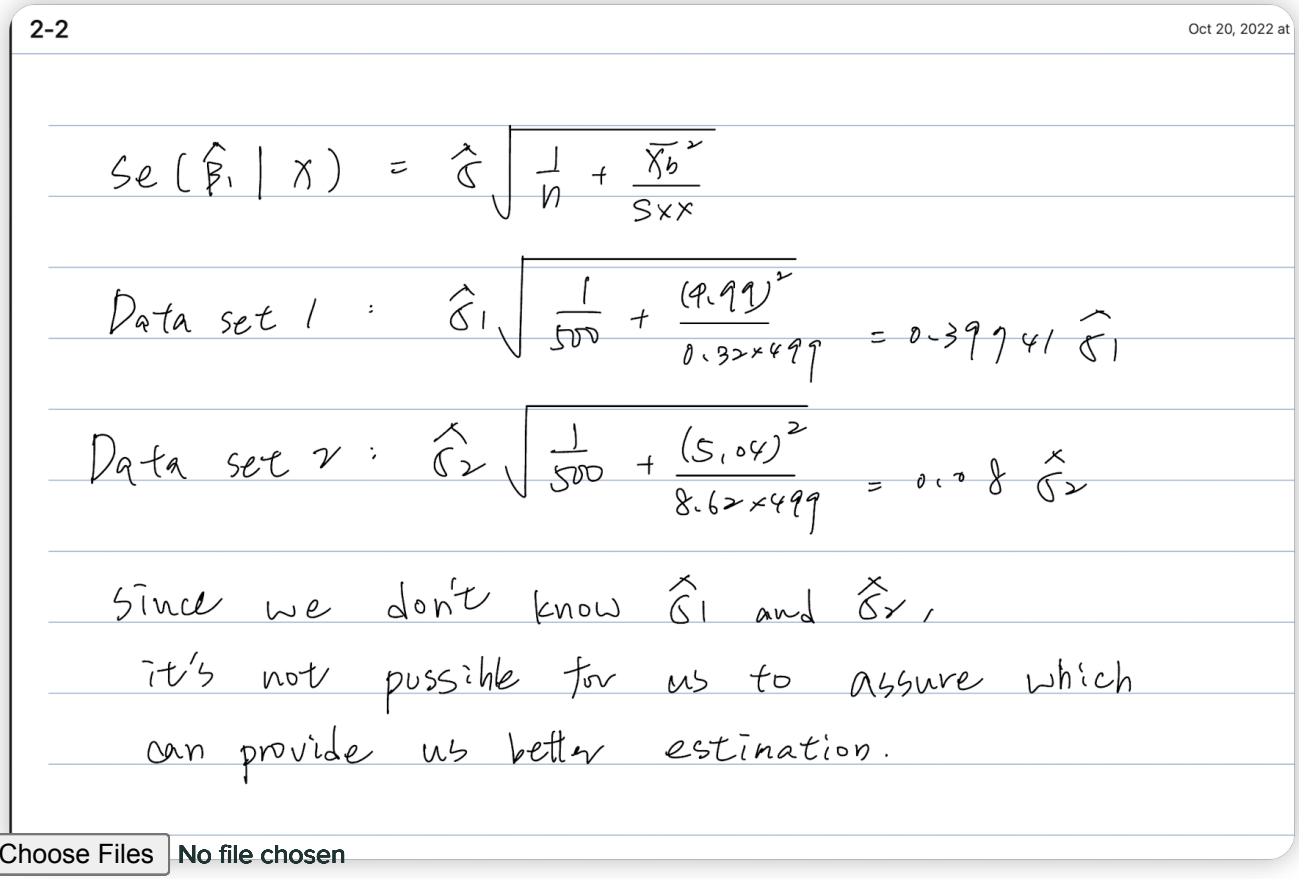
1 Point

The sample mean and sample variance of $\{X_{1i}, 1 \leq i \leq n\}$ are 4.99 and 0.32, respectively. The sample mean and sample variance of $\{X_{2i}, 1 \leq i \leq n\}$ are 5.04 and 8.62, respectively. Which data set will give a more precise estimate for β_1? Why?

We are not able to define which data set has better estimation for beta 1.

CURRENTLY UPLOADED FILES

Q2-2.png Download Remove



Choose Files No file chosen

Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.3

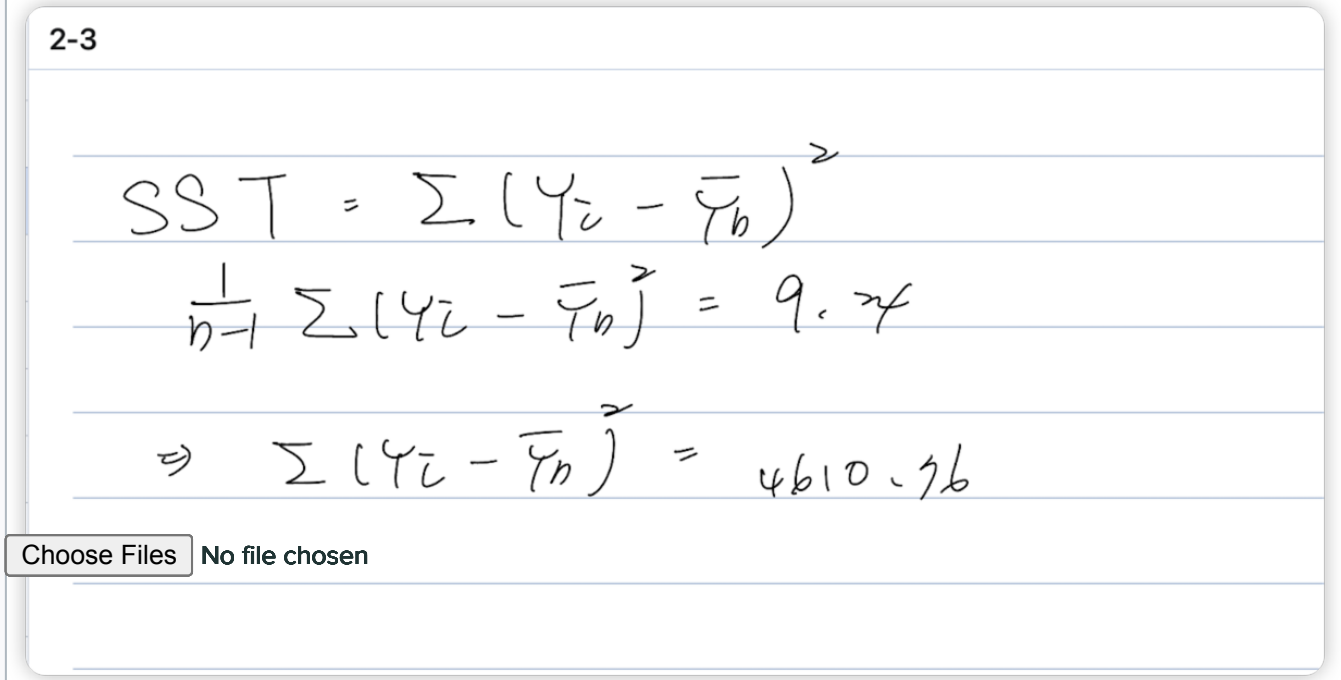
1 Point

In the following, $\{(X_{2i}, Y_{2i}), 1 \leq i \leq n\}$ is used to fit the simple linear regression model. The sample variance of $\{Y_{2i}, 1 \leq i \leq n\}$ is 9.24. What's the total sum of squares?

SST = 4610.76.

CURRENTLY UPLOADED FILES

Q2-3.png Download Remove



Choose Files No file chosen

Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.4

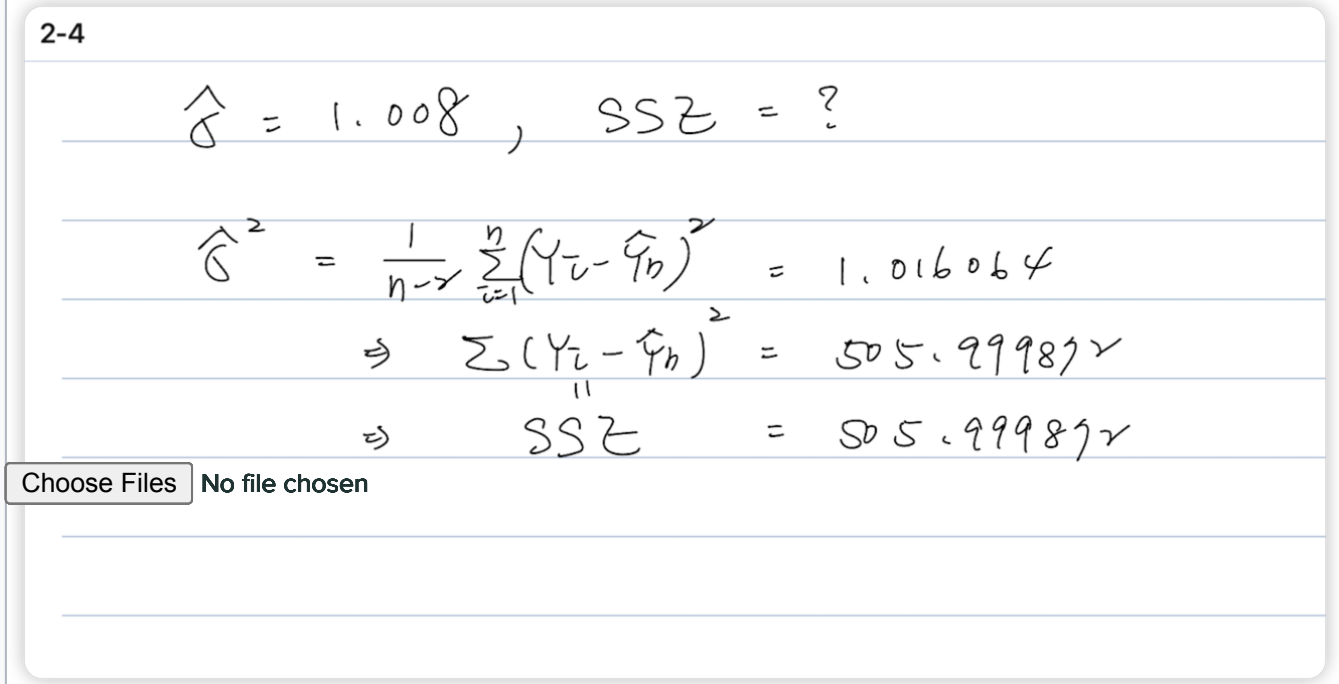
1 Point

The residual standard error is 1.008. What's the corresponding error sum of squares?

SSE = 505.999872

CURRENTLY UPLOADED FILES

Q2-4.png Download Remove



Choose Files No file chosen

Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.5

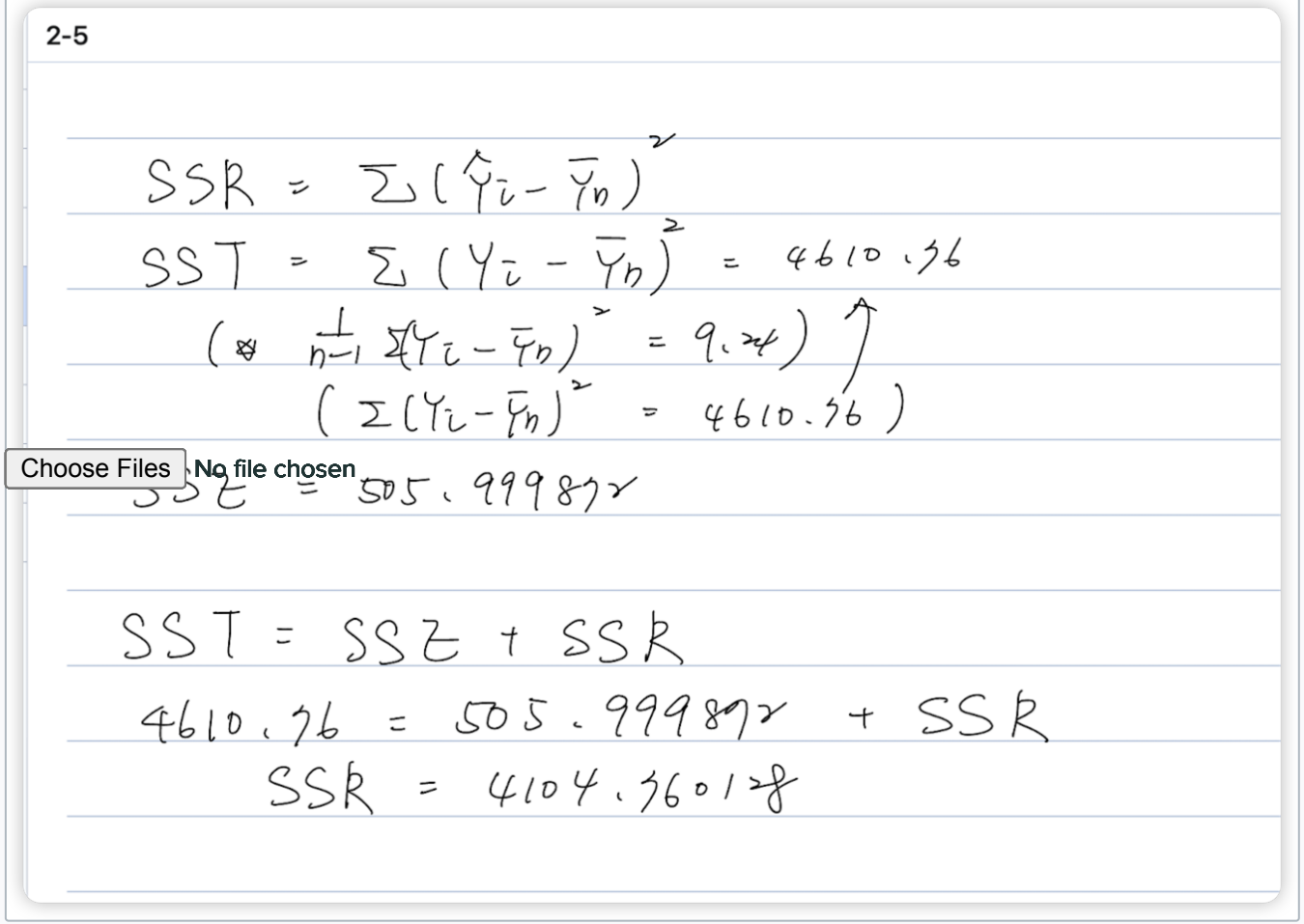
1 Point

What's the regression sum of squares?

SSR = 4104.760128.

CURRENTLY UPLOADED FILES

Q2-5.png Download Remove



Choose Files No file chosen

Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.6

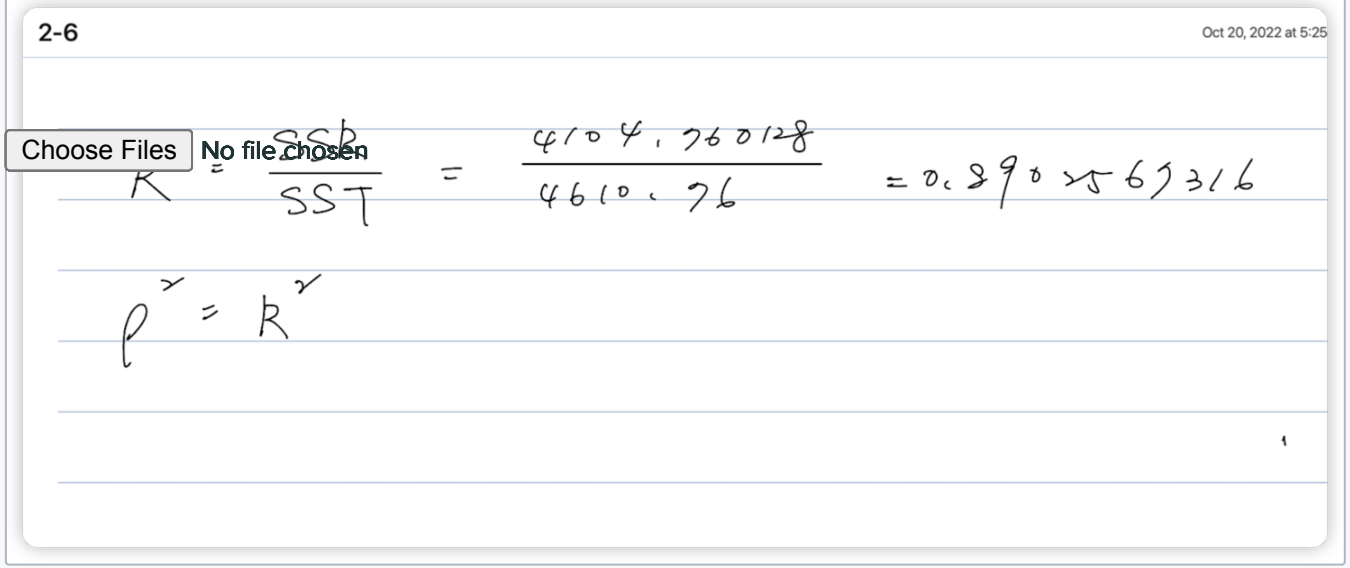
1 Point

What's R^2? What's the sample correlation coefficient between $\{X_{2i}, 1 \leq i \leq n\}$ and $\{Y_{2i}, 1 \leq i \leq n\}$?

R^2 = zho^2 = 0.8902567316

CURRENTLY UPLOADED FILES

Q2-6.png Download Remove



Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.7

1 Point

The sample mean of $\{Y_{2i}, 1 \leq i \leq n\}$ is 15.03. What are $\hat{\beta}_1$ and $\hat{\beta}_2$?

beta 1 hat = 10.10653877
beta 2 hat = 0.9768772279

CURRENTLY UPLOADED FILES

Q2-7.png Download Remove

Choose Files No file chosen

$\bar{Y}_n = 15.03, \quad \bar{X}_n = 5.04$

$\frac{1}{n-1}\sum(Y_i - \bar{Y}_n)^2 = 9.24 \longrightarrow \sum(Y_i - \bar{Y}_n)^2 = 4610.76$

$\frac{1}{n-1}\sum(X_i - \bar{X}_n)^2 = 8.62 \longrightarrow \sum(X_i - \bar{X}_n)^2 = 4301.38$

$\downarrow$

$S_{XX} = 4301.38$

$R^2 = \frac{S_{XY}^2}{S_{XX} SST} \Rightarrow S_{XY} = \sqrt{R^2 S_{XX} SST}$

$= (0.890567316 \times 4301.38 \times 4610.76)^{1/2}$

$\Rightarrow S_{XY} = 4201.92017$

Choose Files No file chosen

$\frac{4201.92017}{4301.38} = \frac{S_{XY}}{S_{XX}} = 0.976877239$

$\hat{\beta}_1 = \bar{Y}_n - \hat{\beta}_2 \bar{X}_n = 15.03 - \hat{\beta}_2 \times 5.04$

$= 10.1065387$

Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.8

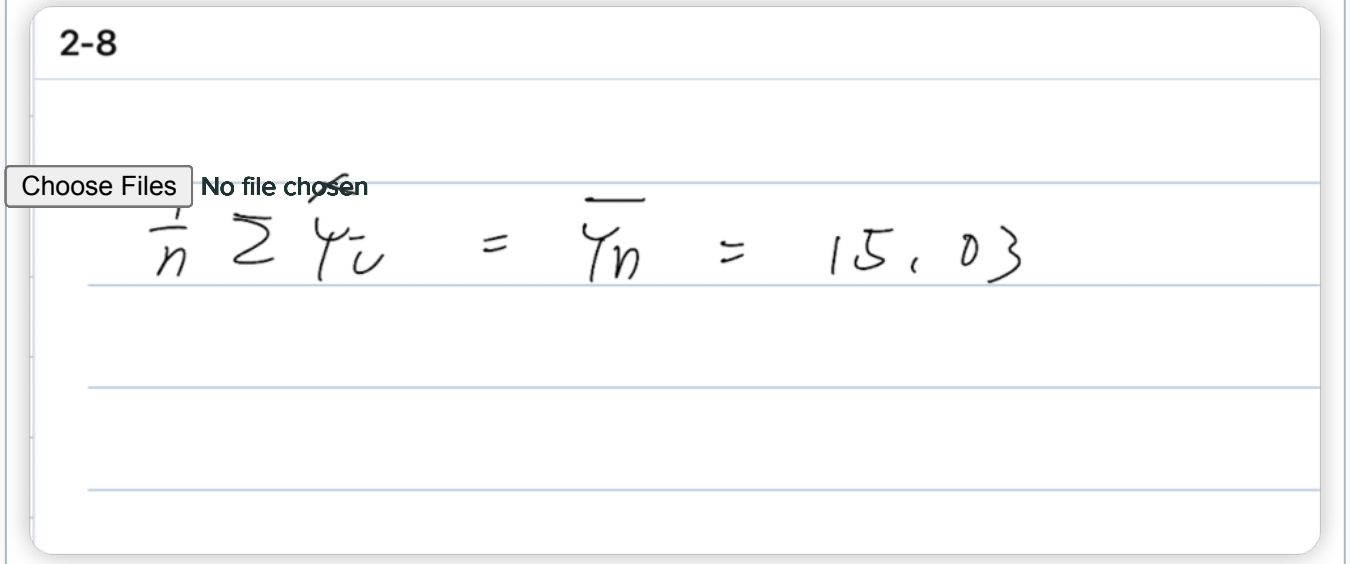
1 Point

What's the average of the fitted values $\{\hat{Y}_{2i}, 1 \leq i \leq n\}$?

15.03

CURRENTLY UPLOADED FILES

Q2-8.png Download Remove



Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Q2.9

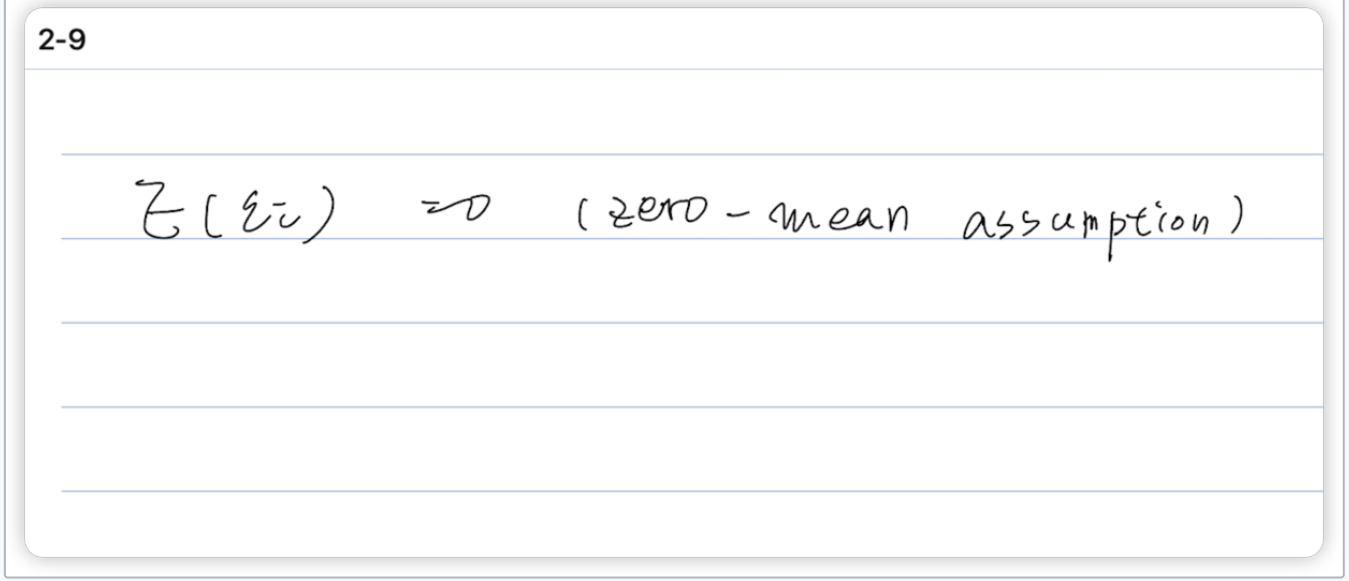
1 Point

What's the average of the residuals?

0

CURRENTLY UPLOADED FILES

Q2-9.png Download Remove



Please select file(s) Select file(s)

Save Answer Last saved on **Oct 20 at 10:02 PM**

Save All Answers

Submit & View Submission