

MATH 4780 (MSSC 5780) In-class Activity 3

October 8, 2023

Samantha and Edward, please present your work on Oct 10. Remember to ask undergraduate students two Non-yes-no questions. You, as a teacher, try to get students understand your work. First clearly state the question, then show your work to answer the question.

Intepretable intercept (Samantha)

Quite often the intercept in the regression model is meaningless because the predictor value is never zero.

1. Rewrite the simple linear regression model as

$$y = \alpha_0 + \beta_1(x - \bar{x}) + \epsilon$$

where $\alpha = \beta_0 + \beta_1\bar{x}$

2. Show that the sampled fitted line can be written as

$$\hat{y} = \bar{y} + b_1(x - \bar{x})$$

where b_1 is the LSE for β_1 .

3. Pick a data set whose predictor is never zero, for example, heights, car horsepower, etc, then fit the rewritten model, and explain the meaning of intercept and slope.

Weighted average of individual slopes (Edward)

For each data pair (x_i, y_i) and (x_j, y_j) , $i \neq j$, we can compute the slope of the line connecting them:

$$\text{slope}_{ij} = \frac{y_j - y_i}{x_j - x_i}.$$

In fact, the weighted average of slope_{ij} weighted by the squared separation $(x_j - x_i)^2$ is the least squares estimator b_1 , i.e.,

$$b_1 = \frac{\sum_{i,j} (x_j - x_i)^2 \text{slope}_{ij}}{\sum_{i,j} (x_j - x_i)^2}.$$

Demonstrate how this work *numerically and graphically* in a case with three data points, $(x, y) = (0, 0), (4, 1), (5, 5)$.