Classwork 4: More on Deriving OLS Estimators

Recall that in the videos from workbook chapter 3, we took the definition of the method of least squares and we derived that $\hat{\beta}_1 = \frac{\sum_i (x_i y_i) - n\bar{x}\bar{y}}{\sum_i (x_i^2) - n\bar{x}^2}$.

1. Show that an equivalent formula for $\hat{\beta}_1$ is: $\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$. *Hint: start with* $\frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$ *and show that is equal to* $\frac{\sum_i (x_i y_i) - n\bar{x}\bar{y}}{\sum_i (x_i^2) - n\bar{x}^2}$. *(2 points)*

2. Consider this new formula: $\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$.

   (a) Recall that the formula for the estimate of the variance of a random variable X is $\hat{\sigma}_X^2 = \frac{\sum_i (x_i - \bar{x})^2}{n-1}$ and the formula for the estimate of the covariance of two random variables X and Y is $\hat{\sigma}_{XY} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n-1}$. Then it must be true that $\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2} = \frac{}{}$. Fill in the blanks for the numerator and denominator with functions of $\hat{\sigma}_X^2$ and $\hat{\sigma}_{XY}$ and make the expression as simple as possible. *(1 point)*

   (b) If X and Y are independent (so $\sigma_{XY} = 0$), what should we expect $\hat{\beta}_1$ to be? *(1 point)*

   (c) If $\hat{\sigma}_{XY}$ is greater than 0, will $\hat{\beta}_1$ be greater than 0? *(1 point)*

3. Use the new formula for $\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2}$ to show that $\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x}) y_i}{\sum_i (x_i - \bar{x})^2}$. *(2 points)*

4. In this question, you'll explore what insight about OLS this new formula $\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x}) y_i}{\sum_i (x_i - \bar{x})^2}$ gives us.

   (a) Let observation i's "weight" or "leverage" be $w_i = \frac{x_i - \bar{x}}{\sum_i (x_i - \bar{x})^2}$. This lets us see that $\hat{\beta}_1$ is a weighted average of y_i: $\hat{\beta}_1 = \sum_i w_i y_i$. Let $x_i = \{-10, 0, 10\}$ and fill out the table below for w_i, since w_i is a function of x_i. Leave the y_i column blank for now. *(1 point)*

x_i	w_i	y_i
-10	_	
0	_	
10	_	

   (b) Suppose $y_i = \{0, 5, 10\}$. Draw a plot of the three data points (x_i, y_i) and sketch the line of best fit. Then calculate $\hat{\beta}_1$ using the formula $\sum_i w_i y_i$ and verify that it's the same as the slope of the line of best fit in your drawing. *(1 point)*

   (c) Consider two other possibilities for y_i: calculate $\hat{\beta}_1$ when we take the second observation and move it up, so $y_i = \{0, 10, 10\}$, and then calculate $\hat{\beta}_1$ when we take the first observation and move it up, so $y_i = \{10, 5, 10\}$. The lesson: $\hat{\beta}_1$ will never be sensitive to outliers with low leverage, that is, outliers near $\bar{x}$. It's the high leverage outliers that we need to be most concerned about. *(1 point)*