# Chapter 2 The Regression Line

Arthur Chen

January 17, 2023

## Exercise Set A

### Problem 1

In the Pearson-Lee data, the average height of the fathers was 67.78 inches; the SD was 2.74 inches. The average height of the sons was 68.7 inches; the SD was 2.81 inches. The correlation was .501.

#### Part a

True or false and explain: because the sons average an inch taller than the fathers, if the father is 72 inches tall, it's 50-50 whether the son is taller than 73 inches.

False. The statement as given states that given the point $(\bar{x}, \bar{y})$, moving to $\bar{x} + c$ causes the conditional mean of $y$ to increase to $\bar{y} + c$. This is incorrect. The correct statement is that by moving to $\bar{x} + SD_x$, the conditional mean of $y$ increases to $\bar{y} + rSD_y$.

#### Part b

Find the regression line of son's height on father's height, and its RMS error.

The line is

$$y - 68.7 = .501\frac{2.81}{2.74}(x - 67.7)$$

From the formulas, the MSE is

$$MSE = (1 - .501^2)(2.81)^2 = 1.966$$

and RMS is $\sqrt{1.966} = 1.402$.