Homework 6 Chapters 7 and 12 (Problems 1 to 6)

Due on March 24

Math 125 *Kovitz* Spring 2025

The key problem on this assignment is problem 3.

1. A small sample (size 20) of men led to the following results (the scatter diagram is football-shaped):

$$\begin{aligned} \text{average height} \approx 175 \text{ cm}, &\quad \text{SD} \approx 10 \text{ cm} \\ \text{average weight} \approx 83 \text{ kg}, &\quad \text{SD} \approx 16 \text{ kg}, \quad r \approx 0.48 \end{aligned}$$

   (a) Find the regression equation for predicting height from weight. Use it to predict the height of a 100-kg. man.

   (b) What is the slope of that equation? Explain how it could be intepreted in the context of predicting the height of a man who is 30 kilograms heavier than his friend.

   (c) A prediction made by this equation is likely to be off by about how many centimeters?

2. (a) Plot four different points whose x-coordinates are half their y-coordinates. Do these points lie on a line? If so, what is the equation of the line in the standard form: $y = mx + b$? Look at the point when $y = 14$, and show that the equation wiil produce the correct y when the x is plugged in..

   (b) Plot the points $(-1, 4)$, $(-2, 6.5)$, $(-3, 9)$, $(-4, 11.5)$ on the same graph. These points all lie on a line. What is the equation of this line in the standard form: $y = mx+b$? Find the point on the line when $x = 1$, and plot it. Does the plotted point appear to lie on the line?

3. Find the regression equation for predicting final score from midterm score, based on the following information.

$$\begin{aligned} \text{average midterm score} = 65, &\quad \text{SD} = 15 \\ \text{average final score} = 60, &\quad \text{SD} = 20, \quad r \approx 0.60 \end{aligned}$$

4. A statistician is doing a study on a group of undergraduates. On average, these students drink 4 beers a month, with an SD of 8. They eat 4 pizzas a month, with an SD of 4. There is some positive association between beer and pizza, and the regression equation is

$$\text{predicted number of beers} = \underline{\qquad\qquad} \times \text{number of pizzas} + 2.$$

   However, the statistician lost the data and forgot the slope of the equation. (Perhaps he had too much beer and pizza?). Can you help him remember the slope? Explain.

5. A researcher wants to use a straight line to predict income from height, for a large group of residents of a certain state. There is a weak positive association in the data. True or false, and explain—

   (a) He has to use the regression line.

   (b) He can use many different lines.

   (c) Any line he uses will have an r.m.s. error.

   (d) Only the regression line has an r.m.s. error.

   (e) Among all lines, the regression line has the smallest r.m.s. error.

6. True or false:

   (a) If a line does not go through the point (average x, average y), it cannot be the regression line for y on x.

   (b) If a line does not go through the point of averages (the point mentioned in part (a)), it cannot be the SD line for y on x.