

1 The regression line

We know how to create a scatterplot for two variables and compute the correlation coefficient r . Should there be a good correlation between two variables we may want to use one variable to predict the other. To do so we use a regression line. Other curves and functions can be utilized but we'll work with lines for now. The equation of a line is give as

$$y = a + bx. \tag{1}$$

One can also use function notations as well.

$$f(x) = a + bx \tag{2}$$

Example 1 *Sean starts playing video poker with an initial stake of \$80 and loses an average of \$2.50 on each game. Write the equation of the regression line for predicting the amount of money Sean has after playing x games.*

$$f(x) = 80 - 2.50x$$

Example 2 *After 20 games of video poker how much money does Sean have? $f(20) = 80 - 2.50 * 20 = 30.0$.*

Example 3 *When does Sean go broke?*

$$80 - 2.50 * x < 0 \tag{3}$$

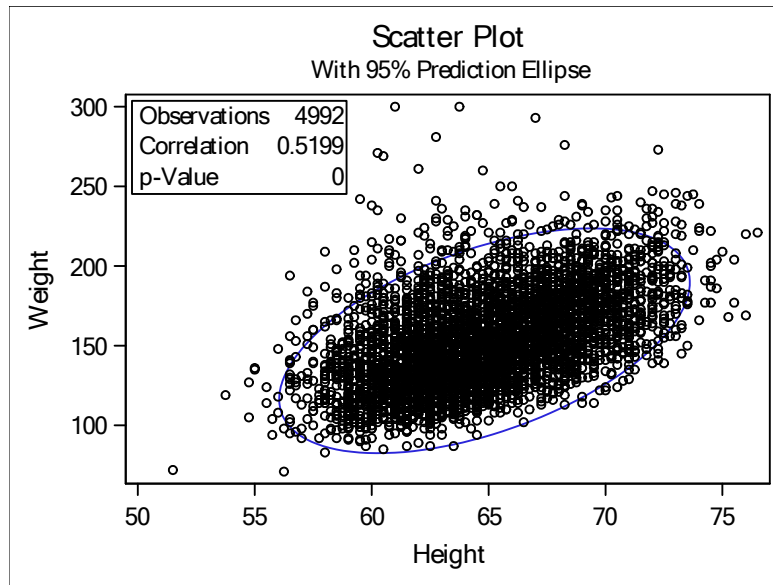
$$80 < 2.5 * x \tag{4}$$

$$\frac{80}{2.5} < x \tag{5}$$

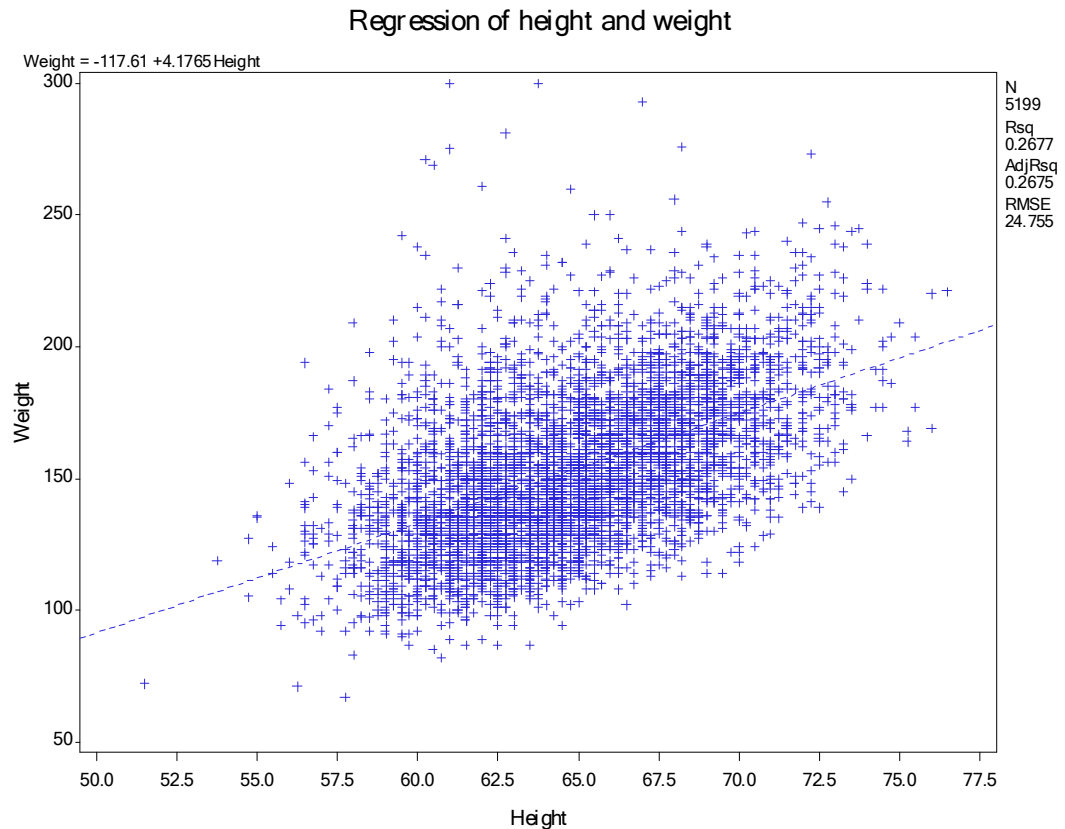
$$32 < x \tag{6}$$

After 32 games Sean is broke.

Now let's focus on using the regression line and interpreting its results. Consider the health data set from before. There is a medium correlation between height and weight.



How can we measure this relationship as a line? We can use technology to find a line of best fit through the data.



How useful is the regression line? That depends on the amount of correlation between the variables. The correlation coefficient $r = 0.5199$ indicates a medium positive correlation of the two variables. The square of the correlation coefficient also provides information. The value r^2 indicates the percentage of data in y that is explained strictly by the linear relationship between x and y . Thus, $1 - r^2$ is the percentage of data with variation from the linear relationship. In this instance $r^2 = .2677$ of the one's weight is directly correlated to height. So, $1 - .2677 = 0.7323$ of the one's weight is variation from height. The equation of our regression line is $f(x) = 4.1765x - 117.61$.

Problem 4 *What is the predicted weight of someone who is 5'7"?* $4.1765 * 67 - 117.61 = 162.22$

Problem 5 *What is the predicted weight of someone who is 5'2"?*

Problem 6 *What is the predicted height of someone who weighs 170 lbs?* $170 = 4.1765x - 117.61$, *Solution is:* 68.864

Problem 7 *What is the predicted height of someone who weighs 125 lbs?*

Problem 8 *If two people differ in height by one inch, how much do you predict they will differ in weight?*

Problem 9 *Does the prediction of the weight of a person who is 24 inches tall make sense (https://en.wikipedia.org/wiki/List_of_the_verified_shortest_people)?*

2 Homework

1. Navidi/Monk Section 11.2: 13-16, 21, 23