Discussion 6

9/3/2024

Dan Chen

Practice problem (10.1.7)

Which of the following is not a property of correlation, r?

A. $-1 \leq r \leq 1$

B. Correlation measures the strength of a linear relationship between two quantitative variables.

C. The sign on r tells the direction of the linear relationship between two quantitative variables.

D. If the correlation between two quantitative variables is zero, then there is no relationship between these two variables.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Practice problem (10.1.10)

For each of the following statements, say what, if anything, is wrong.

a. Because the correlation coefficient between test time and test score is −0.56, the correlation coefficient between test score and test time must be 0.56 (the positive value).

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

b. There is a strong positive correlation between a person's yard size and whether or not they have a dog.

c. For a sample of 50 students, the correlation coefficient between weight (kg) and height (inches) was found to be 0.78 kg/inches.

d. A correlation coefficient of r = 0.84 denotes just as strong a relationship as a correlation coefficient of r = −0.84.

Linear Regression

$$y = \alpha + \beta x$$

$$\widehat{\alpha} = \bar{y} - (\widehat{\beta}\,\bar{x}),$$

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = r_{xy}\frac{s_y}{s_x}$$

Standard deviation of y

Correlation

Standard deviation of x

Practice problem (10.3.15)

Harris and Steudel (2002) studied factors that might be associated with the jumping performance of domestic cats. They studied 18 cats, using takeoff velocity (in centimeters per second) as the response variable. They used body mass (in grams), hind limb length (in centimeters), muscle mass (in grams), and percent body fat in addition to sex as potential explanatory variables. The data can be found in the CatJumping data file. *(See Discussion 6.Rmd)*

a. Describe the association between these variables.

b. Use R to determine the equation of the least squares line for predicting a cat's takeoff velocity from its mass.

c. Interpret the value of the slope coefficient in this context.

d. Interpret the value of the intercept coefficient. Is this a context in which the intercept coefficient is meaningful?

e. Determine the proportion of variability in takeoff velocity that is explained by the least squares line with mass.