2.6 The Least-Squares Regression Line (Part 2)

It is also possible to calculate the equation of the least-squares regression line using the means and standard deviations of each variable, along with their correlation.

$\hat{y} = a + bx$

How to Calculate the Least-Squares Regression Line Using Summary Statistics

If $\bar{x}$ and s_x are the mean and standard deviation of the explanatory variable, $\bar{y}$ and s_y are the mean and the standard deviation of the response variable, and r is the correlation between the variables:

$$\text{slope} = b = r\frac{s_y}{s_x} \qquad y\text{ intercept} = a = \bar{y} - b\bar{x}$$

Example: In the last section, we used tapping time to predict the amount of soda remaining in a vigorously shaken can. For these cans, the mean tapping time was 6 seconds, with a standard deviation of 4.53 seconds. The mean soda remaining was 264.45 milliliters, with a standard deviation of 12.92 milliliters. The correlation between tapping time and soda remaining was r = 0.924. Calculate the equation of the least-squares regression line for predicting the amount of soda remaining (y) from tapping time (x).

$\bar{x} = 6 \quad s_x = 4.53 \quad \bar{y} = 264.45 \quad s_y = 12.92$

$b = \text{slope} = r\frac{s_y}{s_x} = (0.924)\left(\frac{12.92}{4.53}\right) = 2.635$

$y\text{-int}: \bar{y} - b\bar{x} = 264.45 - (2.635)(6) = 248.64$

$$\hat{y} = 248.64 + 2.635x$$