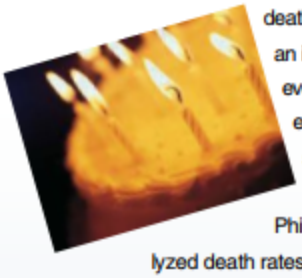
Postponing Death

Several studies addressed the ability of people to postpone their death until after an important event. For example, sociologist David Phillips analyzed death rates of Jewish men who died near Passover, and he found that the death rate dropped dramatically in the week before Passover, but rose the week after. Other researchers of cancer patients concluded that there is "no pattern to support the concept that 'death takes a holiday.'" (See "Holidays, Birthdays, and Postponement of Cancer Death," by Young and Hade, *Journal of the American Medical Association*, Vol. 292, No. 24.) Based on records of 1.3 million deaths, this more recent study found no relationship between the time of death and Christmas, Thanksgiving, or the person's birthday. The findings were disputed by David Phillips, who said that the study focused on cancer patients, but they are least likely to have psychosomatic effects.



s_y is the standard deviation of the sample y values, and s_x is the standard deviation of the sample x values.

$$b_1 = r\frac{s_y}{s_x} = 0.591269 \cdot \frac{4.87391}{1.66823} = 1.72745$$

After finding the slope b_1, we can now use Formula 10-4 to find the y-intercept as follows:

$$b_0 = \bar{y} - b_1\bar{x} = 177.3 - (1.72745)(30.04) = 125.40740$$

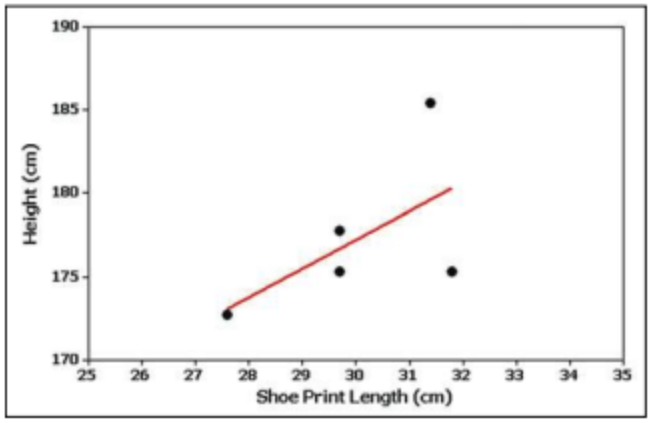
After rounding, the slope and y-intercept are $b_1 = 1.73$ and $b_0 = 125$. We can now express the regression equation as $\hat{y} = 125 + 1.73x$, where $\hat{y}$ is the predicted height of a person and x is the length of the shoe print.

Example 3 Graphing the Regression Line

Graph the regression equation $\hat{y} = 125 + 1.73x$ (found in Examples 1 and 2) on the scatterplot of the shoe print and height data from Table 10-1 and examine the graph to subjectively determine how well the regression line fits the data.

Solution

Shown below is the Minitab display of the scatterplot with the graph of the regression line included. We can see that the regression line doesn't fit the data very well.



Using the Regression Equation for Predictions

Regression equations are often useful for *predicting* the value of one variable, given some specific value of the other variable. When making predictions, we should consider the following:

1. Use the regression equation for predictions only if the graph of the regression line on the scatterplot confirms that the regression line fits the points reasonably well.
2. Use the regression equation for predictions only if the linear correlation coefficient r indicates that there is a linear correlation between the two variables (as described in Section 10-2).
3. Use the regression line for predictions only if the data do not go much beyond the scope of the available sample data. (Predicting too far beyond the scope of the available sample data is called *extrapolation*, and it could result in bad predictions.)
4. If the regression equation does not appear to be useful for making predictions, the best predicted value of a variable is its sample mean.