using TECHNOLOGY

Because of the messy calculations involved, the linear correlation coefficient r and the slope and y-intercept of the regression line are usually found using a calculator or computer software.

STATDISK First enter the paired data in columns of the Statdisk Data Window. Select **Analysis** from the main menu bar, then use the option **Correlation and Regression.** Enter a value for the significance level and select the columns of data. Click on the **Evaluate** button. The display will include the value of the linear correlation coefficient along with the critical value of r and the P-value, the conclusion about correlation, and the intercept and slope of the regression equation, as well as some other results. Click on **Plot** to get a graph of the scatterplot with the regression line included.

MINITAB First enter the x values in column C1 and enter the y values in column C2 (or use any other columns). In Section 10-2 we saw that we could find the value of the linear correlation coefficient r by selecting **Stat/Basic Statistics/Correlation.** To get the equation of the regression line, select **Stat/Regression/Regression,** and enter C2 for "response" and C1 for "predictor." To get the graph of the scatterplot with the regression line, select **Stat/Regression/Fitted Line Plot,** then enter C2 for the response variable and C1 for the predictor variable. Select the "linear" model.

With Minitab Release 16 or later, you can also click on **Assistant,** then select **Regression.** Click on the icon to perform the analysis, then complete the dialog box that appears. There will be much helpful information displayed.

EXCEL Enter the paired data in columns A and B. Use either XLSTAT or Excel's Data Analysis add-in.

XLSTAT Click on **XLSTAT,** then select **Modeling Data,** then **Linear Regression.** In the dialog box that appears, first enter the range of the sample values for the dependent y variable in the box identified as "Y/Dependent variable." Next, enter the range of quantitative values for the independent x variable in the box identified as "X/Explanatory variable." Put a checkmark next to the "Variable labels" box only if the first row of each column consists of names or labels. Click **OK.** After the results are displayed, look for the value of the linear correlation coefficient r in the "Correlation matrix" table. The P-value can be found in the last column of the "Analysis of Variance" table. The equation of the regression line is identified as "Equation of the model."

Data Analysis add-in: If using Excel 2103, 2010, or 2007, click on **Data,** then click on **Data Analysis;** if using Excel 2003, click on **Tools,** then click on **Data Analysis.** Select **Regression,** then click on **OK.** Enter the range for the y values, such as B1:B10. Enter the range for the x values, such as A1:A10. Click on the box adjacent to Line Fit Plots, then click **OK.** Among all of the information provided by Excel, the slope and intercept of the regression equation can be found under the table heading "Coefficient." The displayed graph will include a scatterplot of the original sample points along with the points that would be predicted by the regression equation. You can easily get the regression line by connecting the "predicted y" points.

TI-83/84 PLUS Enter the paired data in lists L1 and L2, then press **STAT** and select **TESTS,** then choose the option **LinRegTTest.** The displayed results will include the y-intercept and slope of the regression equation. Instead of b_0 and b_1, the TI-83/84 display represents these values as a and b.

STATCRUNCH Click on **Open StatCrunch.** Enter the columns of data or open a data set. Click on **Stat,** then select **Regression,** then select **Simple Linear.** Enter the columns to be used, then click on **Calculate.** The equation of the regression line will be found in the fourth line of the display.

10-3 Basic Skills and Concepts

Statistical Literacy and Critical Thinking

1. Notation and Terminology If we use the paired height/pulse data for females from Data Set 1 in Appendix B, we get this regression equation: $\hat{y} = 73.9 + 0.0223x$, where x represents height (cm) and the pulse rate is in beats per minute. What does the symbol $\hat{y}$ represent? In this case, what does the predictor variable represent? What does the response variable represent?

2. Best-Fit Line In what sense is the regression line the straight line that "best" fits the points in a scatterplot?

3. Correlation and Slope Formula 10-3 shows that the slope of a regression line can be found by evaluating $r \cdot s_y/s_x$. What do we know about the graph of the regression line if r is a positive value? What do we know about the graph of the regression line if r is a negative value?

4. Notation What is the difference between the regression equation $\hat{y} = b_0 + b_1x$ and the regression equation $y = \beta_0 + \beta_1x$?