value is not included among the algebraic operations.) Although the use of absolute values would be simple and easy, it would create algebraic difficulties in inferential methods of statistics discussed in later chapters. For example, Section 9-3 presents a method for making inferences about the means of two populations, and that method is built around an additive property of variances, but the mean absolute deviation has no such additive property. (Here is a simplified version of the additive property of variances: If you have two independent populations and you randomly select one value from each population and add them, such sums will have a variance equal to the sum of the variances of the two populations.) Also, the mean absolute deviation is a *biased* estimator, meaning that when you find mean absolute deviations of samples, you do not tend to target the mean absolute deviation of the population. The standard deviation has the advantage of using only algebraic operations. Because it is based on the square root of a sum of squares, the standard deviation closely parallels distance formulas found in algebra. There are many instances where a statistical procedure is based on a similar sum of squares. Therefore, instead of using absolute values, we square all deviations $(x - \bar{x})$ so that they are nonnegative. This approach leads to the standard deviation. For these reasons, scientific calculators typically include a standard deviation function, but they almost never include the mean absolute deviation.

Why Divide by $n - 1$? After finding all of the individual values of $(x - \bar{x})^2$, we combine them by finding their sum. We then divide by $n - 1$ because there are only $n - 1$ independent values. With a given mean, only $n - 1$ values can be freely assigned any number before the last value is determined. Exercise 45 illustrates that division by $n - 1$ yields a better result than division by n. That exercise shows how division by $n - 1$ causes the sample variance s^2 to target the value of the population variance σ^2, whereas division by n causes the sample variance s^2 to underestimate the value of the population variance σ^2.

How Do We Make Sense of a Value of Standard Deviation? Part 1 of this section included the range rule of thumb for interpreting a known value of a standard deviation or estimating a value of a standard deviation. (See Examples 4 and 5.) We now discuss two other approaches for interpreting standard deviation: the empirical rule and Chebyshev's theorem.

Empirical (or 68–95–99.7) Rule for Data with a Bell-Shaped Distribution

A concept helpful in interpreting the value of a standard deviation is the **empirical rule.** This rule states that *for data sets having a distribution that is approximately bell-shaped,* the following properties apply. (See Figure 3-3 on the next page.)

- About 68% of all values fall within 1 standard deviation of the mean.
- About 95% of all values fall within 2 standard deviations of the mean.
- About 99.7% of all values fall within 3 standard deviations of the mean.

Example 6 **The Empirical Rule**

IQ scores have a bell-shaped distribution with a mean of 100 and a standard deviation of 15. What percentage of IQ scores are between 70 and 130?

Solution

The key to solving this problem is to recognize that 70 and 130 are each exactly 2 standard deviations away from the mean of 100, as shown below:.

$$2 \text{ standard deviations} = 2s = 2(15) = 30$$