Example 3 Calculating Standard Deviation with Formula 3-5

Use Formula 3-5 to find the standard deviation of the sample values 22, 22, 26, and 24 from Example 1.

Solution

Shown below is the computation of the standard deviation of 22, 22, 26, and 24 using Formula: 3-5.

$n = 4$ (because there are 4 values in the sample)

$\Sigma x = 94$ (found by adding the sample values: $22 + 22 + 26 + 24 = 94$)

$\Sigma x^2 = 2220$ (found by adding the squares of the sample values, as in $22^2 + 22^2 + 26^2 + 24^2 = 2220$)

Using Formula 3-5, we get

$$s = \sqrt{\frac{n(\Sigma x^2) - (\Sigma x)^2}{n(n-1)}} = \sqrt{\frac{4(2220) - (94)^2}{4(4-1)}} = \sqrt{\frac{44}{12}} = 1.9 \text{ chips}$$

The result of $s = 1.9$ chips is the same as the result in Example 2.

Comparing Variation in Different Samples Examples 1, 2, and 3 used only four values (22, 22, 26, 24) so that calculations are relatively simple. If we use all 80 chocolate chip counts depicted in Figure 3-2, we get the measures of center and measures of variation listed in Table 3-3. From the table we see that the range for Chips Ahoy cookies is much larger than the range for Triola cookies. Table 3-3 also shows that the standard deviation for Chips Ahoy is much larger than the standard deviation for Triola cookies, but *it's a good practice to compare two sample standard deviations only when the sample means are approximately the same.* When comparing variation in samples with very different means, it is better to use the coefficient of variation, which is defined later in this section. We also use the coefficient of variation when we want to compare variation from two samples with different scales or units of values, such as the comparison of variation of heights of men and weights of men. (See Example 8, which involves numbers of chocolate chips and weights of cola.)

Table 3-3 Comparison of Chocolate Chips in Cookies from Figure 3-2

	Chips Ahoy (regular)	Triola Chips
Number	40	40
Mean	24.0	24.0
Range	11.0	2.0
Standard Deviation	2.6	0.7

CAUTION *Compare two sample standard deviations only when the sample means are approximately the same.* When comparing variation in samples with very different means, it is better to use the coefficient of variation, which is defined later in this section.

Where Are the 0.400 Hitters?

The last baseball player to hit above 0.400 was Ted Williams, who hit 0.406 in 1941. There were averages above 0.400 in 1876, 1879, 1887, 1894, 1895, 1896, 1897, 1899, 1901, 1911, 1920, 1922, 1924, 1925, and 1930, but none since 1941. Are there no longer great hitters? The late Stephen Jay Gould of Harvard University noted that the mean batting average has been steady at 0.260 for about 100 years, but the standard deviation has been decreasing from 0.049 in the 1870s to 0.031, where it is now. He argued that today's stars are as good as those from the past, but consistently better pitchers now keep averages below 0.400.