

a larger sample size before they're willing to rely on  $n$  (as opposed to  $n - 1$ ) in the denominator. The approach in this text is to always use  $n - 1$  when calculating a sample standard deviation.

The issue at this point isn't when different statisticians invoke the correction factor and when they don't; the issue is why. Because the answer to that question is one that usually takes some serious thought, let me suggest that you take time out for one of those dark room moments I mentioned earlier.

First, take the time to give some serious thought to the ideas of variability and the standard deviation in general. Then take some time to think about how the standard deviation of a population is related to the standard deviation of a sample. Develop a mental picture of a population and a sample from that population. Mentally focus on why you would expect the standard deviation of the population to be slightly larger than the standard deviation of the sample. You should think about the relationship between the two long enough to fully appreciate why the correction factor is used. It all goes back to the point that the variability of a sample is going to be smaller than the variability of a population, and that's why a correction factor has to be used.

Finally, in an effort to make certain that you fully understand how to calculate the standard deviation of a sample, and the point about  $n - 1$  in the denominator, let me suggest that you take a close look at Table 2-15. It's an illustration of the calculation of the standard deviation for a sample. My suggestion is that you repeat each of the calculations shown in the illustration, working each step on your own, while also paying particular attention to the next to the last step (i.e., dividing by  $n - 1$  before you take the square root).

Assuming you feel comfortable about the different measures of central tendency and measures of variability (and the standard deviation, in particular), we

**Table 2-15** Calculating the Standard Deviation of a Sample

Scores/Values	Deviations	Squared Deviations	
( $N = 9$ )			
( $X$ )	( $X - \text{Mean}$ )		
7	(7 - 4)	3	9
1	(1 - 4)	-3	9
3	(3 - 4)	-1	1
5	(5 - 4)	1	1
6	(6 - 4)	2	4
2	(2 - 4)	-2	4
8	(8 - 4)	4	16
1	(1 - 4)	-3	9
3	(3 - 4)	-1	1
			54
Mean = 4			
			Sum of Squared Deviations = 54
			$54/8 = 6.75$
			Note that $n - 1$ or 8 is used
			Square Root of 6.75 = 2.598
			Standard Deviation = 2.598 or round to 2.60

can move forward. Next we look at data distributions—the work of the next chapter.

## Chapter Summary

In learning about measures of central tendency and variability, some of the fundamentals of statistics are introduced to the business of statistics. Symbols are used when referring to these concepts, and the connection to the practical world you've begun to understand is discussed. A sample statistic is discussed.

As to what you've learned, you have digested several points of information available, and each one is likely to be appropriate in one or more situations you've likely picked up on. A measure of central tendency—a measure that summarizes the data—for example, the mean is calculated, and the standard deviation is calculated.

On the variability or dispersion of the data, reduced to several different measures, you've learned that the standard deviation is complex, you've learned how to calculate it. You've also learned how to interpret the results of each other, and (ideally) how to use these measures in the statistical analysis.

Finally, you have learned about the personal preference in the choice of formula. You've countered different formulas, and you've learned that ideally suited for use with different data. You've learned that different preferences when using the formula for the standard deviation, perhaps, but they help explain the same statistical process.

## Some Other Things You Should Know

At this point, you deserve a break. You've presented in a variety of ways, and you've learned about yourself. The data distribution