Working with Large Data Sets

9.123 Stressed-Out Bus Drivers. Previous studies have shown that urban bus drivers have an extremely stressful job, and a large proportion of drivers retire prematurely with disabilities due to occupational stress. In the paper, "Hassles on the Job: A Study of a Job Intervention With Urban Bus Drivers" (*Journal of Organizational Behavior*, Vol. 20, pp. 199–208), G. Evans et al. examined the effects of an intervention program to improve the conditions of urban bus drivers. Among other variables, the researchers monitored diastolic blood pressure of bus drivers in downtown Stockholm, Sweden. The data, in millimeters of mercury (mm Hg), on the WeissStats site are based on the blood pressures obtained prior to intervention for the 41 bus drivers in the study. Use the technology of your choice to do the following.

a. Obtain a normal probability plot, boxplot, histogram, and stem-and-leaf diagram of the data.

b. Based on your results from part (a), can you reasonably apply the one-mean t-test to the data? Explain your reasoning.

c. At the 10% significance level, do the data provide sufficient evidence to conclude that the mean diastolic blood pressure of bus drivers in Stockholm exceeds the normal diastolic blood pressure of 80 mm Hg?

9.124 How Far People Drive. In 2011, the average car in the United States was driven 13.5 thousand miles, as reported by the Federal Highway Administration in *Highway Statistics*. On the WeissStats site, we provide last year's distance driven, in thousands of miles, by each of 500 randomly selected cars. Use the technology of your choice to do the following.

a. Obtain a normal probability plot and histogram of the data.

b. Based on your results from part (a), can you reasonably apply the one-mean t-test to the data? Explain your reasoning.

c. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean distance driven last year differs from that in 2011?

9.125 Fair Market Rent. According to the document *Out of Reach*, published by the National Low Income Housing Coalition, the fair market rent (FMR) for a two-bedroom unit in the United States is \$949. A sample of 100 randomly selected two-bedroom units yielded the data on monthly rents, in dollars, given on the WeissStats site. Use the technology of your choice to do the following.

a. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean monthly rent for two-bedroom units is greater than the FMR of \$949? Apply the one-mean t-test.

b. Remove the outlier from the data and repeat the hypothesis test in part (a).

c. Comment on the effect that removing the outlier has on the hypothesis test.

d. State your conclusion regarding the hypothesis test and explain your answer.

Extending the Concepts and Skills

9.126 Two-Tailed Hypothesis Tests and CIs. The following relationship holds between hypothesis tests and confidence intervals for one-mean t-procedures: For a two-tailed hypothesis test at the significance level α, the null hypothesis H_0: $\mu = \mu_0$ will be rejected in favor of the alternative hypothesis H_a: $\mu > \mu_0$ if and only if μ_0 lies outside the $(1-\alpha)$-level confidence interval for μ. In each case, illustrate the preceding relationship by obtaining the appropriate one-mean t-interval (Procedure 8.2 on page 377) and comparing the result to the conclusion of the hypothesis test in the specified exercise.

a. Exercise 9.113 **b.** Exercise 9.116

9.127 Left-Tailed Hypothesis Tests and CIs. In Exercise 8.146 on page 384, we introduced one-sided one-mean t-intervals. The following relationship holds between hypothesis tests and confidence intervals for one-mean t-procedures: For a left-tailed hypothesis test at the significance level α, the null hypothesis H_0: $\mu = \mu_0$ will be rejected in favor of the alternative hypothesis H_a: $\mu < \mu_0$ if and only if μ_0 is greater than or equal to the $(1-\alpha)$-level upper confidence bound for μ. In each case, illustrate the preceding relationship by obtaining the appropriate upper confidence bound and comparing the result to the conclusion of the hypothesis test in the specified exercise.

a. Exercise 9.117 **b.** Exercise 9.118

9.128 Right-Tailed Hypothesis Tests and CIs. In Exercise 8.146 on page 384, we introduced one-sided one-mean t-intervals. The following relationship holds between hypothesis tests and confidence intervals for one-mean t-procedures: For a right-tailed hypothesis test at the significance level α, the null hypothesis H_0: $\mu = \mu_0$ will be rejected in favor of the alternative hypothesis H_a: $\mu > \mu_0$ if and only if μ_0 is less than or equal to the $(1-\alpha)$-level lower confidence bound for μ. In each case, illustrate the preceding relationship by obtaining the appropriate lower confidence bound and comparing the result to the conclusion of the hypothesis test in the specified exercise.

a. Exercise 9.114 (both parts)

b. Exercise 9.115

9.6 The Wilcoxon Signed-Rank Test*

Up to this point, we have presented two methods for performing a hypothesis test for a population mean. If the population standard deviation is known, we can use the z-test; if it is unknown, we can use the t-test.

Both procedures require another assumption for their use: The variable under consideration should be approximately normally distributed, or the sample size should be relatively large. For small samples, both procedures should be avoided in the presence of outliers.

In this section, we describe a third method for performing a hypothesis test for a population mean—the **Wilcoxon signed-rank test.**† This test, which is sometimes

†The Wilcoxon signed-rank text is also known as the **one-sample Wilcoxon signed-rank test** and the **one-variable Wilcoxon signed-rank test.**

more appropriate than either the z-test or the t-test, is an example of a *nonparametric method.*

What Is a Nonparametric Method?

Recall that descriptive measures for population data, such as μ and σ, are called parameters. Technically, inferential methods concerned with parameters are called **parametric methods;** those that are not are called **nonparametric methods.** However, common statistical practice is to refer to most methods that can be applied without assuming normality as nonparametric. Thus the term *nonparametric method* as used in contemporary statistics is a misnomer.

Nonparametric methods have both advantages and disadvantages. On one hand, they usually entail fewer and simpler computations than parametric methods and are resistant to outliers and other extreme values. On the other hand, they are not as powerful as parametric methods, such as the z-test and t-test, when the requirements for use of parametric methods are met.†

The Logic Behind the Wilcoxon Signed-Rank Test

The Wilcoxon signed-rank test is based on the assumption that the variable under consideration has a *symmetric distribution*—one that can be divided into two pieces that are mirror images of each other—but does not require that its distribution be normal or have any other specific shape. Thus, for instance, the Wilcoxon signed-rank test applies to a variable that has a normal, triangular, uniform, or symmetric bimodal distribution but not to one that has a right-skewed or left-skewed distribution. The next example explains the reasoning behind this test.

EXAMPLE 9.18 Introducing the Wilcoxon Signed-Rank Test

Weekly Food Costs The U.S. Department of Agriculture publishes information about food costs in *Agricultural Research Service*. According to that document, a typical U.S. family of three spends about \$157 per week on food. Ten randomly selected Kansas families of three have the weekly food costs shown in Table 9.13. Do the data provide sufficient evidence to conclude that the mean weekly food cost for Kansas families of three is less than the national mean of \$157?

TABLE 9.13
Sample of weekly food costs ($)

143	169	149	135	161
138	152	150	141	159

Solution Let μ denote the mean weekly food cost for all Kansas families of three. We want to perform the hypothesis test

$$H_0\text{: } \mu = \$157 \text{ (mean weekly food cost is not less than \$157)}$$
$$H_a\text{: } \mu < \$157 \text{ (mean weekly food cost is less than \$157).}$$

As we said, a condition for the use of the Wilcoxon signed-rank test is that the variable under consideration have a symmetric distribution. If the weekly food costs for Kansas families of three have a symmetric distribution, a graphic of the sample data should be roughly symmetric.

FIGURE 9.23
Stem-and-leaf diagram of sample data in Table 9.13

13	5 8
14	1 3
14	9
15	0 2
15	9
16	1
16	9

Figure 9.23 shows a stem-and-leaf diagram of the sample data in Table 9.13. The diagram is roughly symmetric and so does not reveal any obvious violations of the symmetry condition.‡ We therefore apply the Wilcoxon signed-rank test to carry out the hypothesis test.

To begin, we rank the data in Table 9.13 according to distance and direction from the null hypothesis mean, $\mu_0 = \$157$. The steps for doing so are presented in Table 9.14.

†A precise definition of *power* is presented in Section 9.7.

‡For ease in explaining the Wilcoxon signed-rank test, we have chosen an example in which the sample size is very small. This selection, however, makes it difficult to effectively check the symmetry condition. In general, we must proceed cautiously when dealing with very small samples.

TABLE 9.14
Steps for ranking the data in Table 9.13 according to distance and direction from the null hypothesis mean

Cost ($) x	Difference $D = x - 157$	$\lvert D \rvert$	Rank of $\lvert D \rvert$	Signed rank R
143	−14	14	7	−7
138	−19	19	9	−9
169	12	12	6	6
152	−5	5	3	−3
149	−8	8	5	−5
150	−7	7	4	−4
135	−22	22	10	−10
141	−16	16	8	−8
161	4	4	2	2
159	2	2	1	1

Step 1 *Subtract μ_0 from x.*

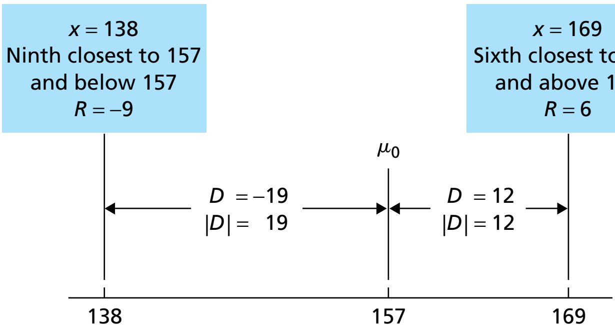
Step 2 *Make each difference positive by taking absolute values.*

Step 3 *Rank the absolute differences in order from smallest (1) to largest (10).*

Step 4 *Give each rank the same sign as the sign in the Difference column.*

The absolute differences, $|D|$, displayed in the third column, identify how far each observation is from 157. The ranks of those absolute differences, displayed in the fourth column, show which observations are closer to 157 and which are farther away. The signed ranks, R, displayed in the last column, indicate in addition whether an observation is greater than 157 (+) or less than 157 (−). Figure 9.24 depicts the information for the second and third rows of Table 9.14.

FIGURE 9.24
Meaning of signed ranks for the observations 138 and 169



The reasoning behind the Wilcoxon signed-rank test is as follows: If the null hypothesis, $\mu = \$157$, is true, then, because the distribution of weekly food costs is symmetric, we expect the sum of the positive ranks and the sum of the negative ranks to be roughly the same in magnitude. For the sample size of 10, the sum of all the ranks must be $1 + 2 + \cdots + 10 = 55$, and half of 55 is 27.5.

Thus, if the null hypothesis is true, we expect the sum of the positive ranks (and the sum of the negative ranks) to be roughly 27.5. If the sum of the positive ranks is too much smaller than 27.5, we conclude that the null hypothesis is false and, therefore, that the mean weekly food cost is less than \$157. From the last column of Table 9.14, the sum of the positive ranks, which we call W, equals $6 + 2 + 1 = 9$. This value is much smaller than 27.5 (the value we would expect if the mean is \$157).

The question now is, can the difference between the observed and expected values of W be reasonably attributed to sampling error, or does it indicate that the mean weekly food cost for Kansas families of three is actually less than \$157? We answer that question and complete the hypothesis test after we discuss some prerequisite material.

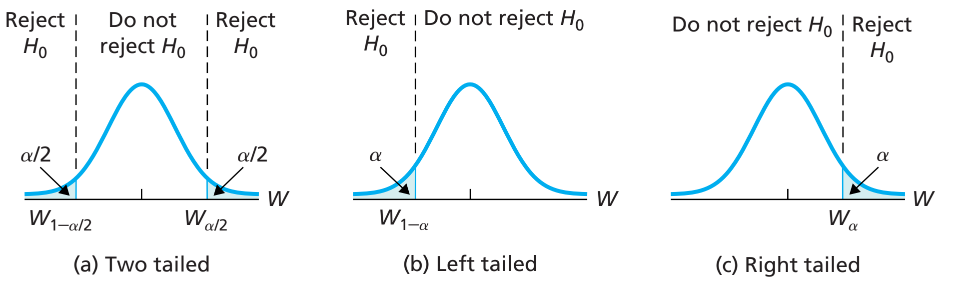
Using the Wilcoxon Signed-Rank Table†

Table V in Appendix A gives values of W_α for a Wilcoxon signed-rank test.‡ The two outside columns of Table V give the sample size, n. As expected, the symbol $\boldsymbol{W_\alpha}$ denotes the W-value with area (percentage, probability) α to its right. Thus the column headed $W_{0.10}$ contains W-values with area 0.10 to their right, the column headed $W_{0.05}$ contains W-values with area 0.05 to their right, and so on.

We can express the critical value(s) for a Wilcoxon signed-rank test at the significance level α as follows:

- For a two-tailed test, the critical values are the W-values with area $\alpha/2$ to its left (or, equivalently, area $1 - \alpha/2$ to its right) and area $\alpha/2$ to its right, which are $W_{1-\alpha/2}$ and $W_{\alpha/2}$, respectively. See Fig. 9.25(a).
- For a left-tailed test, the critical value is the W-value with area α to its left or, equivalently, area $1 - \alpha$ to its right, which is $W_{1-\alpha}$. See Fig. 9.25(b).
- For a right-tailed test, the critical value is the W-value with area α to its right, which is W_α. See Fig. 9.25(c).

FIGURE 9.25
Critical value(s) for a Wilcoxon signed-rank test at the significance level α if the test is (a) two tailed, (b) left tailed, or (c) right tailed



Note the following:

- A critical value from Table V is to be included as part of the rejection region.
- Although the variable W is discrete, we drew the "histograms" in Fig. 9.25 in the shape of a normal curve. This approach is not only convenient, it is also acceptable because W is close to normally distributed except for very small sample sizes. We use this graphical convention throughout this section.

The distribution of the variable W is symmetric about $n(n+1)/4$. This characteristic implies that the W-value with area A to its left (or, equivalently, area $1 - A$ to its right) equals $n(n+1)/2$ minus the W-value with area A to its right. In symbols,

$$W_{1-A} = n(n+1)/2 - W_A. \tag{9.1}$$

Referring to Fig. 9.25, we see that by using Equation (9.1) and Table V, we can determine the critical value for a left-tailed Wilcoxon signed-rank test and the critical values for a two-tailed Wilcoxon signed-rank test. The next example illustrates the use of Table V to determine critical values for a Wilcoxon signed-rank test.

EXAMPLE 9.19 Using the Wilcoxon Signed-Rank Table

In each case, use Table V to determine the critical value(s) for a Wilcoxon signed-rank test. Sketch graphs to illustrate your results.

a. Sample size = 12; significance level = 0.01; right tailed
b. Sample size = 14; significance level = 0.10; left tailed
c. Sample size = 10; significance level = 0.05; two tailed

---

†We can use the Wilcoxon signed-rank table to estimate the P-value of a Wilcoxon signed-rank test. Because doing so can be awkward or tedious, however, using statistical software is preferable. Thus, those concentrating on the P-value approach to hypothesis testing can skip to the subsection "Performing the Wilcoxon Signed-Rank Test" on the next page.

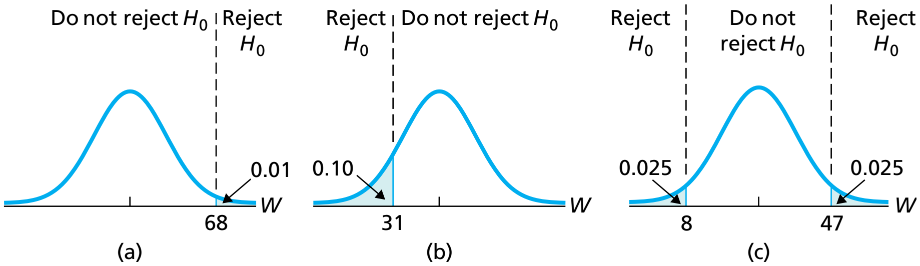
‡Actually, the α-levels in Table V are only approximate but are used in practice.

Solution In solving these problems, it helps to refer to Fig. 9.25.

a. The critical value for a right-tailed test at the 1% significance level is $W_{0.01}$. To find the critical value, we use Table V. First we go down the outside columns, labeled n, to "12." Then, going across that row to the column labeled $W_{0.01}$, we reach 68, the required critical value. See Fig. 9.26(a).

FIGURE 9.26

Critical value(s) for a Wilcoxon signed-rank test: (a) right tailed, $\alpha = 0.01$, $n = 12$; (b) left tailed, $\alpha = 0.10$, $n = 14$; (c) two tailed, $\alpha = 0.05$, $n = 10$



b. The critical value for a left-tailed test at the 10% significance level is $W_{1-0.10}$. To find the critical value, we use Table V and Equation (9.1). First we go down the outside columns, labeled n, to "14." Then, going across that row to the column labeled $W_{0.10}$, we reach 74; thus $W_{0.10} = 74$. Now we apply Equation (9.1) and the result just obtained to get

$$W_{1-0.10} = 14(14 + 1)/2 - W_{0.10} = 105 - 74 = 31,$$

which is the required critical value. See Fig. 9.26(b).

c. The critical values for a two-tailed test at the 5% significance level are $W_{1-0.05/2}$ and $W_{0.05/2}$, that is, $W_{1-0.025}$ and $W_{0.025}$. First we use Table V to find $W_{0.025}$. We go down the outside columns, labeled n, to "10." Then, going across that row to the column labeled $W_{0.025}$, we reach 47; thus $W_{0.025} = 47$. Now we apply Equation (9.1) and the result just obtained to get $W_{1-0.025}$:

$$W_{1-0.025} = 10(10 + 1)/2 - W_{0.025} = 55 - 47 = 8.$$

See Fig. 9.26(c).

Exercise 9.139 on page 440

Performing the Wilcoxon Signed-Rank Test

Procedure 9.3 on the next page provides a step-by-step method for performing a Wilcoxon signed-rank test by using either the critical-value approach or the P-value approach. Note that we often use the phrase **symmetric population** to indicate that the variable under consideration has a symmetric distribution.

EXAMPLE 9.20 The Wilcoxon Signed-Rank Test

Weekly Food Costs Let's complete the hypothesis test of Example 9.18. A random sample of 10 Kansas families of three yielded the data on weekly food costs shown in Table 9.13 on page 430. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean weekly food cost for Kansas families of three is less than the national mean of $157?

Solution We apply Procedure 9.3.

Step 1 State the null and alternative hypotheses.

Let μ denote the mean weekly food cost for all Kansas families of three. Then the null and alternative hypotheses are, respectively,

$$H_0\text{: } \mu = \$157 \text{ (mean weekly food cost is not less than \$157)}$$
$$H_a\text{: } \mu < \$157 \text{ (mean weekly food cost is less than \$157).}$$

Note that the hypothesis test is left tailed.

PROCEDURE 9.3 Wilcoxon Signed-Rank Test

Purpose To perform a hypothesis test for a population mean, μ

Assumptions

1. Simple random sample
2. Symmetric population

Step 1 The null hypothesis is H_0: $\mu = \mu_0$, and the alternative hypothesis is

$$\begin{array}{ccccc} H_a\text{: } \mu \neq \mu_0 & \text{or} & H_a\text{: } \mu < \mu_0 & \text{or} & H_a\text{: } \mu > \mu_0 \\ \text{(Two tailed)} & & \text{(Left tailed)} & & \text{(Right tailed)} \end{array}.$$

Step 2 Decide on the significance level, α.

Step 3 Compute the value of the test statistic

$$W = \text{sum of the positive ranks}$$

and denote that value W_0. To do so, construct a work table of the following form.

Observation x	Difference $D = x - \mu_0$	$\lvert D\rvert$	Rank of $\lvert D\rvert$	Signed rank R
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

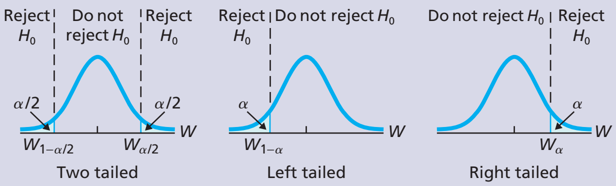
CRITICAL-VALUE APPROACH OR ***P*-VALUE APPROACH**

CRITICAL-VALUE APPROACH

Step 4 The critical value(s) are

$$\begin{array}{ccccc} W_{1-\alpha/2} \text{ and } W_{\alpha/2} & \text{or} & W_{1-\alpha} & \text{or} & W_{\alpha} \\ \text{(Two tailed)} & & \text{(Left tailed)} & & \text{(Right tailed)} \end{array}.$$

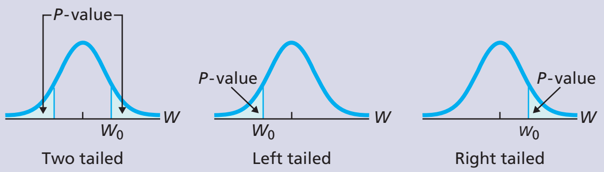
Use Table V to find the critical value(s). For a left-tailed or two-tailed test, you will also need the relation $W_{1-A} = n(n+1)/2 - W_A$.



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

***P*-VALUE APPROACH**

Step 4 Obtain the P-value by using technology.



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

Step 6 Interpret the results of the hypothesis test.

Step 2 Decide on the significance level, α.

The test is to be performed at the 5% significance level, or $\alpha = 0.05$.

Step 3 Compute the value of the test statistic

$$W = \text{sum of the positive ranks.}$$

The last column of Table 9.14 on page 431 shows that the sum of the positive ranks equals

$$W = 6 + 2 + 1 = 9.$$

CRITICAL-VALUE APPROACH OR *P*-VALUE APPROACH

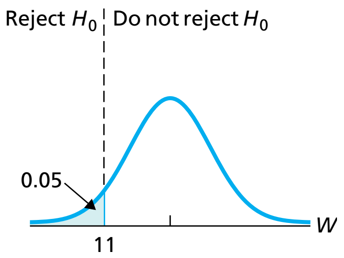
Step 4 The critical value for a left-tailed test is $W_{1-\alpha}$. Use Table V and the relation $W_{1-A} = n(n+1)/2 - W_A$ to find the critical value.

From Table 9.13 on page 430, we see that the sample size is 10. The critical value for a left-tailed test at the 5% significance level is $W_{1-0.05}$. To find the critical value, first we go down the outside columns of Table V, labeled n, to "10." Then, going across that row to the column labeled $W_{0.05}$, we reach 44; thus $W_{0.05} = 44$. Now we apply the aforementioned relation and the result just obtained to get

$$W_{1-0.05} = 10(10+1)/2 - W_{0.05} = 55 - 44 = 11,$$

which is the required critical value. See Fig. 9.27A.

FIGURE 9.27A



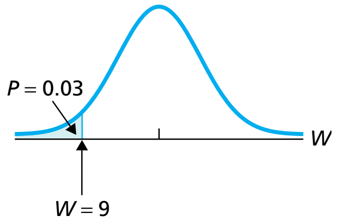
Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

The value of the test statistic is $W = 9$, as found in Step 3, which falls in the rejection region shown in Fig. 9.27A. Thus we reject H_0. The test results are statistically significant at the 5% level.

Step 4 Obtain the *P*-value by using technology.

Using technology, we find that the P-value for the hypothesis test is $P = 0.03$, as shown in Fig. 9.27B.

FIGURE 9.27B



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

From Step 4, $P = 0.03$. Because the P-value is less than the specified significance level of 0.05, we reject H_0. The test results are statistically significant at the 5% level and (see Table 9.8 on page 408) provide strong evidence against the null hypothesis.

Step 6 Interpret the results of the hypothesis test.

StatCrunch Report 9.3

Interpretation At the 5% significance level, the data provide sufficient evidence to conclude that the mean weekly food cost for Kansas families of three is less than the national mean of \$157.

As mentioned earlier, one advantage of nonparametric methods is that they are resistant to outliers. We can illustrate that advantage for the Wilcoxon signed-rank test by referring to Example 9.20.

The stem-and-leaf diagram depicted in Fig. 9.23 on page 430 shows that the sample data presented in Table 9.13 contain no outliers. The smallest observation, and also the farthest from the null hypothesis mean of 157, is 135. Replacing 135 by, say, 85, introduces an outlier but has no effect on the value of the test statistic and hence none on the hypothesis test itself. (Why is that so?)

Note: The following points may be relevant when performing a Wilcoxon signed-rank test:

- If an observation equals μ_0 (the value for the mean in the null hypothesis), that observation should be removed and the sample size reduced by 1.
- If two or more absolute differences are tied, each should be assigned the mean of the ranks they would have had if there were no ties.

To illustrate the second bulleted item, suppose that two absolute differences are tied for second place. Then each should be assigned rank $(2+3)/2 = 2.5$, and rank 4 should be assigned to the next-largest absolute difference, which really is fourth. Similarly, if three absolute differences are tied for fifth place, each should be assigned rank $(5+6+7)/3 = 6$, and rank 8 should be assigned to the next-largest absolute difference.

In Example 9.16, we used the one-mean t-test to decide whether, on average, high mountain lakes in the Southern Alps are nonacidic. Now we do so by using the Wilcoxon signed-rank test.

EXAMPLE 9.21 The Wilcoxon Signed-Rank Test

Acid Rain and Lake Acidity A lake is classified as nonacidic if it has a pH greater than 6. A. Marchetto and A. Lami measured the pH of high mountain lakes in the Southern Alps and reported their findings in the paper "Reconstruction of pH by Chrysophycean Scales in Some Lakes of the Southern Alps" (*Hydrobiologia*, Vol. 274, pp 83–90). Table 9.12, which we repeat here as Table 9.15, shows the pH levels obtained by the researchers for 15 lakes.

TABLE 9.15
pH levels for 15 lakes

7.2	7.3	6.1	6.9	6.6
7.3	6.3	5.5	6.3	6.5
5.7	6.9	6.7	7.9	5.8

At the 5% significance level, do the data provide sufficient evidence to conclude that, on average, high mountain lakes in the Southern Alps are nonacidic? Use the Wilcoxon signed-rank test.

FIGURE 9.28
Stem-and-leaf diagram of pH levels in Table 9.15

```
5 | 5 7 8
6 | 1 3 3
6 | 5 6 7 9 9
7 | 2 3 3
7 | 9
```

Solution Figure 9.28 shows a stem-and-leaf diagram of the sample data in Table 9.15. The diagram is relatively symmetric. Hence, we can reasonably apply Procedure 9.3 to carry out the required hypothesis test.

Step 1 State the null and alternative hypotheses.

Let μ denote the mean pH level of all high mountain lakes in the Southern Alps. Then the null and alternative hypotheses are, respectively,

$$H_0\text{: } \mu = 6 \text{ (on average, the lakes are acidic)}$$

$$H_a\text{: } \mu > 6 \text{ (on average, the lakes are nonacidic).}$$

Note that the hypothesis test is right tailed.

Step 2 Decide on the significance level, α.

We are to perform the test at the 5% significance level, so $\alpha = 0.05$.

Step 3 Compute the value of the test statistic

$$\mathbf{W = \text{sum of the positive ranks.}}$$

To do so, first construct a worktable to obtain the signed ranks.

We construct the following work table. Note that, in several instances, we applied the aforementioned method to deal with tied absolute differences.

pH x	Difference $D = x - 6$	$\lvert D \rvert$	Rank of $\lvert D \rvert$	Signed rank R
7.2	1.2	1.2	12	12
7.3	1.3	1.3	13.5	13.5
6.1	0.1	0.1	1	1
6.9	0.9	0.9	10.5	10.5
6.6	0.6	0.6	8	8
7.3	1.3	1.3	13.5	13.5
6.3	0.3	0.3	4	4
5.5	−0.5	0.5	6.5	−6.5
6.3	0.3	0.3	4	4
6.5	0.5	0.5	6.5	6.5
5.7	−0.3	0.3	4	−4
6.9	0.9	0.9	10.5	10.5
6.7	0.7	0.7	9	9
7.9	1.9	1.9	15	15
5.8	−0.2	0.2	2	−2

Referring to the last column of the work table, we find that the value of the test statistic is

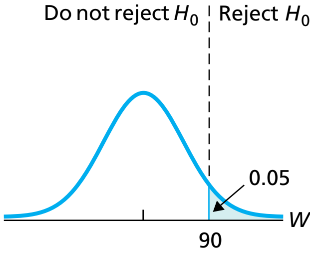
$$W = 12 + 13.5 + 1 + \cdots + 9 + 15 = 107.5.$$

CRITICAL-VALUE APPROACH

Step 4 The critical value for a right-tailed test is W_α. Use Table V to find the critical value.

From Table 9.15, we see that the sample size is 15. The critical value for a right-tailed test at the 5% significance level is $W_{0.05}$. To find the critical value, first we go down the outside columns of Table V, labeled n, to “15.” Then, going across that row to the column labeled $W_{0.05}$, we reach 90, the required critical value. See Fig. 9.29A.

FIGURE 9.29A



Step 5 If the value of the test statistic falls in the rejection region, reject H_0; otherwise, do not reject H_0.

The value of the test statistic is $W = 107.5$, as found in Step 3, which falls in the rejection region shown in Fig. 9.29A. Thus we reject H_0. The test results are statistically significant at the 5% level.

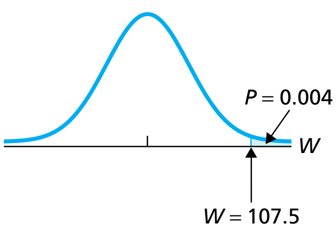
OR

P-VALUE APPROACH

Step 4 Obtain the P-value by using technology.

Using technology, we find that the P-value for the hypothesis test is $P = 0.004$, as shown in Fig. 9.29B.

FIGURE 9.29B



Step 5 If $P \leq \alpha$, reject H_0; otherwise, do not reject H_0.

From Step 4, $P = 0.004$. Because the P-value is less than the specified significance level of 0.05, we reject H_0. The test results are statistically significant at the 5% level and (see Table 9.8 on page 408) provide very strong evidence against the null hypothesis.

You try it!

Exercise 9.149 on page 441

Step 6 Interpret the results of the hypothesis test.

Interpretation At the 5% significance level, the data provide sufficient evidence to conclude that, on average, high mountain lakes in the Southern Alps are nonacidic.

We note that both the one-mean t-test of Example 9.16 and the Wilcoxon signed-rank test of Example 9.21 reject the null hypothesis that high mountain lakes in the Southern Alps are, on average, acidic in favor of the alternative hypothesis that they are, on average, nonacidic. Furthermore, with both tests, the data provide very strong evidence against that null hypothesis (and, hence, in favor of the alternative hypothesis). Indeed, as we have seen, $P = 0.002$ for the one-mean t-test, and $P = 0.004$ for the Wilcoxon signed-rank test.

Comparing the Wilcoxon Signed-Rank Test and the *t*-Test

As you learned in Section 9.5, a t-test can be used to conduct a hypothesis test for a population mean when the variable under consideration is normally distributed. Because normally distributed variables have symmetric distributions, we can also use the Wilcoxon signed-rank test to perform such a hypothesis test.

For a normally distributed variable, the t-test is more powerful than the Wilcoxon signed-rank test because it is designed expressly for such variables; surprisingly, though, the t-test is not much more powerful than the Wilcoxon signed-rank test. However, if the variable under consideration has a symmetric distribution but is not normally distributed, the Wilcoxon signed-rank test is usually more powerful than the t-test and is often considerably more powerful.

KEY FACT 9.8

Wilcoxon Signed-Rank Test Versus the *t*-Test

Suppose that you want to perform a hypothesis test for a population mean. When deciding between the t-test and the Wilcoxon signed-rank test, follow these guidelines:

- If you are reasonably sure that the variable under consideration is normally distributed, use the t-test.
- If you are not reasonably sure that the variable under consideration is normally distributed but are reasonably sure that it has a symmetric distribution, use the Wilcoxon signed-rank test.

Testing a Population Median with the Wilcoxon Signed-Rank Procedure

Because the mean and median of a symmetric distribution are identical, a Wilcoxon signed-rank test can be used to perform a hypothesis test for a population median, η, as well as for a population mean, μ. To use Procedure 9.3 to carry out a hypothesis test for a population median, simply replace μ by η and μ_0 by η_0.

THE TECHNOLOGY CENTER

Some statistical technologies have programs that automatically perform a Wilcoxon signed-rank test, but others do not. In this subsection, we present output and step-by-step instructions for such programs.

As you will see, different programs may report slightly different P-values for a Wilcoxon signed-rank test. These differences are due to the fact that different programs may use different methods for obtaining or approximating such P-values.

Note to Excel users: The Excel program that we use to perform a (one-sample) Wilcoxon signed-rank test is actually designed for a two-sample test. Nonetheless, as we show, it is possible to use that program to perform a (one-sample) Wilcoxon signed-rank test.

Note to TI-83/84 Plus users: At the time of this writing, the TI-83/84 Plus does not have a built-in program for conducting a Wilcoxon signed-rank test. However, we have written a TI program called WILCOX for performing that test. It is located in the TI Programs section on the WeissStats site. Your instructor can show you how to download the program to your calculator. *Warning:* Any data that you may have previously stored in Lists 1–6 will be erased during program execution, so copy those data to other lists prior to program execution if you want to retain them.

As we said earlier, a Wilcoxon signed-rank test can be used to perform a hypothesis test for a population median, η, as well as for a population mean, μ. Many statistical technologies present the output of that procedure in terms of the median, but that output can also be interpreted in terms of the mean.

EXAMPLE 9.22 Using Technology to Conduct a Wilcoxon Signed-Rank Test

Weekly Food Costs Table 9.13 on page 430 gives the weekly food costs for 10 Kansas families of three. Use Minitab, Excel, or the TI-83/84 Plus to decide, at the 5% significance level, whether the data provide sufficient evidence to conclude that the mean weekly food cost for Kansas families of three is less than the national mean of \$157.

Solution Let μ denote the mean weekly food cost for all Kansas families of three. We want to perform the hypothesis test

$$H_0\text{: } \mu = \$157 \text{ (mean weekly food cost is not less than \$157)}$$
$$H_a\text{: } \mu < \$157 \text{ (mean weekly food cost is less than \$157)}$$

at the 5% significance level. Note that the hypothesis test is left tailed.

We applied the Wilcoxon signed-rank test programs to the data, resulting in Output 9.3. Steps for generating that output are presented in Instructions 9.3. *Note to Excel users:* For brevity, we have presented only the essential portions of the actual output.

OUTPUT 9.3 Wilcoxon signed-rank test output on the sample of weekly food costs

MINITAB

Wilcoxon Signed Rank Test: COST

```
Test of median = 157.0 versus median < 157.0

            N for   Wilcoxon               Estimated
       N    Test    Statistic         P       Median
COST  10      10          9.0     0.033        149.5
```

TI-83/84 PLUS

```
NORMAL FLOAT AUTO REAL RADIAN MP
N
                   10
N FOR TEST
                   10
W
                    9
P
          .0332728013
                 Done
```

EXCEL

Hypothesized difference (D): 0
Significance level (%): 5
p-value: Exact p-value

Summary statistics:

Variable	Observations	Mean	Std. deviation
COST	10	149.7000	10.8837

Wilcoxon signed-rank test / Lower-tailed test:

V	9
Expected value	27.5000
Variance (V)	96.2500
p-value (one-tailed)	0.0322
alpha	0.05

As shown in Output 9.3, the *P*-value for the hypothesis test is 0.03. Because the *P*-value is less than the specified significance level of 0.05, we reject H_0. At the 5% significance level, the data provide sufficient evidence to conclude that the mean weekly food cost for Kansas families of three is less than the national mean of \$157.

INSTRUCTIONS 9.3 Steps for generating Output 9.3

MINITAB

1 Store the data from Table 9.13 in a column named COST
2 Choose **Stat ➤ Nonparametrics ➤ 1-Sample Wilcoxon. . .**
3 Press the F3 key to reset the dialog box
4 Specify COST in the **Variables** text box
5 Select the **Test median** option button
6 Click in the **Test median** text box and type 157
7 Click the arrow button at the right of the **Alternative** drop-down list box and select **less than**
8 Click **OK**

EXCEL

1 Store the data from Table 9.13 in a column named COST
2 Store the null hypothesis mean, 157, repeated 10 times (the sample size), in a column named MU_0
3 Choose **XLSTAT ➤ Nonparametric tests ➤ Comparison of two samples (Wilcoxon, Mann-Whitney, . . .)**
4 Click the reset button in the lower left corner of the dialog box
5 Click in the **Sample 1** selection box and then select the column of the worksheet that contains the COST data
6 Click in the **Sample 2** selection box and then select the column of the worksheet that contains the MU_0 data
7 Uncheck the **Sign test** check box
8 Click the **Options** tab
9 Click the arrow button at the right of the **Alternative hypothesis** drop-down list box and select **Sample 1 – Sample 2 < D**
10 Type 5 in the **Significance level (%)** text box
11 Check the **Exact p-value** check box
12 Click **OK**
13 Click the **Continue** button in the **XLSTAT – Selections** dialog box

TI-83/84 PLUS

1 Store the data from Table 9.13 in a list named COST
2 Press **PRGM**
3 Arrow down to WILCOX and press **ENTER** twice
4 Press **2ND ➤ LIST**, arrow down to COST, and press **ENTER** twice
5 Type 157 for **MU0** and press **ENTER**
6 Type -1 for **TYPE** and press **ENTER**

Exercises 9.6

Understanding the Concepts and Skills

9.129 Technically, what is a *nonparametric method?* In current statistical practice, how is that term used?

9.130 What distributional assumption must be met in order to use the Wilcoxon signed-rank test?

9.131 We mentioned that if, in a Wilcoxon signed-rank test, an observation equals μ_0 (the value given for the mean in the null hypothesis), that observation should be removed and the sample size reduced by 1. Why does that need to be done?

In each of Exercises ***9.132–9.137****, suppose that you want to perform a hypothesis test for a population mean. Assume that the population standard deviation is unknown and that the sample size is relatively small. In each exercise, we have given the distribution shape of the variable under consideration. Decide whether you would use the t-test, the Wilcoxon signed-rank test, or neither. Explain your answers.*

9.132 Uniform

9.133 Normal

9.134 Reverse J shaped

9.135 Triangular

9.136 Symmetric bimodal

9.137 Left skewed

9.138 The Wilcoxon signed-rank test can be used to perform a hypothesis test for a population median, η, as well as for a population mean, μ. Why is that so?

Exercises ***9.139–9.142*** *pertain to critical values for a Wilcoxon signed-rank test. Use Table V in Appendix A to determine the critical value(s) in each case. For a left-tailed or two-tailed test, you will also need the relation* $W_{1-A} = n(n+1)/2 - W_A$.

9.139 Sample size = 8; Significance level = 0.05
a. Right tailed **b.** Left tailed **c.** Two tailed

9.140 Sample size = 10; Significance level = 0.01
a. Right tailed **b.** Left tailed **c.** Two tailed

9.141 Sample size = 19; Significance level = 0.10
a. Right tailed **b.** Left tailed **c.** Two tailed

9.142 Sample size = 15; Significance level = 0.05
a. Right tailed **b.** Left tailed **c.** Two tailed

In each of Exercises ***9.143–9.148****, we have provided a null hypothesis and alternative hypothesis and a sample from the population under consideration. In each case, use the Wilcoxon signed-rank test to perform the required hypothesis test at the 10% significance level.*

9.143 H_0: $\mu = 5$, H_a: $\mu > 5$

12	7	11	9	3	2	8	6

9.144 H_0: $\mu = 10$, H_a: $\mu < 10$

7	6	5	12	15	14	13	4

9.145 H_0: $\mu = 6$, H_a: $\mu \neq 6$

6	4	8	4	1	1	4	7

9.146 H_0: $\mu = 3$, H_a: $\mu \neq 3$

6	6	3	3	2	5	4	7	4

9.147 H_0: $\mu = 12$, H_a: $\mu < 12$

16	11	10	14	13	15	5	8	11

9.148 H_0: $\mu = 8$, H_a: $\mu > 8$

8	10	11	11	5	9	9	12

Applying the Concepts and Skills

*In each of Exercises **9.149–9.154**, use the Wilcoxon signed-rank test to perform the required hypothesis test.*

9.149 Global Warming? During the late 1800s, Lake Wingra in Madison, Wisconsin, was frozen over an average of 124.9 days per year. A random sample of eight recent years provided the following data on numbers of days that the lake was frozen over.

103	80	79	135	134	77	80	111

At the 5% significance level, do the data provide sufficient evidence to conclude that the average number of ice days is less now than in the late 1800s?

9.150 Happy-Life Years. In the article, "Apparent Quality-of-Life in Nations: How Long and Happy People Live" (*Social Indicators Research*, Vol. 71, pp. 61–86) R. Veenhoven discussed how the quality of life in nations can be measured by how long and happy people live. In the 1990s, the median number of happy-life years across nations was 46.7. A random sample of eight nations for this year provided the following data on number of happy-life years.

30.3	47.0	56.4	30.5	39.6	47.9	29.7	52.5

At the 5% significance level, do the data provide sufficient evidence to conclude that the median number of happy-life years has changed from that in the 1990s?

9.151 How Old People Are. In 2010, the median age of U.S. residents was 37.2 years, as reported by the U.S. Census Bureau in *Current Population Reports*. A random sample of 10 U.S. residents taken this year yielded the following ages, in years.

44	64	16	59	38
47	51	41	13	28

At the 1% significance level, do the data provide sufficient evidence to conclude that the median age of today's U.S. residents has increased from the 2010 median age of 37.2 years?

9.152 Beverage Expenditures. The Bureau of Labor Statistics publishes information on average annual expenditures by consumers in *Consumer Expenditures*. In 2012, the mean amount spent per consumer unit on nonalcoholic beverages was \$370. A random sample of 12 consumer units yielded the following data, in dollars, on last year's expenditures on nonalcoholic beverages.

511	326	334	415	409	431
390	423	409	399	344	408

At the 5% significance level, do the data provide sufficient evidence to conclude that last year's mean amount spent by consumers on nonalcoholic beverages has increased from the 2012 mean of \$370?

9.153 Pricing Mustangs. According to the *Kelley Blue Book*, the fair purchase price from dealers for a 2-year-old Ford Mustang coupe is about \$18,000. A random sample of 10 purchase prices from private parties yielded the following data, in dollars.

16,594	16,106	16,102	15,914	15,713
15,613	14,614	13,514	15,614	15,714

At the 1% significance level, do the data provide sufficient evidence to conclude that the mean purchase price from private parties for 2-year-old Ford Mustang coupes is less than the fair purchase price from dealers?

9.154 Birth Weights. According to the article "Baby Birth Weight Statistics" by V. Iannelli, which appears on the About.com Pediatrics website, in 2005, the median birth weight of U.S. babies was 3389 g, or about 7.5 lb. A random sample of this year's births provided the following weights, in pounds.

8.7	7.5	5.4	13.9	7.9	5.8	9.3
8.9	8.3	9.3	5.7	6.1	11.7	7.3

Can we conclude that this year's median birth weight differs from that in 2005? Use a significance level of 0.05.

9.155 Death Rolls. Alligators perform a spinning maneuver, referred to as a "death roll", to subdue their prey. Videos were taken of juvenile alligators performing this maneuver in a study for the article "Death Roll of the Alligator, Mechanics of Twist and Feeding in Water" (*Journal of Experimental Biology*, Vol. 210, pp. 2811–2818) by F. Fish et al. One of the variables measured was the degree of the angle between the body and head of the alligator while performing the roll. A sample of 20 rolls yielded the following data, in degrees.

58.6	58.7	57.3	54.5	52.9
59.5	29.4	43.4	31.8	52.3
42.7	34.8	39.2	61.3	60.4
51.5	42.8	57.5	43.6	47.6

a. Do the data provide sufficient evidence to conclude that, on average, the angle between the body and head of an alligator during a death roll is greater than 45°? Perform a Wilcoxon signed-rank test at the 5% significance level.

b. The hypothesis test considered in part (a) was done in Exercise 9.115 with a t-test. The assumption in that exercise is that the angle between the body and head of an alligator during a death roll is (approximately) normally distributed. If that is the case, why is it permissible to perform a Wilcoxon signed-rank test for the mean angle between the body and head of an alligator during a death roll?

9.156 Ethical Food Choice Motives. In the paper "Measurement of Ethical Food Choice Motives" (*Appetite*, Vol. 34, pp. 55–59), research psychologists M. Lindeman and M. Väänänen of the University of Helsinki published a study on the factors that most influence peoples' choice of food. One of the questions asked of the participants was how important, on a scale of 1 to 4 (1 = not at all important, 4 = very important), is ecological welfare in food choice motive, where ecological welfare includes animal welfare and environmental protection. Following are the responses of a random sample of 18 Helsinkians.

3	2	2	3	3	3	2	2	3
3	3	1	3	4	2	1	3	1

At the 5% significance level, do the data provide sufficient evidence to conclude that, on average, Helsinkians respond with an ecological welfare food choice motive greater than 2?

a. Use the Wilcoxon signed-rank test.
b. Use the t-test.
c. Compare the results of the two tests.

9.157 Checking Advertised Contents. A manufacturer of liquid detergent produces a bottle with an advertised content of 450 millilitres (mL). Fourteen bottles are randomly selected and found to have the following contents, in mL.

447	459	439	443	462	449	437
458	453	461	445	467	456	448

A normal probability plot of the data indicates that you can assume the contents are normally distributed. Let μ denote the mean content of all bottles produced. To decide whether the mean content is less than advertised, perform the hypothesis test

$$H_0\text{: } \mu = 450 \text{ mL}$$
$$H_a\text{: } \mu < 450 \text{ mL}$$

at the 10% significance level.

a. Use the t-test.
b. Use the Wilcoxon signed-rank test.
c. If the mean content is in fact less than 450 mL, how do you explain the discrepancy between the two tests?

9.158 Education of Jail Inmates. Thirty years ago, a country's national justice statistics reported that the median educational attainment of its jail inmates was 10.2 years. Ten current inmates are randomly selected and found to have the following educational attainments, in years.

7	6	10	9	7
10	14	9	10	8

Assume that educational attainments of current jail inmates have a symmetric, non-normal distribution. At the 5% significance level, do the data provide sufficient evidence to conclude that this year's median educational attainment has changed from what it was 30 years ago?

a. Use the t-test.
b. Use the Wilcoxon signed-rank test.
c. If this year's median educational attainment has in fact changed from what it was 30 years ago, how do you explain the discrepancy between the two tests?

*In each of Exercises **9.159** and **9.160**, use the technology of your choice to decide whether applying the Wilcoxon signed-rank test is reasonable. Explain your answers.*

9.159 Head Injury Criterion. The Head Injury Criterion (HIC) is a measure of the likelihood of an injury arising from an accident such as a vehicle crash. At an HIC of 1000, one in six people will suffer a life-threatening injury to the brain. The Insurance Institute for Highway Safety performs safety rating tests on vehicles. One of the variables measured is the HIC. The following data provide the HIC levels for a sample of small SUVs.

493	127	95	101	283	82
147	358	81	158	102	196

9.160 Asian Elephants. In the paper "A Survey of African and Asian Elephant Diets and Measured Body Dimensions Compared to Their Estimated Nutrient Requirements" (*Proceedings of the American Zoo and Aquarium Association Nutrition Advisory Group*, 4:13–27), K. Ange et al. studied nutrient levels of African and Asian elephants at a sample of zoos in the United States and Europe. The following table gives the number of Asian elephants at each of the fifteen zoos sampled.

0	2	4	0	3	2	0	0
0	1	0	7	8	2	0	

Working with Large Data Sets

9.161 Delaying Adulthood. The convict surgeonfish is a common tropical reef fish that has been found to delay metamorphosis into adult by extending its larval phase. This delay often leads to enhanced survivorship in the species by increasing the chances of finding suitable habitat. In the paper "Delayed Metamorphosis of a Tropical Reef Fish (*Acanthurus triostegus*): A Field Experiment" (*Marine Ecology Progress Series*, Vol. 176, pp. 25–38), M. McCormick published data that he obtained on the larval duration, in days, of 90 convict surgeonfish. The data are given on the WeissStats site. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean larval duration of convict surgeonfish exceeds 52 days?

a. Employ the Wilcoxon signed-rank test.
b. Employ the t-test.
c. Compare your results from parts (a) and (b).

9.162 Easy Hole at the British Open? The Old Course at St. Andrews in Scotland is home of the British Open, one of the major tournaments in professional golf. The *Hole O'Cross Out*, known by both European and American professional golfers as one of the friendliest holes at St. Andrews, is the fifth hole, a 514-yard, par 5 hole with an open fairway and a large green. As one reporter put it, "If players think before they drive, they will easily walk away with birdies and pars." The scores on the *Hole O'Cross Out* posted by a sample of 156 golf professionals are presented on the WeissStats site. Use those data and the technology of your choice to decide whether, on average, professional golfers score better than par on the *Hole O'Cross Out*. Perform the required hypothesis test at the 0.01 level of significance.

a. Employ the Wilcoxon signed-rank test.
b. Employ the t-test.
c. Compare your results from parts (a) and (b).

*In Exercises **9.163–9.165**, we have repeated the contexts of Exercises 9.123–9.125 from Section 9.5. For each exercise, use the technology of your choice to do the following.*

a. *Apply the Wilcoxon signed-rank test to perform the required hypothesis test.*
b. *Compare your result in part (a) to that obtained in the corresponding exercise in Section 9.5, where the t-test was used.*

9.163 Stressed-Out Bus Drivers. In the paper "Hassles on the Job: A Study of a Job Intervention With Urban Bus Drivers" (*Journal of Organizational Behavior*, Vol. 20, pp. 199–208), G. Evans et al.

examined the effects of an intervention program to improve the conditions of urban bus drivers. Among other variables, the researchers monitored diastolic blood pressure of bus drivers in downtown Stockholm, Sweden. The data, in millimeters of mercury (mm Hg), on the WeissStats site are based on the blood pressures obtained prior to intervention for the 41 bus drivers in the study. At the 10% significance level, do the data provide sufficient evidence to conclude that the mean diastolic blood pressure of bus drivers in Stockholm exceeds the normal diastolic blood pressure of 80 mm Hg?

9.164 How Far People Drive. In 2011, the average car in the United States was driven 13.5 thousand miles, as reported by the Federal Highway Administration in *Highway Statistics*. On the WeissStats site, we provide last year's distance driven, in thousands of miles, by each of 500 randomly selected cars. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean distance driven last year differs from that in 2011?

9.165 Fair Market Rent. According to the document *Out of Reach*, published by the National Low Income Housing Coalition, the fair market rent (FMR) for a two-bedroom unit in the United States is \$949. A sample of 100 randomly selected two-bedroom units yielded the data on monthly rents, in dollars, given on the WeissStats site. At the 5% significance level, do the data provide sufficient evidence to conclude that the mean monthly rent for two-bedroom units is greater than the FMR of \$949? Perform the required hypothesis test both with and without the outlier.

Extending the Concepts and Skills

Normal Approximation for *W*. The Wilcoxon signed-rank table, Table V, stops at $n = 20$. For larger samples, a normal approximation can be used. In fact, the normal approximation works well even for sample sizes as small as 10.

Normal Approximation for *W*

Suppose that the variable under consideration has a symmetric distribution. Then, for samples of size n,

- $\mu_W = n(n+1)/4$,
- $\sigma_W = \sqrt{n(n+1)(2n+1)/24}$, and
- W is approximately normally distributed for $n \geq 10$.

Thus, for samples of size 10 or more, the standardized variable

$$z = \frac{W - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}$$

has approximately the standard normal distribution.

9.166 Large-Sample Wilcoxon Signed-Rank Test. Formulate a hypothesis-testing procedure for a Wilcoxon signed-rank test that uses the test statistic z given in the preceding box. *Note:* Using a continuity correction provides even more accurate results.

9.167 Birth Weights. Refer to Exercise 9.154.

a. Use the procedure you formulated in Exercise 9.166 to perform the hypothesis test in Exercise 9.154.

b. Compare your result in part (a) to the one you obtained in Exercise 9.154, where the normal approximation was not used.

9.168 The Distribution of *W*. In this exercise, you are to obtain the distribution of the variable W for samples of size 3 so that you can see how the Wilcoxon signed-rank table is constructed.

a. The rows of the following table give all possible signs for the signed ranks in a Wilcoxon signed-rank test with $n = 3$. For instance, the first row covers the possibility that all three observations are greater than μ_0 and thus have positive sign ranks. Fill in the empty column with values of W. (*Hint:* The first entry is 6, and the last is 0.)

Rank			
1	2	3	***W***
+	+	+	
+	+	−	
+	−	+	
+	−	−	
−	+	+	
−	+	−	
−	−	+	
−	−	−	

b. If the null hypothesis H_0: $\mu = \mu_0$ is true, what percentages of samples will match any particular row of the table? (*Hint:* The answer is the same for all rows.)

c. Use the answer from part (b) to obtain the distribution of W for samples of size 3.

d. Draw a relative-frequency histogram of the distribution obtained in part (c).

e. Use your histogram from part (d) to find $W_{0.125}$ for a sample size of 3.

One-Median Sign Test. Recall that the Wilcoxon signed-rank test, which can be used to perform a hypothesis test for a population median, η, requires that the variable under consideration has a symmetric distribution. If that is not the case, the **one-median sign test** (or simply the **sign test**) can be used instead. The one-median sign test is also known as the **one-sample sign test** and the **one-variable sign test.** Technically, like the Wilcoxon signed-rank test, use of the sign test requires that the variable under consideration has a continuous distribution. In practice, however, that restriction is usually ignored.

If the null hypothesis H_0: $\eta = \eta_0$ is true, the probability is 0.5 of an observation exceeding η_0. Therefore, in a simple random sample of size n, the number of observations, s, that exceed η_0 has a binomial distribution with parameters n and 0.5.

To perform a sign test, first assign a "+" sign to each observation in the sample that exceeds η_0 and then obtain the number of "+" signs, which we denote s_0. The P-value for the hypothesis test can be found by applying Exercise 9.71 on page 409 and obtaining the required binomial probability.

9.169 Assuming that the null hypothesis H_0: $\eta = \eta_0$ is true, answer the following questions.

a. Why is the probability of an observation exceeding η_0 equal to 0.5?

b. In a simple random sample of size n, why does the number of observations that exceed η_0 have a binomial distribution with parameters n and 0.5?

9.170 The sign test can be used whether or not the variable under consideration has a symmetric distribution. If the distribution is in fact symmetric, the Wilcoxon signed-rank test is preferable. Why do you think that is so?

9.171 What advantage does the sign test have over the Wilcoxon signed-rank test?

9.172 Explain how to proceed with a sign test if one or more of the observations equals η_0, the value specified in the null hypothesis for the population median.