11 CHAPTER

NONPARAMETRIC TESTS

In a recent year, the most common form of reported identity theft was credit card fraud (20%), followed by government documents/benefits fraud (15%), employment fraud (15%), and phone or utilities fraud (13%).

WHERE YOU'VE BEEN

Up to this point in the text, you have studied dozens of different statistical formulas and tests that can help you in a decision-making process. Specific conditions had to be satisfied in order to use these formulas and tests.

Suppose it is believed that as the number of fraud complaints in a state increases, the number of identity theft victims also increases. Can this belief be supported by actual data? The data below show the number of fraud complaints and the number of identify theft victims for 25 randomly selected states in a recent year.

Fraud complaints	29,506	22,805	1535	10,556	8099	106,623	2630	8978
Identity theft victims	8237	8363	296	3292	2696	51,140	759	3819

Fraud complaints	5895	57,472	19,585	15,159	2253	12,584	4807	7101
Identity theft victims	1347	24,440	5412	4589	490	2937	2081	2005

Fraud complaints	8173	24,695	7345	15,515	21,730	20,610	13,259	4498	30,578
Identity theft victims	2396	6349	1775	5408	5855	9683	3528	2367	13,726

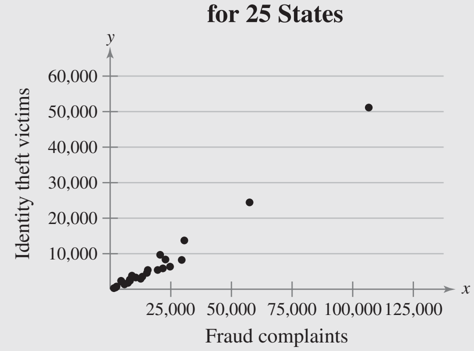
WHERE YOU'RE GOING

In this chapter, you will study additional statistical tests that do not require the population distribution to meet any specific conditions. Each of these tests has usefulness in real-life applications.

With the data above, the number of fraud complaints F and the number of identity theft victims V can be related by the regression equation $V = 0.472F - 1802.101$. The correlation coefficient is approximately 0.987, so there is a strong positive correlation. You can determine that the correlation is significant by using Table 11 in Appendix B, but the V-values do not pass the normality requirement.

So, although a simple correlation test might indicate a relationship between the number of fraud complaints and the number of identity theft victims, one might question the results because the data do not fit the requirements for the test. Similar tests you will study in this chapter, such as Spearman's rank correlation test, will give you additional information. The Spearman's rank correlation coefficient for this data is approximately 0.971. At $\alpha = 0.01$, there is in fact a significant correlation between the number of fraud complaints and the number of identity theft victims for each state.

Number of Fraud Complaints and Identity Theft Victims for 25 States



11.1 The Sign Test

WHAT YOU SHOULD LEARN

- How to use the sign test to test a population median
- How to use the paired-sample sign test to test the difference between two population medians (dependent samples)

The Sign Test for a Population Median ▸ The Paired-Sample Sign Test

▸ THE SIGN TEST FOR A POPULATION MEDIAN

Many of the hypothesis tests studied so far have imposed one or more requirements for a population distribution. For instance, some tests require that a population must have a normal distribution, and other tests require that population variances be equal. What if, for a given test, such requirements cannot be met? For these cases, statisticians have developed hypothesis tests that are "distribution free." Such tests are called *nonparametric tests*.

DEFINITION

A **nonparametric test** is a hypothesis test that does not require any specific conditions concerning the shapes of population distributions or the values of population parameters.

Nonparametric tests are usually easier to perform than corresponding parametric tests. However, they are usually less efficient than parametric tests. Stronger evidence is required to reject a null hypothesis using the results of a nonparametric test. Consequently, whenever possible, you should use a parametric test. One of the easiest nonparametric tests to perform is the *sign test*.

INSIGHT

For many nonparametric tests, statisticians test the median instead of the mean.

DEFINITION

The **sign test** is a nonparametric test that can be used to test a population median against a hypothesized value k.

The sign test for a population median can be left-tailed, right-tailed, or two-tailed. The null and alternative hypotheses for each type of test are as follows.

Left-tailed test:
H_0: median $\geq k$ and H_a: median $< k$

Right-tailed test:
H_0: median $\leq k$ and H_a: median $> k$

Two-tailed test:
H_0: median $= k$ and H_a: median $\neq k$

To use the sign test, first compare each entry in the sample with the hypothesized median k. If the entry is below the median, assign it a $-$ sign; if the entry is above the median, assign it a $+$ sign; and if the entry is equal to the median, assign it a 0. Then compare the number of $+$ and $-$ signs. (The 0's are ignored.) If there is a large difference between the number of $+$ signs and the number of $-$ signs, it is likely that the median is different from the hypothesized value and the null hypothesis should be rejected.

Table 8 in Appendix B lists the critical values for the sign test for selected levels of significance and sample sizes. When the sign test is used, the sample size n is the total number of + and − signs. If the sample size is greater than 25, you can use the standard normal distribution to find the critical values.

INSIGHT

Because the 0's are ignored, there are two possible outcomes when comparing a data entry with a hypothesized median: a + or a − sign. If the median is k, then about half of the values will be above k and half will be below. As such, the probability for each sign is 0.5. Table 8 in Appendix B is constructed using the binomial distribution where $p = 0.5$.

When $n > 25$, you can use the normal approximation (with a continuity correction) for the binomial. In this case, use $\mu = np = 0.5n$ and $\sigma = \sqrt{npq} = \dfrac{\sqrt{n}}{2}$.

TEST STATISTIC FOR THE SIGN TEST

When $n \leq 25$, the **test statistic** for the sign test is x, the smaller number of + or − signs.

When $n > 25$, the **test statistic** for the sign test is

$$z = \frac{(x + 0.5) - 0.5n}{\dfrac{\sqrt{n}}{2}}$$

where x is the smaller number of + or − signs and n is the sample size, i.e., the total number of + and − signs.

Because x is defined to be the smaller number of + or − signs, the rejection region is always in the left tail. Consequently, the sign test for a population median is always a left-tailed test or a two-tailed test. When the test is two-tailed, use only the left-tailed critical value. (If x is defined to be the larger number of + or − signs, the rejection region is always in the right tail. Right-tailed sign tests are presented in the exercises.)

GUIDELINES

Performing a Sign Test for a Population Median

IN WORDS	IN SYMBOLS
1. Identify the claim. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Determine the sample size n by assigning + signs and − signs to the sample data.	n = total number of + and − signs
4. Determine the critical value.	If $n \leq 25$, use Table 8 in Appendix B. If $n > 25$, use Table 4 in Appendix B.
5. Find the test statistic.	If $n \leq 25$, use x. If $n > 25$, use $z = \dfrac{(x + 0.5) - 0.5n}{\dfrac{\sqrt{n}}{2}}$.
6. Make a decision to reject or fail to reject the null hypothesis.	If the test statistic is less than or equal to the critical value, reject H_0. Otherwise, fail to reject H_0.
7. Interpret the decision in the context of the original claim.	

EXAMPLE 1

SC Report 50

Using the Sign Test

A website administrator for a company claims that the median number of visitors per day to the company's website is no more than 1500. An employee doubts the accuracy of this claim. The number of visitors per day for 20 randomly selected days are listed below. At $\alpha = 0.05$, can the employee reject the administrator's claim?

1469	1462	1634	1602	1500
1463	1476	1570	1544	1452
1487	1523	1525	1548	1511
1579	1620	1568	1492	1649

Solution

The claim is "the median number of visitors per day to the company's website is no more than 1500." So, the null and alternative hypotheses are

H_0: median ≤ 1500 (Claim) and H_a: median > 1500.

The results of comparing each data entry with the hypothesized median 1500 are shown.

−	−	+	+	0
−	−	+	+	−
−	+	+	+	+
+	+	+	−	+

You can see that there are 7 − signs and 12 + signs. So, $n = 12 + 7 = 19$. Because $n \leq 25$, use Table 8 to find the critical value. The test is a one-tailed test with $\alpha = 0.05$ and $n = 19$. So, the critical value is 5. Because $n \leq 25$, the test statistic x is the smaller number of + or − signs. So, $x = 7$. Because $x = 7$ is greater than the critical value, the employee should fail to reject the null hypothesis.

Interpretation There is not enough evidence at the 5% level of significance for the employee to reject the website administrator's claim that the median number of visitors per day to the company's website is no more than 1500.

Try It Yourself 1

A real estate agency claims that the median number of days a home is on the market in its city is greater than 120. A homeowner wants to verify the accuracy of this claim. The number of days on the market for 24 randomly selected homes is shown below. At $\alpha = 0.025$, can the homeowner support the agency's claim?

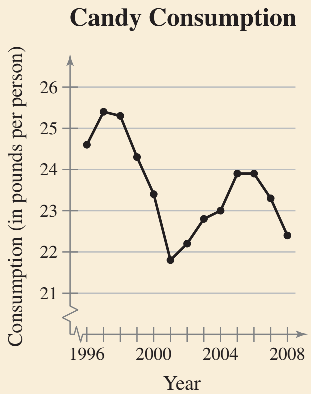
118	167	72	79	76	106	102	113
73	119	162	114	120	93	135	147
77	157	115	88	152	70	65	91

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine the *sample size* n.
d. Determine the *critical value*.
e. Find the *test statistic* x.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A47

PICTURING THE WORLD

In 2008, people in the United States spent a total of about \$15.5 billion on candy. The U.S. Department of Commerce reported that in 2008, the average person in the United States ate about 22.4 pounds of candy.



If you were to test the U.S. Department of Commerce's claim concerning per capita candy consumption, would you use a parametric test or a nonparametric test? What factors must you consider?

EXAMPLE 2

Using the Sign Test

An organization claims that the median annual attendance for museums in the United States is at least 39,000. A random sample of 125 museums reveals that the annual attendances for 79 museums were less than 39,000, the annual attendances for 42 museums were more than 39,000, and the annual attendances for 4 museums were 39,000. At $\alpha = 0.01$, is there enough evidence to reject the organization's claim? *(Adapted from American Association of Museums)*

Solution

The claim is "the median annual attendance for museums in the United States is at least 39,000." So, the null and alternative hypotheses are

H_0: median $\geq$ 39,000 (Claim) and H_a: median $<$ 39,000.

Because $n > 25$, use Table 4, the Standard Normal Table, to find the critical value. Because the test is a left-tailed test with $\alpha = 0.01$, the critical value is $z_0 = -2.33$. Of the 125 museums, there are 79 $-$ signs and 42 $+$ signs. When the 0s are ignored, the sample size is

$$n = 79 + 42 = 121, \quad \text{and} \quad x = 42.$$

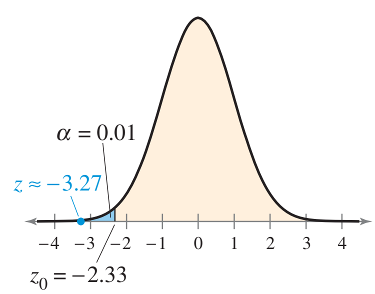
With these values, the test statistic is

$$z = \frac{(42 + 0.5) - 0.5(121)}{\sqrt{121}/2}$$

$$= \frac{-18}{5.5}$$

$$\approx -3.27.$$

The graph at the right shows the location of the rejection region and the test statistic z. Because z is less than the critical value, it is in the rejection region. So, you should reject the null hypothesis.



Interpretation There is enough evidence at the 1% level of significance to reject the organization's claim that the median annual attendance for museums in the United States is at least 39,000.

Try It Yourself 2

An organization claims that the median age of automobiles in operation in the United States is 9.4 years. A random sample of 95 automobiles reveals that 41 automobiles were less than 9.4 years old and 51 automobiles were more than 9.4 years old. At $\alpha = 0.10$, can you reject the organization's claim? *(Source: Bureau of Transportation Statistics)*

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine the *sample size* n.
d. Determine the *critical value*.
e. Find the *test statistic* z.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A47

STUDY TIP

When performing a two-tailed sign test, remember to use only the left-tailed critical value.

THE PAIRED-SAMPLE SIGN TEST

In Section 8.3, you learned how to use a t-test for the difference between means of dependent samples. That test required both populations to be normally distributed. If the parametric condition of normality cannot be satisfied, you can use the paired-sample sign test to test the difference between two population medians. To perform the paired-sample sign test for the difference between two population medians, the following conditions must be met.

1. A sample must be randomly selected from each population.
2. The samples must be dependent (paired).

The paired-sample sign test can be left-tailed, right-tailed, or two-tailed. This test is similar to the sign test for a single population median. However, instead of comparing each data entry with a hypothesized median and recording a $+$, $-$, or 0, you find the difference between corresponding data entries and record the sign of the difference. Generally, to find the difference, subtract the entry representing the second variable from the entry representing the first variable. Then compare the number of $+$ and $-$ signs. (The 0's are ignored.) If the number of $+$ signs is approximately equal to the number of $-$ signs, the null hypothesis should not be rejected. If, however, there is a significant difference between the number of $+$ signs and the number of $-$ signs, the null hypothesis should be rejected.

GUIDELINES

Performing a Paired-Sample Sign Test

IN WORDS	IN SYMBOLS
1. Identify the claim. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Determine the sample size n by finding the difference for each data pair. Assign a $+$ sign for a positive difference, a $-$ sign for a negative difference, and a 0 for no difference.	n = total number of $+$ and $-$ signs
4. Determine the critical value.	Use Table 8 in Appendix B.
5. Find the test statistic.	x = smaller number of $+$ or $-$ signs
6. Make a decision to reject or fail to reject the null hypothesis.	If the test statistic is less than or equal to the critical value, reject H_0. Otherwise, fail to reject H_0.
7. Interpret the decision in the context of the original claim.	

EXAMPLE 3

Using the Paired-Sample Sign Test

A psychologist claims that the number of repeat offenders will decrease if first-time offenders complete a particular rehabilitation course. You randomly select 10 prisons and record the number of repeat offenders during a two-year period. Then, after first-time offenders complete the course, you record the number of repeat offenders at each prison for another two-year period. The results are shown in the following table. At $\alpha = 0.025$, can you support the psychologist's claim?

Prison	1	2	3	4	5	6	7	8	9	10
Before	21	34	9	45	30	54	37	36	33	40
After	19	22	16	31	21	30	22	18	17	21

Solution

To support the psychologist's claim, you could use the following null and alternative hypotheses.

H_0: The number of repeat offenders will not decrease.

H_a: The number of repeat offenders will decrease. (Claim)

The table below shows the sign of the differences between the "before" and "after" data.

Prison	1	2	3	4	5	6	7	8	9	10
Before	21	34	9	45	30	54	37	36	33	40
After	19	22	16	31	21	30	22	18	17	21
Sign	+	+	−	+	+	+	+	+	+	+

You can see that there is 1 − sign and there are 9 + signs. So, $n = 1 + 9 = 10$. In Table 8 with $\alpha = 0.025$ (one-tailed) and $n = 10$, the critical value is 1. The test statistic x is the smaller number of + or − signs. So, $x = 1$. Because x is equal to the critical value, you should reject the null hypothesis.

Interpretation There is enough evidence at the 2.5% level of significance to support the psychologist's claim that the number of repeat offenders will decrease.

Try It Yourself 3

A medical researcher claims that a new vaccine will decrease the number of colds in adults. You randomly select 14 adults and record the number of colds each has in a one-year period. After giving the vaccine to each adult, you again record the number of colds each has in a one-year period. The results are shown in the table at the left. At $\alpha = 0.05$, can you support the researcher's claim?

Adult	**Before vaccine**	**After vaccine**
1	3	2
2	4	1
3	2	0
4	1	1
5	3	1
6	6	3
7	4	3
8	5	2
9	2	2
10	0	2
11	2	3
12	5	4
13	3	3
14	3	2

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine the *sample size* n.
d. Determine the *critical value*.
e. Find the *test statistic* x.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A47

11.1 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. What is a nonparametric test? How does a nonparametric test differ from a parametric test? What are the advantages and disadvantages of using a nonparametric test?

2. When the sign test is used, what population parameter is being tested?

3. Describe the test statistic for the sign test when the sample size n is less than or equal to 25 and when n is greater than 25.

4. In your own words, explain why the hypothesis test discussed in this section is called the sign test.

5. Explain how to use the sign test to test a population median.

6. List the two conditions that must be met in order to use the paired-sample sign test.

USING AND INTERPRETING CONCEPTS

Performing a Sign Test *In Exercises 7–22, (a) identify the claim and state H_0 and H_a, (b) determine the critical value, (c) find the test statistic, (d) decide whether to reject or fail to reject the null hypothesis, and (e) interpret the decision in the context of the original claim.*

7. Credit Card Charges In order to estimate the median amount of new credit card charges for the previous month, a financial service accountant randomly selects 12 credit card accounts and records the amount of new charges for each account for the previous month. The amounts (in dollars) are listed below. At $\alpha = 0.01$, can the accountant conclude that the median amount of new credit card charges for the previous month was more than \$300? *(Adapted from Board of Governors of the Federal Reserve System)*

346.71	382.59	255.03	202.17	309.80	265.88
299.41	270.38	296.54	318.46	245.92	309.47

8. Temperature A meteorologist estimates that the median daily high temperature for the month of July in Pittsburgh is 83° Fahrenheit. The high temperatures (in degrees Fahrenheit) for 15 randomly selected July days in Pittsburgh are listed below. At $\alpha = 0.01$, is there enough evidence to reject the meteorologist's claim? *(Adapted from U.S. National Oceanic and Atmospheric Administration)*

74 79 81 86 90 79 81 83 81 74 78 76 84 82 85

9. Sales Prices of Homes A real estate agent believes that the median sales price of new privately owned one-family homes sold in the past year is \$198,000 or less. The sales prices (in dollars) of 10 randomly selected homes are listed below. At $\alpha = 0.05$, is there enough evidence to reject the agent's claim? *(Adapted from National Association of Realtors)*

205,800	234,500	210,900	195,700	145,200
198,900	254,000	175,900	189,500	212,500

10. Temperature During a weather report, a meteorologist states that the median daily high temperature for the month of January in San Diego is 66° Fahrenheit. The high temperatures (in degrees Fahrenheit) for 16 randomly selected January days in San Diego are listed below. At $\alpha = 0.01$, can you reject the meteorologist's claim? *(Adapted from U.S. National Oceanic and Atmospheric Administration)*

78 74 72 72 70 70 72 78 74 71 72 74 77 79 75 73

11. Credit Card Debt A financial services institution reports that the median amount of credit card debt for families holding such debts is at least $3000. In a random sample of 104 families holding debt, the debts of 60 families were less than $3000 and the debts of 44 families were greater than $3000. At $\alpha = 0.02$, can you reject the institution's claim? *(Adapted from Board of Governors of the Federal Reserve System)*

12. Financial Debt A financial services accountant estimates that the median amount of financial debt for families holding such debts is less than $65,000. In a random sample of 70 families holding debts, the debts of 24 families were less than $65,000 and the debts of 46 families were greater than $65,000. At $\alpha = 0.025$, can you support the accountant's estimate? *(Adapted from Board of Governors of the Federal Reserve System)*

13. Twitter® Users A research group claims that the median age of Twitter® users is greater than 30 years old. In a random sample of 24 Twitter® users, 11 are less than 30 years old, 10 are more than 30 years old, and 3 are 30 years old. At $\alpha = 0.01$, can you support the research group's claim? *(Adapted from Pew Research Center)*

14. Facebook® Users A research group claims that the median age of Facebook® users is less than 32 years old. In a random sample of 20 Facebook® users, 5 are less than 32 years old, 13 are more than 32 years old, and 2 are 32 years old. At $\alpha = 0.05$, can you support the research group's claim? *(Adapted from Pew Research Center)*

15. Unit Size A renters' organization claims that the median number of rooms in renter-occupied units is four. You randomly select 120 renter-occupied units and obtain the results shown below. At $\alpha = 0.05$, can you reject the organization's claim? *(Adapted from U.S. Census Bureau)*

Unit size	Number of units
Fewer than 4 rooms	31
4 rooms	40
More than 4 rooms	49

TABLE FOR EXERCISE 15

Unit size	Number of units
Less than 1350	7
1350	3
More than 1350	12

TABLE FOR EXERCISE 16

16. Square Footage A renters' organization believes that the median square footage of renter-occupied units is 1350 square feet. To test this claim, you randomly select 22 renter-occupied units and obtain the results shown above. At $\alpha = 0.10$, can you reject the organization's claim? *(Adapted from U.S. Census Bureau)*

17. Hourly Wages A labor organization estimates that the median hourly wage of computer systems analysts is $37.06. In a random sample of 45 computer systems analysts, 18 are paid less than $37.06 per hour, 25 are paid more than $37.06 per hour, and 2 are paid $37.06 per hour. At $\alpha = 0.01$, can you reject the labor organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

18. Hourly Wages A labor organization estimates that the median hourly wage of podiatrists is at least \$55.89. In a random sample of 23 podiatrists, 17 are paid less than \$55.89 per hour, 5 are paid more than \$55.89 per hour, and 1 is paid \$55.89 per hour. At $\alpha = 0.05$, can you reject the labor organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

19. Lower Back Pain The table shows the lower back pain intensity scores for eight patients before and after receiving acupuncture for eight weeks. At $\alpha = 0.05$, is there enough evidence to conclude that the lower back pain intensity scores decreased after the acupuncture? *(Adapted from Archives of Internal Medicine)*

Patient	1	2	3	4	5	6	7	8
Intensity score (before)	59.2	46.3	65.4	74.0	79.3	81.6	44.4	59.1
Intensity score (after)	12.4	22.5	18.6	59.3	70.1	70.2	13.2	25.9

20. Lower Back Pain The table shows the lower back pain intensity scores for 12 patients before and after taking anti-inflammatory drugs for 8 weeks. At $\alpha = 0.05$, is there enough evidence to conclude that the lower back pain intensity scores decreased after taking anti-inflammatory drugs? *(Adapted from Archives of Internal Medicine)*

Patient	1	2	3	4	5	6
Intensity score (before)	71.0	42.1	79.1	57.5	64.0	60.4
Intensity score (after)	60.1	23.4	86.2	62.1	44.2	49.7

Patient	7	8	9	10	11	12
Intensity score (before)	68.3	95.2	48.1	78.6	65.4	59.9
Intensity score (after)	58.3	72.6	51.8	82.5	63.2	47.9

21. Improving SAT Scores A tutoring agency believes that by completing a special course, students can improve their critical reading SAT scores. In part of a study, 12 students take the critical reading part of the SAT, complete the special course, then take the critical reading part of the SAT again. The students' scores are shown below. At $\alpha = 0.05$, is there enough evidence to conclude that the students' critical reading SAT scores improved?

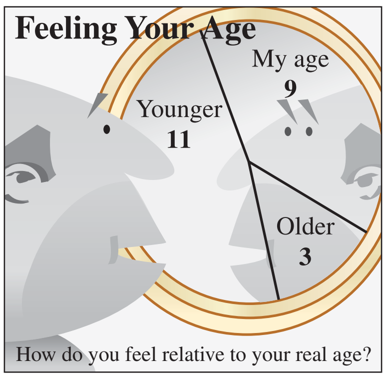
Student	1	2	3	4	5	6
Score on first SAT	308	456	352	433	306	471
Score on second SAT	300	524	409	419	304	483

Student	7	8	9	10	11	12
Score on first SAT	538	207	205	351	360	251
Score on second SAT	708	253	399	350	480	303

22. SAT Scores Students at a certain school are required to take the SAT twice. The table shows both critical reading SAT scores for 12 students. At $\alpha = 0.01$, can you conclude that the students' critical reading scores improved the second time they took the SAT?

Student	1	2	3	4	5	6
Score on first SAT	445	510	429	452	629	453
Score on second SAT	446	571	517	478	610	453

Student	7	8	9	10	11	12
Score on first SAT	358	477	325	513	636	571
Score on second SAT	378	532	299	501	648	603



Adapted from USA TODAY Snapshot, September 1, 2009.

FIGURE FOR EXERCISE 23

23. Feeling Your Age A research organization conducts a survey by randomly selecting adults and asking them how they feel relative to their real age. The results are shown in the figure. *(Adapted from Pew Research Center)*

(a) Use a sign test to test the null hypothesis that the proportion of adults who feel older than their real age is equal to the proportion of adults who feel younger than their real age. Assign a + sign to adults who feel older than their real age, assign a − sign to adults who feel younger than their real age, and assign a 0 to adults who feel their age. Use $\alpha = 0.05$.

(b) What can you conclude?

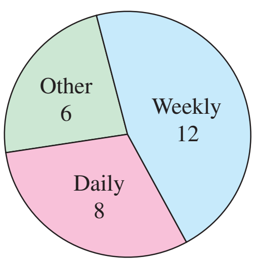


FIGURE FOR EXERCISE 24

24. Contacting Parents A research organization conducts a survey by randomly selecting adults and asking them how frequently they contact their parents by phone. The results are shown in the figure. *(Adapted from Pew Research Center)*

(a) Use a sign test to test the null hypothesis that the proportion of adults who contact their parents by phone weekly is equal to the proportion of adults who contact their parents by phone daily. Assign a + sign to adults who contact their parents by phone weekly, assign a − sign to adults who contact their parents by phone daily, and assign a 0 to adults who answer "other." Use $\alpha = 0.05$.

(b) What can you conclude?

SC *In Exercises 25 and 26, use StatCrunch to help you test the claim about the population median.*

25. Hourly Wages A labor organization claims that the median hourly wage of tool and die makers is $22.55. The hourly wages (in dollars) of 14 randomly selected tool and die makers are listed below. At $\alpha = 0.05$, is there enough evidence to reject the labor organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

21.75	23.10	20.50	25.80	29.25	26.35	27.40
22.90	23.50	22.55	32.70	30.05	29.80	34.85

26. Viewing Audience A television network claims that the median age of viewers for the Masters Golf Tournament is greater than 57 years. The ages of 24 randomly selected viewers are listed below. At $\alpha = 0.01$, is there enough evidence to support the network's claim? *(Adapted from ESPN)*

60	85	70	59	42	21	57	25	65	71	33	40
54	50	57	49	50	30	27	57	17	90	35	46

EXTENDING CONCEPTS

More on Sign Tests *When you are using a sign test for $n > 25$ and the test is left-tailed, you know you can reject the null hypothesis if the test statistic*

$$z = \frac{(x + 0.5) - 0.5n}{\frac{\sqrt{n}}{2}}$$

is less than or equal to the ***left-tailed*** *critical value, where x is the* ***smaller*** *number of + or − signs. For a right-tailed test, you can reject the null hypothesis if the test statistic*

$$z = \frac{(x - 0.5) - 0.5n}{\frac{\sqrt{n}}{2}}$$

is greater than or equal to the ***right-tailed*** *critical value, where x is the* ***larger*** *number of + or − signs.*

In Exercises 27–30, (a) write the claim mathematically and identify H_0 and H_a, (b) determine the critical value, (c) find the test statistic, (d) decide whether to reject or fail to reject the null hypothesis, and (e) interpret the decision in the context of the original claim.

27. Weekly Earnings A labor organization claims that the median weekly earnings of female workers is less than or equal to \$638. To test this claim, you randomly select 50 female workers and ask each to provide her weekly earnings. The results are shown in the table. At $\alpha = 0.01$, can you reject the organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

Weekly earnings	Number of workers
Less than \$638	18
\$638	3
More than \$638	29

TABLE FOR EXERCISE 27

Weekly earnings	Number of workers
Less than \$798	23
\$798	2
More than \$798	45

TABLE FOR EXERCISE 28

28. Weekly Earnings A labor organization states that the median weekly earnings of male workers is greater than \$798. To test this claim, you randomly select 70 male workers and ask each to provide his weekly earnings. The results are shown in the table. At $\alpha = 0.01$, can you support the organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

29. Ages of Brides A marriage counselor estimates that the median age of brides at the time of their first marriage is less than or equal to 26 years. In a random sample of 65 brides, 24 are less than 26 years old, 35 are more than 26 years old, and 6 are 26 years old. At $\alpha = 0.05$, can you reject the counselor's claim? *(Adapted from U.S. Census Bureau)*

30. Ages of Grooms A marriage counselor estimates that the median age of grooms at the time of their first marriage is greater than 28 years. In a random sample of 56 grooms, 33 are less than 28 years old, 23 are more than 28 years old, and none are 28 years old. At $\alpha = 0.05$, can you support the counselor's claim? *(Adapted from U.S. Census Bureau)*

11.2 The Wilcoxon Tests

WHAT YOU SHOULD LEARN

- How to use the Wilcoxon signed-rank test to determine if two dependent samples are selected from populations having the same distribution
- How to use the Wilcoxon rank sum test to determine if two independent samples are selected from populations having the same distribution

The Wilcoxon Signed-Rank Test ▸ The Wilcoxon Rank Sum Test

▸ THE WILCOXON SIGNED-RANK TEST

In this section, you will study the Wilcoxon signed-rank test and the Wilcoxon rank sum test. Unlike the sign test from Section 11.1, the strength of these two nonparametric tests is that each considers the magnitude, or size, of the data entries.

In Section 8.3, you used a t-test together with dependent samples to determine whether there was a difference between two populations. To use the t-test to test such a difference, you must assume (or know) that the dependent samples are randomly selected from populations having a normal distribution. But, what if this assumption cannot be made? Instead of using the two-sample t-test, you can use the *Wilcoxon signed-rank test*.

DEFINITION

The **Wilcoxon signed-rank test** is a nonparametric test that can be used to determine whether two *dependent* samples were selected from populations having the same distribution.

STUDY TIP

The absolute value of a number is its value, disregarding its sign. A pair of vertical bars, $|\;|$, is used to denote absolute value. For example, $|3| = 3$ and $|-7| = 7$.

GUIDELINES

Performing a Wilcoxon Signed-Rank Test

IN WORDS	IN SYMBOLS
1. Identify the claim. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Determine the sample size n, which is the number of pairs of data for which the difference is not 0.	
4. Determine the critical value.	Use Table 9 in Appendix B.
5. Find the test statistic w_s. **a.** Complete a table using the headers listed at the right. **b.** Find the sum of the positive ranks and the sum of the negative ranks. **c.** Select the smaller absolute value of the sums.	Headers: **Sample 1, Sample 2, Difference, Absolute value, Rank,** and **Signed rank.** Signed rank takes on the same sign as its corresponding difference.
6. Make a decision to reject or fail to reject the null hypothesis.	If w_s is less than or equal to the critical value, reject H_0. Otherwise, fail to reject H_0.
7. Interpret the decision in the context of the original claim.	

EXAMPLE 1

Performing a Wilcoxon Signed-Rank Test

A golf club manufacturer believes that golfers can lower their scores by using the manufacturer's newly designed golf clubs. The scores of 10 golfers while using the old design and while using the new design are shown in the table. At $\alpha = 0.05$, can you support the manufacturer's claim?

Golfer	1	2	3	4	5	6	7	8	9	10
Score (old design)	89	84	96	74	91	85	95	82	92	81
Score (new design)	83	83	92	76	91	80	87	85	90	77

Solution

The claim is "golfers can lower their scores." To test this claim, use the following null and alternative hypotheses.

H_0: The new design does not lower scores.

H_a: The new design lowers scores. (Claim)

This Wilcoxon signed-rank test is a one-tailed test with $\alpha = 0.05$, and because one data pair has a difference of 0, $n = 9$ instead of 10. From Table 9 in Appendix B, the critical value is 8. To find the test statistic w_s, complete a table as shown below.

Score (old design)	Score (new design)	Difference	Absolute value	Rank	Signed rank
89	83	6	6	8	8
84	83	1	1	1	1
96	92	4	4	5.5	5.5
74	76	−2	2	2.5	−2.5
91	91	0	0	—	—
85	80	5	5	7	7
95	87	8	8	9	9
82	85	−3	3	4	−4
92	90	2	2	2.5	2.5
81	77	4	4	5.5	5.5

The sum of the negative ranks is

$$-2.5 + (-4) = -6.5.$$

The sum of the positive ranks is

$$8 + 1 + 5.5 + 7 + 9 + 2.5 + 5.5 = 38.5.$$

The test statistic is the smaller absolute value of these two sums. Because $|-6.5| < |38.5|$, the test statistic is $w_s = 6.5$. Because the test statistic is less than the critical value, that is, $6.5 < 8$, you should decide to reject the null hypothesis.

Interpretation There is enough evidence at the 5% level of significance to support the claim that golfers can lower their scores by using the newly designed clubs.

STUDY TIP

Do not assign a rank to any difference of 0. In the case of a tie between data entries, use the average of the corresponding ranks. For instance, if two data entries are tied for the fifth rank, use the average of 5 and 6, which is 5.5, as the rank for both entries. The next data entry will be assigned a rank of 7, not 6.

If three entries are tied for the fifth rank, use the average of 5, 6, and 7, which is 6, as the rank for all three data entries. The next data entry will be assigned a rank of 8.

PICTURING THE WORLD

To help determine when knee arthroscopy patients can resume driving after surgery, the driving reaction times (in milliseconds) of 10 right knee arthroscopy patients were measured before surgery and 4 weeks after surgery using a computer-linked car simulator. The results are shown in the table. (Adapted from Knee Surgery, Sports Traumatology, Arthroscopy Journal)

Patient	Reaction time before surgery	Reaction time 4 weeks after surgery
1	720	730
2	750	645
3	735	745
4	730	640
5	755	660
6	745	670
7	730	650
8	725	730
9	770	675
10	700	705

At $\alpha = 0.05$, can you conclude that the reaction times changed significantly four weeks after surgery?

▸ Try It Yourself 1

A quality control inspector wants to test the claim that a spray-on water repellent is effective. To test this claim, he selects 12 pieces of fabric, sprays water on each, and measures the amount of water repelled (in milliliters). He then applies the water repellent and repeats the experiment. The results are shown in the table. At $\alpha = 0.01$, can he conclude that the water repellent is effective?

Fabric	1	2	3	4	5	6	7	8	9	10	11	12
No repellent	8	7	7	4	6	10	9	5	9	11	8	4
Repellent applied	15	12	11	6	6	8	8	6	12	8	14	8

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine the *sample size* n.
d. Determine the *critical value*.
e. Find the *test statistic* w_s by making a table, finding the sum of the positive ranks and the sum of the negative ranks, and finding the absolute value of each.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A47

▸ THE WILCOXON RANK SUM TEST

In Sections 8.1 and 8.2, you used a z-test or a t-test together with independent samples to determine whether there was a difference between two populations. To use the z-test to test such a difference, you must assume (or know) that the independent samples are randomly selected and that either each sample size is at least 30 or each population has a normal distribution with a known standard deviation. To use the t-test to test such a difference, you must assume (or know) that the independent samples are randomly selected from populations having a normal distribution. But, what if these assumptions cannot be made? You can still compare the populations using the *Wilcoxon rank sum test*.

DEFINITION

The **Wilcoxon rank sum test** is a nonparametric test that can be used to determine whether two *independent* samples were selected from populations having the same distribution.

STUDY TIP

Use the Wilcoxon signed-rank test for dependent samples and the Wilcoxon rank sum test for independent samples.

A requirement for the Wilcoxon rank sum test is that the sample sizes of both samples must be at least 10. When calculating the test statistic for the Wilcoxon rank sum test, let n_1 represent the sample size of the smaller sample and n_2 represent the sample size of the larger sample. If the two samples have the same size, it does not matter which one is n_1 or n_2.

When calculating the sum of the ranks R, combine both samples and rank the combined data. Then sum the ranks for the smaller of the two samples. If the two samples have the same size, you can use the ranks from either sample, but you must use the ranks from the sample you associate with n_1.

TEST STATISTIC FOR THE WILCOXON RANK SUM TEST

Given two independent samples, the **test statistic** z for the Wilcoxon rank sum test is

$$z = \frac{R - \mu_R}{\sigma_R}$$

where

R = sum of the ranks for the smaller sample,

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2},$$

and

$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}.$$

GUIDELINES

Performing a Wilcoxon Rank Sum Test

IN WORDS	IN SYMBOLS
1. Identify the claim. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Determine the critical value(s) and the rejection region(s).	Use Table 4 in Appendix B.
4. Determine the sample sizes.	$n_1 \le n_2$
5. Find the sum of the ranks for the smaller sample.	R
a. List the combined data in ascending order.	
b. Rank the combined data.	
c. Add the sum of the ranks for the smaller sample, n_1.	
6. Find the test statistic and sketch the sampling distribution.	$z = \frac{R - \mu_R}{\sigma_R}$
7. Make a decision to reject or fail to reject the null hypothesis.	If z is in the rejection region, reject H_0. Otherwise, fail to reject H_0.
8. Interpret the decision in the context of the original claim.	

EXAMPLE 2

Performing a Wilcoxon Rank Sum Test

The table shows the earnings (in thousands of dollars) of a random sample of 10 male and 12 female pharmaceutical sales representatives. At $\alpha = 0.10$, can you conclude that there is a difference between the males' and females' earnings?

Male earnings	78	93	114	101	98	94	86	95	117	99		
Female earnings	86	77	101	93	85	98	91	87	84	97	100	90

Solution

The claim is "there is a difference between the males' and females' earnings." The null and alternative hypotheses for this test are as follows.

H_0: There is no difference between the males' and the females' earnings.

H_a: There is a difference between the males' and the females' earnings. (Claim)

Because the test is a two-tailed test with $\alpha = 0.10$, the critical values are $-z_0 = -1.645$ and $z_0 = 1.645$. The rejection regions are $z < -1.645$ and $z > 1.645$.

The sample size for men is 10 and the sample size for women is 12. Because $10 < 12$, $n_1 = 10$ and $n_2 = 12$. Before calculating the test statistic, you must find the values of R, μ_R, and σ_R. The table shows the combined data listed in ascending order and the corresponding ranks.

STUDY TIP

Remember that in the case of a tie between data entries, use the average of the corresponding ranks.

Ordered data	Sample	Rank	Ordered data	Sample	Rank
77	F	1	94	M	12
78	M	2	95	M	13
84	F	3	97	F	14
85	F	4	98	M	15.5
86	M	5.5	98	F	15.5
86	F	5.5	99	M	17
87	F	7	100	F	18
90	F	8	101	M	19.5
91	F	9	101	F	19.5
93	M	10.5	114	M	21
93	F	10.5	117	M	22

Because the smaller sample is the sample of males, R is the sum of the male rankings.

$$R = 2 + 5.5 + 10.5 + 12 + 13 + 15.5 + 17 + 19.5 + 21 + 22$$
$$= 138$$

Using $n_1 = 10$ and $n_2 = 12$, you can find μ_R and σ_R as follows.

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} = \frac{10(10 + 12 + 1)}{2} = \frac{230}{2} = 115$$

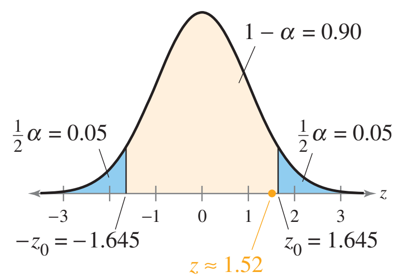
$$\sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}$$
$$= \sqrt{\frac{(10)(12)(10 + 12 + 1)}{12}}$$
$$= \sqrt{\frac{2760}{12}}$$
$$= \sqrt{230}$$
$$\approx 15.17$$

When $R = 138$, $\mu_R = 115$, and $\sigma_R \approx 15.17$, the test statistic is

$$z = \frac{R - \mu_R}{\sigma_R}$$
$$\approx \frac{138 - 115}{15.17}$$
$$\approx 1.52.$$

From the graph at the right, you can see that the test statistic z is not in the rejection region. So, you should decide to fail to reject the null hypothesis.

Interpretation There is not enough evidence at the 10% level of significance to conclude that there is a difference between the males' and females' earnings.



▸ Try It Yourself 2

You are investigating the automobile insurance claims paid (in thousands of dollars) by two insurance companies. The table shows a random, independent sample of 12 claims paid by the two insurance companies. At $\alpha = 0.05$, can you conclude that there is a difference in the claims paid by the companies?

Company A	6.2	10.6	2.5	4.5	6.5	7.4
Company B	7.3	5.6	3.4	1.8	2.2	4.7

Company A	9.9	3.0	5.8	3.9	6.0	6.3
Company B	10.8	4.1	1.7	3.0	4.4	5.3

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine the *critical value(s)* and the *rejection region(s)*.
d. Determine the *sample sizes* n_1 and n_2.
e. *List* the combined data in ascending order, *rank* the data, and *find* the sum of the ranks of the smaller sample.
f. Find the *test statistic* z. *Sketch* a graph.
g. *Decide* whether to reject the null hypothesis.
h. *Interpret* the decision in the context of the original claim.

Answer: Page A47

11.2 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. How do you know whether to use a Wilcoxon signed-rank test or a Wilcoxon rank sum test?

2. What is the requirement for the sample size of both samples when using the Wilcoxon rank sum test?

USING AND INTERPRETING CONCEPTS

Performing a Wilcoxon Test *In Exercises 3–8,*

(a) identify the claim and state H_0 and H_a.

(b) decide whether to use a Wilcoxon signed-rank test or a Wilcoxon rank sum test.

(c) determine the critical value(s).

(d) find the test statistic.

(e) decide whether to reject or fail to reject the null hypothesis.

(f) interpret the decision in the context of the original claim.

3. Calcium Supplements and Blood Pressure In a study testing the effects of calcium supplements on blood pressure in men, 12 men were randomly chosen and given a calcium supplement for 12 weeks. The measurements shown in the table are for each subject's diastolic blood pressure taken before and after the 12-week treatment period. At $\alpha = 0.01$, can you reject the claim that there was no reduction in diastolic blood pressure? *(Adapted from The Journal of the American Medical Association)*

Patient	1	2	3	4	5	6
Before treatment	108	109	120	129	112	111
After treatment	99	115	105	116	115	117

Patient	7	8	9	10	11	12
Before treatment	117	135	124	118	130	115
After treatment	108	122	120	126	128	106

4. Wholesale Trade and Manufacturing A private industry analyst claims that there is no difference in the salaries earned by workers in the wholesale trade and manufacturing industries. A random sample of 10 wholesale trade and 10 manufacturing workers and their salaries (in thousands of dollars) are shown in the table. At $\alpha = 0.10$, can you reject the analyst's claim? *(Adapted from U.S. Bureau of Economic Analysis)*

Wholesale trade	62	55	56	70	53	59	64	67	65	62
Manufacturing	62	58	47	65	45	56	67	49	55	43

5. Drug Prices A researcher wants to determine whether the cost of prescription drugs is lower in Canada than in the United States. The researcher selects seven of the most popular brand-name prescription drugs and records the cost per pill (in U.S. dollars) of each. The results are shown in the table. At $\alpha = 0.05$, can the researcher conclude that the cost of prescription drugs is lower in Canada than in the United States? *(Adapted from Annals of Internal Medicine)*

Drug	1	2	3	4	5	6	7
Cost in U.S.	1.26	1.76	4.19	3.36	1.80	9.91	3.95
Cost in Canada	1.04	0.82	2.22	2.22	1.31	11.47	2.63

6. Earnings by Degree A college administrator believes that there is a difference in the earnings of people with bachelor's degrees and those with associate's degrees. The table shows the earnings (in thousands of dollars) of a random sample of 11 people with bachelor's degrees and 10 people with associate's degrees. At $\alpha = 0.05$, is there enough evidence to support the administrator's belief? *(Adapted from U.S. Census Bureau)*

Bachelor's degree	54	50	63	76	70	50	44	56	60	52	54
Associate's degree	36	39	47	33	38	38	45	45	42	34	

7. Teacher Salaries A teacher's union representative claims that there is a difference in the salaries earned by teachers in Wisconsin and Michigan. The table shows the salaries (in thousands of dollars) of a random sample of 11 teachers from Wisconsin and 12 teachers from Michigan. At $\alpha = 0.05$, is there enough evidence to support the representative's claim? *(Adapted from National Education Association)*

Wisconsin	51	59	52	46	51	55	53	51	50	50	64	
Michigan	57	61	51	58	53	63	57	63	55	49	54	72

8. Heart Rate A physician wants to determine whether an experimental medication affects an individual's heart rate. The physician selects 15 patients and measures the heart rate of each. The subjects then take the medication and have their heart rates measured after one hour. The results are shown in the table. At $\alpha = 0.05$, can the physician conclude that the experimental medication affects an individual's heart rate?

Patient	1	2	3	4	5	6	7	8
Heart rate (before)	72	81	75	76	79	74	65	67
Heart rate (after)	73	80	75	79	74	76	73	67

Patient	9	10	11	12	13	14	15
Heart rate (before)	76	83	66	75	76	78	68
Heart rate (after)	74	77	70	77	76	75	74

EXTENDING CONCEPTS

Wilcoxon Signed-Rank Test for $n > 30$ *If you are performing a Wilcoxon signed-rank test and the sample size n is greater than 30, you can use the Standard Normal Table and the following formula to find the test statistic.*

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

In Exercises 9 and 10, perform the indicated Wilcoxon signed-rank test using the test statistic for $n > 30$.

9. Fuel Additive A petroleum engineer wants to know whether a certain fuel additive improves a car's gas mileage. To decide, the engineer records the gas mileages (in miles per gallon) of 33 cars with and without the additive. The results are shown in the table. At $\alpha = 0.10$, can the engineer conclude that the gas mileage is improved?

Car	1	2	3	4	5	6	7	8	9	10	11
Without additive	36.4	36.4	36.6	36.6	36.8	36.9	37.0	37.1	37.2	37.2	36.7
With additive	36.7	36.9	37.0	37.5	38.0	38.1	38.4	38.7	38.8	38.9	36.3

Car	12	13	14	15	16	17	18	19	20	21	22
Without additive	37.5	37.6	37.8	37.9	37.9	38.1	38.4	40.2	40.5	40.9	35.0
With additive	38.9	39.0	39.1	39.4	39.4	39.5	39.8	40.0	40.0	40.1	36.3

Car	23	24	25	26	27	28	29	30	31	32	33
Without additive	32.7	33.6	34.2	35.1	35.2	35.3	35.5	35.9	36.0	36.1	37.2
With additive	32.8	34.2	34.7	34.9	34.9	35.3	35.9	36.4	36.6	36.6	38.3

10. Fuel Additive A petroleum engineer claims that a fuel additive improves gas mileage. The table shows the gas mileages (in miles per gallon) of 32 cars measured with and without the fuel additive. Test the petroleum engineer's claim at $\alpha = 0.05$.

Car	1	2	3	4	5	6	7	8	9	10	11
Without additive	34.0	34.2	34.4	34.4	34.6	34.8	35.6	35.7	30.2	31.6	32.3
With additive	36.6	36.7	37.2	37.2	37.3	37.4	37.6	37.7	34.2	34.9	34.9

Car	12	13	14	15	16	17	18	19	20	21	22
Without additive	33.0	33.1	33.7	33.7	33.8	35.7	36.1	36.1	36.6	36.6	36.8
With additive	34.9	35.7	36.0	36.2	36.5	37.8	38.1	38.2	38.3	38.3	38.7

Car	23	24	25	26	27	28	29	30	31	32
Without additive	37.1	37.1	37.2	37.9	37.9	38.0	38.0	38.4	38.8	42.1
With additive	38.8	38.9	39.1	39.1	39.2	39.4	39.8	40.3	40.8	43.2

CASE STUDY

College Ranks

Each year, Forbes and the Center for College Affordability and Productivity release a list of the best colleges in America. Six hundred undergraduate colleges and universities are ranked according to quality of education, 4-year graduation rate, post-graduate success, average student debt after 4 years, and number of students or faculty who have won competitive awards, such as Rhodes Scholarships or Nobel Prizes.

The table shows freshman class size by state for randomly selected colleges on the 2009 list.

Freshman Class Size			
CA	**MA**	**NC**	**PA**
236	540	3865	372
1703	1666	1699	327
320	596	1201	366
382	439	2073	588
202	1048	2781	957
202	2167	1291	453
458	643	2492	2400
252	754	3090	601
467	1297	4538	613
574	518	4804	399

EXERCISES

1. Construct a side-by-side box-and-whisker plot for the four states. Do any of the median freshman class sizes appear to be the same? Do any appear to be different?

In Exercises 2–5, use the sign test to test the claim. What can you conclude? Use $\alpha = 0.05$.

2. The median freshman class size at a California college is less than or equal to 400.

3. The median freshman class size at a Massachusetts college is greater than or equal to 750.

4. The median freshman class size at a Pennsylvania college is 500.

5. The median freshman class size at a North Carolina college is different from 2400.

In Exercises 6 and 7, use the Wilcoxon rank sum test to test the claim. Use $\alpha = 0.01$.

6. There is no difference between freshman class sizes for Pennsylvania colleges and California colleges.

7. There is a difference between freshman class sizes for Massachusetts colleges and North Carolina colleges.

11.3 The Kruskal-Wallis Test

WHAT YOU SHOULD LEARN

- How to use the Kruskal-Wallis test to determine whether three or more samples were selected from populations having the same distribution

The Kruskal-Wallis Test

THE KRUSKAL-WALLIS TEST

In Section 10.4, you learned how to use one-way ANOVA techniques to compare the means of three or more populations. When using one-way ANOVA, you should verify that each independent sample is selected from a population that is normally, or approximately normally, distributed. If, however, you cannot verify that the populations are normal, you can still compare the distributions of three or more populations. To do so, you can use the *Kruskal-Wallis test*.

DEFINITION

The **Kruskal-Wallis test** is a nonparametric test that can be used to determine whether three or more independent samples were selected from populations having the same distribution.

The null and alternative hypotheses for the Kruskal-Wallis test are as follows.

H_0: There is no difference in the distribution of the populations.

H_a: There is a difference in the distribution of the populations.

Two conditions for using the Kruskal-Wallis test are that each sample must be randomly selected and the size of each sample must be at least 5. If these conditions are met, then the sampling distribution for the Kruskal-Wallis test is approximated by a chi-square distribution with $k - 1$ degrees of freedom, where k is the number of samples. You can calculate the Kruskal-Wallis test statistic using the following formula.

TEST STATISTIC FOR THE KRUSKAL-WALLIS TEST

Given three or more independent samples, the **test statistic** for the Kruskal-Wallis test is

$$H = \frac{12}{N(N+1)}\left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right) - 3(N+1)$$

where

k represents the number of samples,
n_i is the size of the ith sample,
N is the sum of the sample sizes,

and

R_i is the sum of the ranks of the ith sample.

Performing a Kruskal-Wallis test consists of combining and ranking the sample data. The data are then separated according to sample and the sum of the ranks of each sample is calculated.

These sums are then used to calculate the test statistic H, which is an approximation of the variance of the rank sums. If the samples are selected from populations having the same distribution, the sums of the ranks will be approximately equal, H will be small, and the null hypothesis should not be rejected.

If, however, the samples are selected from populations not having the same distribution, the sums of the ranks will be quite different, H will be large, and the null hypothesis should be rejected.

Because the null hypothesis is rejected only when H is significantly large, the Kruskal-Wallis test is always a right-tailed test.

GUIDELINES

Performing a Kruskal-Wallis Test

IN WORDS	IN SYMBOLS
1. Identify the claim. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Identify the degrees of freedom.	d.f. $= k - 1$
4. Determine the critical value and the rejection region.	Use Table 6 in Appendix B.
5. Find the sum of the ranks for each sample.	
a. List the combined data in ascending order.	
b. Rank the combined data.	
6. Find the test statistic and sketch the sampling distribution.	$H = \frac{12}{N(N+1)} \cdot \left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right) - 3(N+1)$
7. Make a decision to reject or fail to reject the null hypothesis.	If H is in the rejection region, reject H_0. Otherwise, fail to reject H_0.
8. Interpret the decision in the context of the original claim.	

EXAMPLE 1 SC Report 51

Performing a Kruskal-Wallis Test

You want to compare the number of crimes reported in three police precincts in a city. To do so, you randomly select 10 weeks for each precinct and record the number of crimes reported. The results are shown in the table. At $\alpha = 0.01$, can you conclude that the distributions of crimes reported in the three police precincts are different?

Number of Crimes Reported for the Week		
101st Precinct (Sample 1)	**106th Precinct (Sample 2)**	**113th Precinct (Sample 3)**
60	65	69
52	55	51
49	64	70
52	66	61
50	53	67
48	58	65
57	50	62
45	54	59
44	70	60
56	62	63

Solution

You want to test the claim that there is a difference in the number of crimes reported in the three precincts. The null and alternative hypotheses are as follows.

H_0: There is no difference in the number of crimes reported in the three precincts.

H_a: There is a difference in the number of crimes reported in the three precincts. (Claim)

The test is a right-tailed test with $\alpha = 0.01$ and d.f. $= k - 1 = 3 - 1 = 2$. From Table 6, the critical value is $\chi_0^2 = 9.210$. Before calculating the test statistic, you must find the sum of the ranks for each sample. The table shows the combined data listed in ascending order and the corresponding ranks.

Ordered data	Sample	Rank	Ordered data	Sample	Rank	Ordered data	Sample	Rank
44	101st	1	54	106th	11	62	113th	20.5
45	101st	2	55	106th	12	63	113th	22
48	101st	3	56	101st	13	64	106th	23
49	101st	4	57	101st	14	65	106th	24.5
50	101st	5.5	58	106th	15	65	113th	24.5
50	106th	5.5	59	113th	16	66	106th	26
51	113th	7	60	101st	17.5	67	113th	27
52	101st	8.5	60	113th	17.5	69	113th	28
52	101st	8.5	61	113th	19	70	106th	29.5
53	106th	10	62	106th	20.5	70	113th	29.5

PICTURING THE WORLD

The following randomly collected data were used to compare the water temperatures (in degrees Fahrenheit) of cities bordering the Gulf of Mexico. (Adapted from National Oceanographic Data Center)

Cedar Key, FL (Sample 1)	Eugene Island, LA (Sample 2)	Dauphin Island, AL (Sample 3)
62	51	63
69	55	51
77	57	54
59	63	60
60	74	75
75	82	80
83	85	70
65	60	78
79	64	82
86	76	84
82	83	
	86	

At $\alpha = 0.05$, can you conclude that the temperature distributions of the three cities are different?

The sum of the ranks for each sample is as follows.

$$R_1 = 1 + 2 + 3 + 4 + 5.5 + 8.5 + 8.5 + 13 + 14 + 17.5 = 77$$

$$R_2 = 5.5 + 10 + 11 + 12 + 15 + 20.5 + 23 + 24.5 + 26 + 29.5 = 177$$

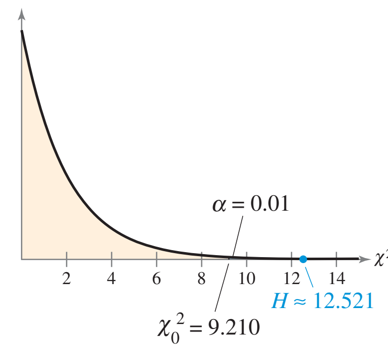
$$R_3 = 7 + 16 + 17.5 + 19 + 20.5 + 22 + 24.5 + 27 + 28 + 29.5 = 211$$

Using these sums and the values $n_1 = 10$, $n_2 = 10$, $n_3 = 10$, and $N = 30$, the test statistic is

$$H = \frac{12}{30(30+1)}\left(\frac{77^2}{10} + \frac{177^2}{10} + \frac{211^2}{10}\right) - 3(30+1) \approx 12.521.$$

From the graph at the right, you can see that the test statistic H is in the rejection region. So, you should decide to reject the null hypothesis.

Interpretation There is enough evidence at the 1% level of significance to support the claim that there is a difference in the number of crimes reported in the three police precincts.



Try It Yourself 1

You want to compare the salaries of veterinarians who work in California, New York, and Pennsylvania. To compare the salaries, you randomly select several veterinarians in each state and record their salaries. The salaries (in thousands of dollars) are listed in the table. At $\alpha = 0.05$, can you conclude that the distributions of the veterinarians' salaries in these three states are different? (*Adapted from U.S. Bureau of Labor Statistics*)

Sample Salaries		
CA (Sample 1)	**NY (Sample 2)**	**PA (Sample 3)**
99.95	94.40	99.20
97.50	99.75	103.70
98.85	97.50	110.45
100.75	101.97	95.15
101.20	93.10	88.80
96.25	102.35	99.99
99.70	97.89	100.55
88.28	92.50	97.25
113.90	101.55	97.44
103.20		

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Identify the *degrees of freedom*.
d. Determine the *critical value* and the *rejection region*.
e. *List* the combined data in ascending order, *rank* the data, and *find* the sum of the ranks of each sample.
f. Find the *test statistic* H. *Sketch* a graph.
g. *Decide* whether to reject the null hypothesis.
h. *Interpret* the decision in the context of the original claim.

Answer: Page A48

11.3 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. What are the conditions for using a Kruskal-Wallis test?

2. Explain why the Kruskal-Wallis test is always a right-tailed test.

USING AND INTERPRETING CONCEPTS

Performing a Kruskal-Wallis Test *In Exercises 3–6, (a) identify the claim and state H_0 and H_a, (b) determine the critical value, (c) find the sums of the ranks for each sample and calculate the test statistic, (d) decide whether to reject or fail to reject the null hypothesis, and (e) interpret the decision in the context of the original claim.*

3. Home Insurance The table shows the annual premiums for a random sample of home insurance policies in Connecticut, Massachusetts, and Virginia. At $\alpha = 0.05$, can you conclude that the distributions of the annual premiums in these three states are different? *(Adapted from National Association of Insurance Commissioners)*

State	Annual Premium (in dollars)						
Connecticut	930	725	890	1040	1165	806	947
Massachusetts	1105	1025	980	1295	1110	889	757
Virginia	815	730	546	625	912	618	535

4. Hourly Rates A researcher wants to determine whether there is a difference in the hourly pay rates for registered nurses in three states: Indiana, Kentucky, and Ohio. The researcher randomly selects several registered nurses in each state and records the hourly pay rate for each in the table shown. At $\alpha = 0.05$, can the researcher conclude that the distributions of the registered nurses' hourly pay rates in these three states are different? *(Adapted from U.S. Bureau of Labor Statistics)*

State	Hourly Pay Rate (in dollars)						
Indiana	27.80	28.25	26.65	27.40	30.24	25.10	29.44
Kentucky	26.95	25.58	28.10	30.20	28.55	31.60	24.60
Ohio	25.75	30.15	31.55	31.82	25.25	27.80	

5. Annual Salaries The table shows the annual salaries for a random sample of workers in Kentucky, North Carolina, South Carolina, and West Virginia. At $\alpha = 0.10$, can you conclude that the distributions of the annual salaries in these four states are different? *(Adapted from U.S. Bureau of Labor Statistics)*

State	Annual Salary (in thousands of dollars)					
Kentucky	32.5	34.2	43.1	54.7	30.9	25.5
North Carolina	40.5	38.9	33.6	51.3	32.5	36.6
South Carolina	27.8	35.4	41.5	40.9	32.7	34.1
West Virginia	27.1	38.2	28.9	37.4	42.6	30.4

6. Caffeine Content The table shows the amounts of caffeine (in milligrams) in 16-ounce servings for a random sample of beverages. At $\alpha = 0.01$, can you conclude that the distributions of the amounts of caffeine in these four beverages are different? *(Source: Center for Science in the Public Interest)*

Beverage	Amount of Caffeine in 16-ounce Serving (in milligrams)						
Coffees	320	300	206	150	266		
Soft drinks	95	96	56	51	71	72	47
Energy drinks	200	141	160	152	154	166	
Teas	100	106	42	15	32	10	

Number of Job Offers			
A	B	C	D
5	8	5	2
4	10	4	3
7	9	3	5
6	7	5	4
5	10	7	2
4	6	8	3

TABLE FOR EXERCISES 7 AND 8

SC *In Exercises 7 and 8, use StatCrunch and the table at the left, which shows the number of job offers received by mechanical engineers who recently graduated from four colleges (A, B, C, D).*

7. At $\alpha = 0.01$, can you conclude that the distributions of the number of job offers at Colleges A, B, and C are different?

8. At $\alpha = 0.01$, can you conclude that the distributions of the number of job offers at all four colleges are different?

EXTENDING CONCEPTS

Comparing Two Tests *In Exercises 9 and 10, perform the indicated test using (a) a Kruskal-Wallis test and (b) a one-way ANOVA test, assuming that each population is normally distributed and the population variances are equal. Compare the results. If convenient, use technology to solve the problem.*

9. Hospital Patient Stays An insurance underwriter reports that the mean number of days patients spend in a hospital differs according to the region of the United States in which the patient lives. The table shows the number of days randomly selected patients spent in a hospital in four U.S. regions. At $\alpha = 0.01$, can you support the underwriter's claim? *(Adapted from U.S. National Center for Health Statistics)*

Region	Number of Days									
Northeast	8	6	6	3	5	11	3	8	1	6
Midwest	5	4	3	9	1	4	6	3	4	7
South	5	8	1	5	8	7	5	1		
West	2	3	6	6	5	4	3	6	5	

10. Energy Consumption The table shows the energy consumed (in millions of Btu) in one year for a random sample of households from four U.S. regions. At $\alpha = 0.01$, can you conclude that the mean energy consumptions are different? *(Adapted from U.S. Energy Information Administration)*

Region	Energy Consumed (in millions of Btu)										
Northeast	72	106	151	138	104	108	95	134	100	174	
Midwest	84	183	194	165	120	212	148	129	113	62	97
South	91	40	72	91	147	74	70	67			
West	74	32	78	28	106	39	118	63	70	56	

11.4 Rank Correlation

WHAT YOU SHOULD LEARN

- How to use the Spearman rank correlation coefficient to determine whether the correlation between two variables is significant

The Spearman Rank Correlation Coefficient

THE SPEARMAN RANK CORRELATION COEFFICIENT

In Section 9.1, you learned how to measure the strength of the relationship between two variables using the Pearson correlation coefficient r. Two requirements for the Pearson correlation coefficient are that the variables are linearly related and that the population represented by each variable is normally distributed. If these requirements cannot be met, you can examine the relationship between two variables using the nonparametric equivalent to the Pearson correlation coefficient—the *Spearman rank correlation coefficient*.

The Spearman rank correlation coefficient has several advantages over the Pearson correlation coefficient. For instance, the Spearman rank correlation coefficient can be used to describe the relationship between linear or nonlinear data. The Spearman rank correlation coefficient can be used for data at the ordinal level. And, the Spearman rank correlation coefficient is easier to calculate by hand than the Pearson coefficient.

DEFINITION

The **Spearman rank correlation coefficient r_s** is a measure of the strength of the relationship between two variables. The Spearman rank correlation coefficient is calculated using the ranks of paired sample data entries. If there are no ties in the ranks of either variable, then the formula for the Spearman rank correlation coefficient is

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$

where n is the number of paired data entries and d is the difference between the ranks of a paired data entry. If there are ties in the ranks and the number of ties is small relative to the number of data pairs, then the formula can still be used to approximate r_s.

The values of r_s range from -1 to 1, inclusive. If the ranks of corresponding data pairs are exactly identical, r_s is equal to 1. If the ranks are in "reverse" order, r_s is equal to -1. If the ranks of corresponding data pairs have no relationship, r_s is equal to 0.

After calculating the Spearman rank correlation coefficient, you can determine whether the correlation between the variables is significant. You can make this determination by performing a hypothesis test for the population correlation coefficient ρ_s. The null and alternative hypotheses for this test are as follows.

H_0: $\rho_s = 0$ (There is no correlation between the variables.)

H_a: $\rho_s \neq 0$ (There is a significant correlation between the variables.)

The critical values for the Spearman rank correlation coefficient are listed in Table 10 of Appendix B. Table 10 lists critical values for selected levels of significance and for sample sizes of 30 or less. The test statistic for the hypothesis test is the Spearman rank correlation coefficient r_s.

GUIDELINES

Testing the Significance of the Spearman Rank Correlation Coefficient

IN WORDS	IN SYMBOLS
1. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance.	Identify α.
3. Determine the critical value.	Use Table 10 in Appendix B.
4. Find the test statistic.	$r_s = 1 - \dfrac{6\sum d^2}{n(n^2 - 1)}$
5. Make a decision to reject or fail to reject the null hypothesis.	If $\lvert r_s \rvert$ is greater than the critical value, reject H_0. Otherwise, fail to reject H_0.
6. Interpret the decision in the context of the original claim.	

EXAMPLE 1

▸ The Spearman Rank Correlation Coefficient

The table shows the school enrollments (in millions) at all levels of education for males and females from 2000 to 2007. At $\alpha = 0.05$, can you conclude that there is a correlation between the number of males and females enrolled in school? *(Source: U.S. Census Bureau)*

Year	Male	Female
2000	35.8	36.4
2001	36.3	36.9
2002	36.8	37.3
2003	37.3	37.6
2004	37.4	38.0
2005	37.4	38.4
2006	37.2	38.0
2007	37.6	38.4

▸ Solution

The null and alternative hypotheses are as follows.

H_0: $\rho_s = 0$ (There is no correlation between the number of males and females enrolled in school.)

H_a: $\rho_s \neq 0$ (There is a correlation between the number of males and females enrolled in school.) (Claim)

STUDY TIP

Remember that in the case of a tie between data entries, use the average of the corresponding ranks.

PICTURING THE WORLD

The table shows the retail prices (in dollars per pound) for 100% ground beef and fresh whole chicken in the United States from 2000 to 2008. (Source: U.S. Bureau of Labor Statistics)

Year	Beef	Chicken
2000	1.63	1.08
2001	1.71	1.11
2002	1.69	1.05
2003	2.23	1.05
2004	2.14	1.03
2005	2.30	1.06
2006	2.26	1.06
2007	2.23	1.17
2008	2.41	1.31

Does a correlation exist between ground beef and chicken prices in the United States from 2000 to 2008? Use $\alpha = 0.10$.

Each data set has eight entries. From Table 10 with $\alpha = 0.05$ and $n = 8$, the critical value is 0.738. Before calculating the test statistic, you must find Σd^2, the sum of the squares of the differences of the ranks of the data sets. You can use a table to calculate d^2, as shown below.

Male	Rank	Female	Rank	d	d^2
35.8	1	36.4	1	0	0
36.3	2	36.9	2	0	0
36.8	3	37.3	3	0	0
37.3	5	37.6	4	1	1
37.4	6.5	38.0	5.5	1	1
37.4	6.5	38.4	7.5	−1	1
37.2	4	38.0	5.5	−1.5	2.25
37.6	8	38.4	7.5	0.5	0.25
					$\Sigma d^2 = 5.5$

When $n = 8$ and $\Sigma d^2 = 5.5$, the test statistic is

$$r_s \approx 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$
$$= 1 - \frac{6(5.5)}{8(8^2 - 1)}$$
$$\approx 0.935.$$

Because $|0.935| > 0.738$, you should reject the null hypothesis.

Interpretation There is enough evidence at the 5% level of significance to conclude that there is a correlation between the number of males and females enrolled in school.

Try It Yourself 1

The table shows the number of males and females (in thousands) who received their doctoral degrees from 2001 to 2007. At $\alpha = 0.01$, can you conclude that there is a correlation between the number of males and females who received doctoral degrees? *(Source: U.S. National Center for Education Statistics)*

Year	2001	2002	2003	2004	2005	2006	2007
Male	25	24	24	25	27	29	30
Female	20	20	22	23	26	27	30

a. State the *null* and *alternative hypotheses*.
b. Specify the *level of significance* α.
c. Determine the *critical value*.
d. *Use a table* to calculate Σd^2.
e. Find the *test statistic* r_s.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A48

11.4 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. What are some advantages of the Spearman rank correlation coefficient over the Pearson correlation coefficient?

2. Describe the ranges of the Spearman rank correlation coefficient and the Pearson correlation coefficient.

3. What does it mean when r_s is equal to 1? What does it mean when r_s is equal to -1? What does it mean when r_s is equal to 0?

4. Explain, in your own words, what r_s and ρ_s represent in Example 1.

USING AND INTERPRETING CONCEPTS

Testing a Claim *In Exercises 5–8, (a) identify the claim and state H_0 and H_a, (b) determine the critical value using Table 10 in Appendix B, (c) find the test statistic (d) decide whether to reject or fail to reject the null hypothesis, and (e) interpret the decision in the context of the original claim.*

5. Farming: Debt and Income In an agricultural report, a commodities analyst suggests that there is a correlation between debt and income in the farming business. The table shows the total debts and total incomes for farms in seven states for a recent year. At $\alpha = 0.01$, is there enough evidence to support the analyst's claim? *(Adapted from U.S. Department of Agriculture)*

State	Debt (in millions of dollars)	Income (in millions of dollars)
California	19,955	28,926
Illinois	10,480	8,630
Iowa	14,434	12,942
Minnesota	9,982	8,807
Nebraska	10,085	11,028
North Carolina	4,235	7,008
Texas	13,286	15,268

6. Exercise Machines The table shows the overall scores and the prices for 11 different models of elliptical exercise machines. The overall score represents the ergonomics, exercise range, ease of use, construction, heart-rate monitoring, and safety. At $\alpha = 0.05$, can you conclude that there is a correlation between the overall score and the price? *(Source: Consumer Reports)*

Overall score	85	78	77	75	73	71
Price (in dollars)	2600	2800	3700	1700	1300	900

Overall score	66	66	64	62	58
Price (in dollars)	1000	1400	1800	1000	700

7. Crop Prices The table shows the prices (in dollars per bushel) received by U.S. farmers for oat and wheat from 2000 to 2008. At $\alpha = 0.01$, can you conclude that there is a correlation between the oat and wheat prices? *(Source: U.S. Department of Agriculture)*

Year	2000	2001	2002	2003	2004	2005	2006	2007	2008
Oat	1.10	1.59	1.81	1.48	1.48	1.63	1.87	2.63	3.10
Wheat	2.62	2.78	3.56	3.40	3.40	3.42	4.26	6.48	6.80

8. Vacuum Cleaners The table shows the overall scores and the prices for 12 different models of vacuum cleaners. The overall score represents carpet and bare-floor cleaning, airflow, handling, noise, and emissions. At $\alpha = 0.10$, can you conclude that there is a correlation between the overall score and the price? *(Source: Consumer Reports)*

Overall score	73	65	60	71	62	39
Price (in dollars)	230	400	600	350	100	300

Overall score	67	64	68	60	70	55
Price (in dollars)	600	700	140	200	80	300

Test Scores and GNI *In Exercises 9–12, use the following table. The table shows the average achievement scores of 15-year-olds in science and mathematics along with the gross national income (GNI) of nine countries for a recent year. (The GNI is a measure of the total value of goods and services produced by the economy of a country.) (Adapted from Organization for Economic Cooperation and Development; The World Bank)*

Country	Science average	Mathematics average	GNI (in billions of dollars)
Canada	534	527	1307
France	495	496	2467
Germany	516	504	3207
Italy	475	462	1988
Japan	531	523	4829
Mexico	410	406	989
Spain	488	480	1314
Sweden	503	502	438
United States	489	474	13,886

9. Science and GNI At $\alpha = 0.05$, can you conclude that there is a correlation between science achievement scores and GNI?

10. Math and GNI At $\alpha = 0.05$, can you conclude that there is a correlation between mathematics achievement scores and GNI?

11. Science and Math At $\alpha = 0.05$, can you conclude that there is a correlation between science and mathematics achievement scores?

12. Writing a Summary Use the results from Exercises 9–11 to write a summary about the correlation (or lack of correlation) between test scores and GNI.

EXTENDING CONCEPTS

Testing the Rank Correlation Coefficient for $n > 30$ *If you are testing the significance of the Spearman rank correlation coefficient and the sample size n is greater than 30, you can use the following expression to find the critical value.*

$$\frac{\pm z}{\sqrt{n-1}}, \quad z \text{ corresponds to the level of significance}$$

In Exercises 13 and 14, perform the indicated test.

13. Work Injuries The table shows the average hours worked per week and the number of on-the-job injuries for a random sample of U.S. industries in a recent year. At $\alpha = 0.05$, can you conclude that there is a correlation between average hours worked and the number of on-the-job injuries? *(Adapted from U.S. Bureau of Labor Statistics; National Safety Council)*

Hours worked	47.6	44.1	45.6	45.5	44.5	47.3	44.6	45.9	45.5	43.7	44.8	42.5
Injuries	16	33	25	33	18	20	21	18	21	28	15	26

Hours worked	46.5	42.3	45.5	41.8	43.1	44.4	44.5	43.7	44.9	47.8	46.6	45.5
Injuries	34	32	26	28	22	19	23	20	28	24	26	29

Hours worked	43.5	42.8	44.8	43.5	47.0	44.5	50.1	46.7	43.1
Injuries	21	28	23	26	24	20	28	26	25

14. Work Injuries in Construction The table shows the average hours worked per week and the number of on-the-job injuries for a random sample of U.S. construction companies in a recent year. At $\alpha = 0.05$, can you conclude that there is a correlation between average hours worked and the number of on-the-job injuries? *(Adapted from U.S. Bureau of Labor Statistics; National Safety Council)*

Hours worked	40.5	38.3	37.8	38.2	38.6	41.2	39.0	41.0	40.6	44.1	39.7	41.2
Injuries	12	13	19	18	22	22	17	13	15	10	18	19

Hours worked	41.1	38.2	42.3	39.2	36.1	36.2	38.7	36.0	37.3	36.5	37.9	38.0
Injuries	13	24	12	12	13	15	18	11	24	16	13	23

Hours worked	36.7	40.1	35.5	38.2	42.3	39.0	39.6	39.1	39.6	39.1
Injuries	14	10	5	14	13	18	15	23	15	23

11.5 The Runs Test

WHAT YOU SHOULD LEARN

- How to use the runs test to determine whether a data set is random

The Runs Test for Randomness

THE RUNS TEST FOR RANDOMNESS

In obtaining a sample of data, it is important for the data to be selected randomly. But how do you know if the sample data are truly random? One way to test for randomness in a data set is to use a *runs test for randomness*.

Before using a runs test for randomness, you must first know how to determine the number of runs in a data set.

DEFINITION

A **run** is a sequence of data having the same characteristic. Each run is preceded by and followed by data with a different characteristic or by no data at all. The number of data in a run is called the **length** of the run.

EXAMPLE 1

Finding the Number of Runs

A liquid-dispensing machine has been designed to fill one-liter bottles. A quality control inspector decides whether each bottle is filled to an acceptable level and passes inspection (P) or fails inspection (F). Determine the number of runs for each sequence and find the length of each run.

1. $P\ P\ P\ P\ P\ P\ P\ P\ F\ F\ F\ F\ F\ F\ F\ F$

2. $P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F$

3. $P\ P\ F\ F\ F\ F\ P\ F\ F\ F\ P\ P\ P\ P\ P\ P$

Solution

1. There are two runs. The first 8 P's form a run of length 8 and the first 8 F's form another run of length 8, as shown below.

$$\underbrace{P\ P\ P\ P\ P\ P\ P\ P}_{\text{1st run}}\qquad\underbrace{F\ F\ F\ F\ F\ F\ F\ F}_{\text{2nd run}}$$

2. There are 16 runs each of length 1, as shown below.

$$\underbrace{P}_{\text{1st run}}\ \underbrace{F}_{\text{2nd run}\ldots}\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ \underbrace{F}_{\ldots\text{16th run}}$$

3. There are 5 runs, the first of length 2, the second of length 4, the third of length 1, the fourth of length 3, and the fifth of length 6, as shown below.

$$\underbrace{P\ P}_{\text{1st run}}\quad\underbrace{F\ F\ F\ F}_{\text{2nd run}}\quad\underbrace{P}_{\text{3rd run}}\quad\underbrace{F\ F\ F}_{\text{4th run}}\quad\underbrace{P\ P\ P\ P\ P\ P}_{\text{5th run}}$$

▸ Try It Yourself 1

A machine produces engine parts. An inspector measures the diameter of each engine part and determines if the part passes inspection (P) or fails inspection (F). The results are shown below. Determine the number of runs in the sequence and find the length of each run.

$P\ P\ P\ F\ P\ F\ P\ P\ P\ P\ F\ F\ P\ F\ P\ P\ F\ F\ F\ P\ P\ P\ F\ P\ P\ P$

a. *Separate the data* each time there is a change in the characteristic of the data.
b. *Count the number of groups* to determine the number of runs.
c. *Count the number of data* within each run to determine the length.

Answer: Page A48

When each value in a set of data can be categorized into one of two separate categories, you can use the runs test for randomness to determine whether the data are random.

DEFINITION

The **runs test for randomness** is a nonparametric test that can be used to determine whether a sequence of sample data is random.

The runs test for randomness considers the number of runs in a sequence of sample data in order to test whether a sequence is random. If a sequence has too few or too many runs, it is usually not random. For instance, the sequence

$P\ P\ P\ P\ P\ P\ P\ P\ F\ F\ F\ F\ F\ F\ F\ F$

from Example 1, Part 1, has too few runs (only 2 runs). The sequence

$P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F\ P\ F$

from Example 1, Part 2, has too many runs (16 runs). So, these sample data are probably not random.

You can use a hypothesis test to determine whether the number of runs in a sequence of sample data is too high or too low. The runs test is a two-tailed test, and the null and alternative hypotheses are as follows.

H_0: The sequence of data is random.

H_a: The sequence of data is not random.

When using the runs test, let n_1 represent the number of data that have one characteristic and let n_2 represent the number of data that have the second characteristic. It does not matter which characteristic you choose to be represented by n_1. Let G represent the number of runs.

n_1 = number of data with one characteristic

n_2 = number of data with the other characteristic

G = number of runs

Table 12 in Appendix B lists the critical values for the runs test for selected values of n_1 and n_2 at the $\alpha = 0.05$ level of significance. (In this text, you will use only the $\alpha = 0.05$ level of significance when performing runs tests.) If n_1 or n_2 is greater than 20, you can use the standard normal distribution to find the critical values.

You can calculate the test statistic for the runs test as follows.

TEST STATISTIC FOR THE RUNS TEST

When $n_1 \leq 20$ and $n_2 \leq 20$, the **test statistic** for the runs test is G, the number of runs.

When $n_1 > 20$ or $n_2 > 20$, the **test statistic** for the runs test is

$$z = \frac{G - \mu_G}{\sigma_G}$$

where

$$\mu_G = \frac{2n_1n_2}{n_1 + n_2} + 1 \quad \text{and} \quad \sigma_G = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}.$$

GUIDELINES

Performing a Runs Test for Randomness

IN WORDS	IN SYMBOLS
1. Identify the claim. State the null and alternative hypotheses.	State H_0 and H_a.
2. Specify the level of significance. (Use $\alpha = 0.05$ for the runs test.)	Identify α.
3. Determine the number of data that have each characteristic and the number of runs.	Determine n_1, n_2, and G.
4. Determine the critical values.	If $n_1 \leq 20$ and $n_2 \leq 20$, use Table 12 in Appendix B. If $n_1 > 20$ or $n_2 > 20$, use Table 4 in Appendix B.
5. Find the test statistic.	If $n_1 \leq 20$ and $n_2 \leq 20$, use G. If $n_1 > 20$ or $n_2 > 20$, use $z = \frac{G - \mu_G}{\sigma_G}$.
6. Make a decision to reject or fail to reject the null hypothesis.	If G is less than or equal to the lower critical value or greater than or equal to the upper critical value, reject H_0. Otherwise, fail to reject H_0. Or, if z is in the rejection region, reject H_0. Otherwise, fail to reject H_0.
7. Interpret the decision in the context of the original claim.	

EXAMPLE 2

Using the Runs Test

As people enter a concert, an usher records where they are sitting. The results for 13 people are shown, where L represents a lawn seat and P represents a pavilion seat. At $\alpha = 0.05$, can you conclude that the sequence of seat locations is not random?

$L\ L\ L\ P\ P\ L\ P\ P\ P\ L\ L\ P\ L$

Solution

The claim is "the sequence of seat locations is not random." To test this claim, use the following null and alternative hypotheses.

H_0: The sequence of seat locations is random.

H_a: The sequence of seat locations is not random. (Claim)

To find the critical values, first determine n_1, the number of L's; n_2, the number of P's; and G, the number of runs.

$$\underbrace{L\,L\,L}_{\text{1st run}}\quad \underbrace{P\,P}_{\text{2nd run}}\quad \underbrace{L}_{\text{3rd run}}\quad \underbrace{P\,P\,P}_{\text{4th run}}\quad \underbrace{L\,L}_{\text{5th run}}\quad \underbrace{P}_{\text{6th run}}\quad \underbrace{L}_{\text{7th run}}$$

n_1 = number of L's = 7

n_2 = number of P's = 6

G = number of runs = 7

Because $n_1 \leq 20$, $n_2 \leq 20$, and $\alpha = 0.05$, use Table 12 to find the lower critical value 3 and the upper critical value 12. The test statistic is the number of runs $G = 7$. Because the test statistic G is between the critical values 3 and 12, you should fail to reject the null hypothesis.

Interpretation There is not enough evidence at the 5% level of significance to support the claim that the sequence of seat locations is not random. So, it appears that the sequence of seat locations is random.

Try It Yourself 2

The genders of 15 students as they enter a classroom are shown below, where F represents a female and M represents a male. At $\alpha = 0.05$, can you conclude that the sequence of genders is not random?

$M\ F\ F\ F\ M\ M\ F\ F\ M\ F\ M\ M\ F\ F\ F$

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine n_1, n_2, and G.
d. Determine the *critical values*.
e. Find the *test statistic* G.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A48

EXAMPLE 3

Using the Runs Test

You want to determine whether the selection of recently hired employees in a large company is random with respect to gender. The genders of 36 recently hired employees are shown below, where F represents a female and M represents a male. At $\alpha = 0.05$, can you conclude that the sequence of employees is not random?

$M\ M\ F\ F\ F\ F\ M\ M\ M\ M\ M\ M$
$F\ F\ F\ F\ F\ M\ M\ M\ M\ M\ M\ M$
$F\ F\ F\ M\ M\ M\ M\ F\ M\ M\ F\ M$

Solution

The claim is "the sequence of employees is not random." To test this claim, use the following null and alternative hypotheses.

H_0: The sequence of employees is random.

H_a: The sequence of employees is not random. (Claim)

To find the critical values, first determine n_1, the number of F's; n_2, the number of M's; and G, the number of runs.

$$\underbrace{M\,M}_{\text{1st run}}\quad\underbrace{F\,F\,F\,F}_{\text{2nd run}}\quad\underbrace{M\,M\,M\,M\,M\,M}_{\text{3rd run}}$$

$$\underbrace{F\,F\,F\,F\,F}_{\text{4th run}}\quad\underbrace{M\,M\,M\,M\,M\,M\,M}_{\text{5th run}}$$

$$\underbrace{F\,F\,F}_{\text{6th run}}\quad\underbrace{M\,M\,M\,M}_{\text{7th run}}\quad\underbrace{F}_{\text{8th run}}\quad\underbrace{M\,M}_{\text{9th run}}\quad\underbrace{F}_{\text{10th run}}\quad\underbrace{M}_{\text{11th run}}$$

n_1 = number of F's = 14

n_2 = number of M's = 22

G = number of runs = 11

Because $n_2 > 20$, use Table 4 in Appendix B to find the critical values. Because the test is a two-tailed test with $\alpha = 0.05$, the critical values are

$-z_0 = -1.96$

and

$z_0 = 1.96.$

Before calculating the test statistic, find the values of μ_G and σ_G, as follows.

$$\begin{aligned}\mu_G &= \frac{2n_1n_2}{n_1 + n_2} + 1\\ &= \frac{2(14)(22)}{14 + 22} + 1\\ &= \frac{616}{36} + 1\\ &\approx 18.11\end{aligned}$$

PICTURING THE WORLD

The table shows the National Football League conference of each winning team from Super Bowl I to Super Bowl XLIV, where *A* represents the American Football Conference and *N* represents the National Football Conference. (Source: National Football League)

Year	Conference	Year	Conference
1967	*N*	1989	*N*
1968	*N*	1990	*N*
1969	*A*	1991	*N*
1970	*A*	1992	*N*
1971	*A*	1993	*N*
1972	*N*	1994	*N*
1973	*A*	1995	*N*
1974	*A*	1996	*N*
1975	*A*	1997	*N*
1976	*A*	1998	*A*
1977	*A*	1999	*A*
1978	*N*	2000	*N*
1979	*A*	2001	*A*
1980	*A*	2002	*A*
1981	*A*	2003	*N*
1982	*N*	2004	*A*
1983	*N*	2005	*A*
1984	*A*	2006	*A*
1985	*N*	2007	*A*
1986	*N*	2008	*N*
1987	*N*	2009	*A*
1988	*N*	2010	*N*

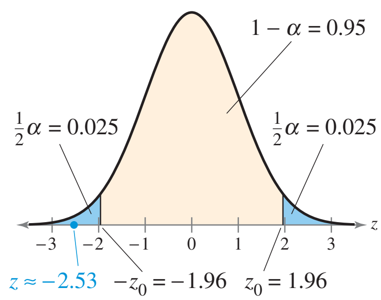
At $\alpha = 0.05$, can you conclude that the sequence of conferences of Super Bowl winning teams is not random?

$$\sigma_G = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$
$$= \sqrt{\frac{2(14)(22)[2(14)(22) - 14 - 22]}{(14 + 22)^2(14 + 22 - 1)}}$$
$$\approx 2.81$$

You can find the test statistic as follows.

$$z = \frac{G - \mu_G}{\sigma_G}$$
$$\approx \frac{11 - 18.11}{2.81}$$
$$\approx -2.53$$

From the graph below, you can see that the test statistic z is in the rejection region. So, you should decide to reject the null hypothesis.



Interpretation You have enough evidence at the 5% level of significance to support the claim that the sequence of employees with respect to gender is not random.

▸ Try It Yourself 3

Let *S* represent a day in a small town in which it snowed and let *N* represent a day in the same town in which it did not snow. The following are the snowfall results for the entire month of January. At $\alpha = 0.05$, can you conclude that the sequence is not random?

N N N S S N N S N S N N N N N S
N S N S N N S N S S N N N N N

a. Identify the *claim* and state H_0 and H_a.
b. Specify the *level of significance* α.
c. Determine n_1, n_2, and G.
d. Determine the *critical values*.
e. Find the *test statistic* z.
f. *Decide* whether to reject the null hypothesis.
g. *Interpret* the decision in the context of the original claim.

Answer: Page A48

When n_1 or n_2 is greater than 20, you can also use a P-value to perform a hypothesis test for the randomness of the data. In Example 3, you can calculate the P-value to be 0.0114. Because $P < \alpha$, you should reject the null hypothesis.

11.5 EXERCISES

BUILDING BASIC SKILLS AND VOCABULARY

1. In your own words, explain why the hypothesis test discussed in this section is called the runs test.
2. Describe the test statistic for the runs test when the sample sizes n_1 and n_2 are less than or equal to 20 and when either n_1 or n_2 is greater than 20.

USING AND INTERPRETING CONCEPTS

Finding the Number of Runs *In Exercises 3–6, determine the number of runs in the given sequence. Then find the length of each run.*

3. *T F T F T T T F F F T F*
4. *U U D D U D U U D D U D U U*
5. *M F M F M F F F F F F M M M F F M M M M*
6. *A A A B B B A B B A A A A A A B A A B A B B*
7. Find the values of n_1 and n_2 in Exercise 3.
8. Find the values of n_1 and n_2 in Exercise 4.
9. Find the values of n_1 and n_2 in Exercise 5.
10. Find the values of n_1 and n_2 in Exercise 6.

Finding Critical Values *In Exercises 11–14, use the given sequence and Table 12 in Appendix B to determine the number of runs that are considered too high and the number of runs that are considered too low for the data to be in random order.*

11. *T F T F T F T F T F T F*
12. *M F M M M M M M F F M M*
13. *N S S S N N N N N S N S N S S N N N*
14. *X X X X X X X Y Y Y Y Y Y Y Y Y Y Y Y Y*

Performing a Runs Test *In Exercises 15–20, use the runs test to (a) identify the claim and state H_0 and H_a, (b) determine the critical values using Table 4 or Table 12 in Appendix B, (c) find the test statistic, (d) decide whether to reject or fail to reject the null hypothesis, and (e) interpret the decision in the context of the original claim. Use $\alpha = 0.05$.*

15. **Coin Toss** A coach records the results of the coin toss at the beginning of each football game for a season. The results are shown, where *H* represents heads and *T* represents tails. The coach claimed the tosses were not random. Use the runs test to test the coach's claim.

    H T T T H T H H T T T T H T H H

16. **Senate** The sequence shows the majority party of the U.S. Senate after each election for a recent group of years, where *R* represents the Republican party and *D* represents the Democratic party. Can you conclude that the sequence is not random? *(Source: United States Senate)*

    R D D D R R R R R R R D D D D D D D
    R D D R D D D D D D D D D D D D D
    R R R D D D D R R R D R R D D

17. Baseball The sequence shows the Major League Baseball league of each World Series winning team from 1969 to 2009, where *N* represents the National League and *A* represents the American League. Can you conclude that the sequence of leagues of World Series winning teams is not random? *(Source: Major League Baseball)*

N A N A A A N N A A N N N N A A A N A N
A N A A A N A N A A A N A N A A N A N A

18. Number Generator A number generator outputs the sequence of digits shown, where *O* represents an odd digit and *E* represents an even digit. Test the claim that the digits were not randomly generated.

O O O E E E E O O O O O E E E E
O O E E E E O O O O E E E E O O

19. Dog Identifications A team of veterinarians record, in order, the genders of every dog that is microchipped at their pet hospital in one month. The genders of recently microchipped dogs are shown, where *F* represents a female and *M* represents a male. A veterinarian claims that the microchips are random by gender. Do you have enough evidence to reject the doctor's claim?

M M F M F F F F F M M M F F F
M F F F F F M F F F M F F F

20. Golf Tournament A golf tournament official records whether each past winner is American-born (*A*) or foreign-born (*F*). The results are shown for every year the tournament has existed. Can you conclude that the sequence is not random?

F F A F F A F F A F F A F F A F F A F F F F F F
A F F A F F A F F A F F A F A F F A F F F F F A
F F F F F A F F F A

EXTENDING CONCEPTS

Runs Test with Quantitative Data *In Exercises 21–23, use the following information to perform a runs test. You can also use the runs test for randomness with quantitative data. First, calculate the median. Then assign + to those values above the median and − to those values below the median. Ignore any values that are equal to the median. Use* $\alpha = 0.05$.

21. Daily High Temperatures The sequence shows the daily high temperatures (in degrees Fahrenheit) for a city during the month of July. Test the claim that the daily high temperatures do not occur randomly.

84 87 92 93 95 84 82 83 81 87 92 98 99 93 84 85
86 92 91 95 84 92 83 81 87 92 98 89 93 84 85

22. Exam Scores The sequence shows the exam scores of a class based on the order in which the students finished the test. Test the claim that the scores occur randomly.

83 94 80 76 92 89 65 75 82 87 90 91 81 99 97 72
72 89 90 92 87 76 74 66 88 81 90 92 89 76 80

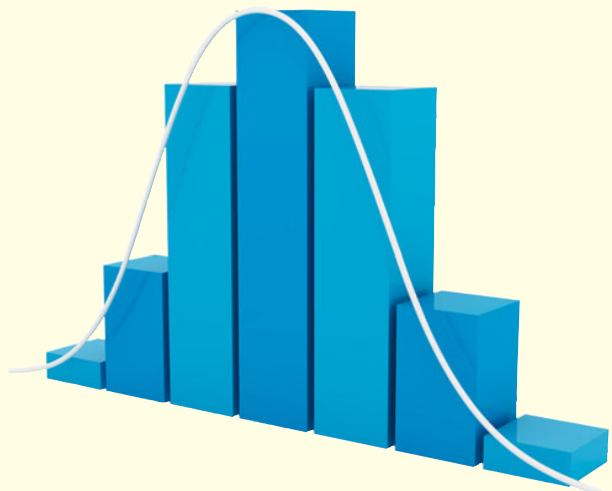
23. Use a technology tool to generate a sequence of 30 numbers from 1 to 99, inclusive. Test the claim that the sequence of numbers is not random.

USES AND ABUSES

Statistics in the Real World

Uses

Nonparametric Tests Before you could perform many of the hypothesis tests you learned about in previous chapters, you had to ensure that certain conditions about the population were satisfied. For instance, before you could run a t-test, you had to verify that the population was normally distributed. One advantage of the nonparametric tests shown in this chapter is that they are distribution free. That is, they do not require any particular information about the population or populations being tested. Another advantage of nonparametric tests is that they are easier to perform than their parametric counterparts. This means that they are easier to understand and quicker to use. Nonparametric tests can often be used when data are at the nominal or ordinal level.



Abuses

Insufficient Evidence Stronger evidence is needed to reject a null hypothesis in a nonparametric test than in a corresponding parametric test. That is, when you are trying to support a claim represented by the alternative hypothesis, you might need a larger sample when performing a nonparametric test. If the outcome of a nonparametric test results in failure to reject the null hypothesis, you should investigate the sample size used. It may be that a larger sample will produce different results.

Using an Inappropriate Test In general, when information about the population (such as the condition of normality) is known, it is more efficient to use a parametric test. However, if information about the population is not known, nonparametric tests can be helpful.

EXERCISES

1. ***Insufficient Evidence*** Give an example of a nonparametric test in which there is not enough evidence to reject the null hypothesis.
2. ***Using an Inappropriate Test*** Discuss the nonparametric tests described in this chapter and match each test with its parametric counterpart, which you studied in earlier chapters.

11 CHAPTER SUMMARY

What did you learn?	EXAMPLE(S)	REVIEW EXERCISES
Section 11.1		
■ How to use the sign test to test a population median $$z = \frac{(x + 0.5) - 0.5n}{\frac{\sqrt{n}}{2}}$$	*1, 2*	*1–3, 6*
■ How to use the paired-sample sign test to test the difference between two population medians (dependent samples)	*3*	*4, 5*
Section 11.2		
■ How to use the Wilcoxon signed-rank test and the Wilcoxon rank sum test to test the difference between two population distributions $$z = \frac{R - \mu_R}{\sigma_R}, \quad \mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}, \quad \sigma_R = \sqrt{\frac{n_1 n_2(n_1 + n_2 + 1)}{12}}$$	*1, 2*	*7, 8*
Section 11.3		
■ How to use the Kruskal-Wallis test to determine whether three or more samples were selected from populations having the same distribution $$H = \frac{12}{N(N + 1)}\left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \cdots + \frac{R_k^2}{n_k}\right) - 3(N + 1)$$	*1*	*9, 10*
Section 11.4		
■ How to use the Spearman rank correlation coefficient to determine whether the correlation between two variables is significant $$r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$$	*1*	*11, 12*
Section 11.5		
■ How to use the runs test to determine whether a data set is random $$G = \text{number of runs}, \quad z = \frac{G - \mu_G}{\sigma_G}, \quad \mu_G = \frac{2n_1 n_2}{n_1 + n_2} + 1, \quad \sigma_G = \sqrt{\frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$	*1–3*	*13, 14*

The table summarizes parametric and nonparametric tests. Always use the parametric test if the conditions for that test are satisfied.

Test application	Parametric test	Nonparametric test
One-sample tests	z-test for a population mean t-test for a population mean	Sign test for a population median
Two-sample tests		
Dependent samples	t-test for the difference between means	Paired-sample sign test Wilcoxon signed-rank test
Independent samples	z-test for the difference between means t-test for the difference between means	Wilcoxon rank sum test
Tests involving three or more samples	One-way ANOVA	Kruskal-Wallis test
Correlation	Pearson correlation coefficient	Spearman rank correlation coefficient
Randomness	(No parametric test)	Runs test

11 REVIEW EXERCISES

SECTION 11.1

In Exercises 1–6, use a sign test to test the claim by doing the following.

(a) Identify the claim and state H_0 and H_a.
(b) Determine the critical value.
(c) Find the test statistic.
(d) Decide whether to reject or fail to reject the null hypothesis.
(e) Interpret the decision in the context of the original claim.

1. A bank manager claims that the median number of customers per day is no more than 650. The number of bank customers per day for 17 randomly selected days are listed below. At $\alpha = 0.01$, can you reject the bank manager's claim?

675 665 601 642 554 653 639 650 645
550 677 569 650 660 682 689 590

2. A company claims that the median credit score for U.S. adults is at least 710. The credit scores for 13 randomly selected U.S. adults are listed below. At $\alpha = 0.05$, can you reject the company's claim? *(Adapted from Fair Isaac Corporation)*

750 782 805 695 700 706 625
589 690 772 745 704 710

3. A government agency estimates that the median sentence length for all federal prisoners is 2 years. In a random sample of 180 federal prisoners, 65 have sentence lengths that are less than 2 years, 109 have sentence lengths that are more than 2 years, and 6 have sentence lengths that are 2 years. At $\alpha = 0.10$, can you reject the agency's claim? *(Adapted from United States Sentencing Commission)*

4. In a study testing the effects of calcium supplements on blood pressure in men, 10 randomly selected men were given a calcium supplement for 12 weeks. The following measurements are for each subject's diastolic blood pressure taken before and after the 12-week treatment period. At $\alpha = 0.05$, can you reject the claim that there was no reduction in diastolic blood pressure? *(Adapted from the American Medical Association)*

Patient	1	2	3	4	5	6	7
Before treatment	107	110	123	129	112	111	107
After treatment	100	114	105	112	115	116	106

Patient	8	9	10
Before treatment	112	136	102
After treatment	102	125	104

5. In a study testing the effects of an herbal supplement on blood pressure in men, 11 randomly selected men were given an herbal supplement for 12 weeks. The following measurements are for each subject's diastolic blood pressure taken before and after the 12-week treatment period. At $\alpha = 0.05$, can you reject the claim that there was no reduction in diastolic blood pressure? *(Adapted from The Journal of the American Medical Association)*

Patient	1	2	3	4	5	6	7	8	9	10	11
Before treatment	123	109	112	102	98	114	119	112	110	117	130
After treatment	124	97	113	105	95	119	114	114	121	118	133

6. An association claims that the median annual salary of lawyers 9 months after graduation from law school is \$68,500. In a random sample of 125 lawyers 9 months after graduation from law school, 76 were paid less than \$68,500, and 49 were paid more than \$68,500. At $\alpha = 0.05$, can you reject the association's claim? *(Adapted from National Association of Law Placement)*

SECTION 11.2

In Exercises 7 and 8, use a Wilcoxon test to test the claim by doing the following.

(a) Decide whether the samples are dependent or independent; then choose the appropriate Wilcoxon test.

(b) Identify the claim and state H_0 and H_a.

(c) Determine the critical values.

(d) Find the test statistic.

(e) Decide whether to reject or fail to reject the null hypothesis.

(f) Interpret the decision in the context of the original claim.

7. A career placement advisor estimates that there is a difference in the total times required to earn a doctorate degree by female and male graduate students. The table shows the total times to earn a doctorate for a random sample of 12 female and 12 male graduate students. At $\alpha = 0.01$, can you support the advisor's claim? *(Adapted from National Opinion Research Council)*

Gender	Total Time (in years)											
Female	13	12	10	13	12	9	11	14	7	7	9	10
Male	11	8	9	11	10	8	8	10	11	9	10	8

8. A medical researcher claims that a new drug affects the number of headache hours experienced by headache sufferers. The number of headache hours (per day) experienced by eight randomly selected patients before and after taking the drug are shown in the table. At $\alpha = 0.05$, can you support the researcher's claim?

Patient	1	2	3	4	5	6	7	8
Headache hours (before)	0.9	2.3	2.7	2.4	2.9	1.9	1.2	3.1
Headache hours (after)	1.4	1.5	1.4	1.8	1.3	0.6	0.7	1.9

SECTION 11.3

In Exercises 9 and 10, use the Kruskal-Wallis test to test the claim by doing the following.

(a) Identify the claim and state H_0 and H_a.
(b) Determine the critical value.
(c) Find the sums of the ranks for each sample and calculate the test statistic.
(d) Decide whether to reject or fail to reject the null hypothesis.
(e) Interpret the decision in the context of the original claim.

9. The table shows the ages for a random sample of doctorate recipients in three fields of study. At $\alpha = 0.01$, can you conclude that the distributions of the ages of the doctorate recipients in these three fields of study are different? *(Adapted from Survey of Earned Doctorates)*

Field of Study	**Age**										
Life sciences	31	32	34	31	30	32	35	31	32	34	29
Physical sciences	30	31	32	31	30	29	31	30	32	33	30
Social sciences	32	35	31	33	34	31	35	36	32	30	33

10. The table shows the starting salary offers for a random sample of college graduates in four fields of engineering. At $\alpha = 0.05$, can you conclude that the distributions of the starting salaries in these four fields of engineering are different? *(Adapted from National Association of Colleges and Employers)*

Field of Engineering	**Starting Salary (in thousands of dollars)**									
Chemical engineering	66.4	63.9	69.7	68.5	62.3	67.9	65.5	63.7	67.4	69.1
Computer engineering	61.1	60.5	58.7	59.3	62.4	65.5	59.9	63.1	61.4	59.3
Electrical engineering	59.3	57.9	58.5	56.8	60.0	59.7	61.3	60.5	59.5	59.8
Mechanical engineering	58.9	58.2	59.0	57.1	59.0	58.7	61.5	62.0	58.3	56.1

SECTION 11.4

In Exercises 11 and 12, use the Spearman rank correlation coefficient to test the claim by doing the following.

(a) Identify the claim mathematically and state H_0 and H_a.
(b) Determine the critical value using Table 10 in Appendix B.
(c) Find the test statistic.
(d) Decide whether to reject the null hypothesis.
(e) Interpret the decision in the context of the original claim.

11. The table shows the overall scores and the prices for seven randomly selected Blu-ray™ players. The overall score is based mainly on picture quality. At $\alpha = 0.05$, can you conclude that there is a correlation between overall score and price? *(Source: Consumer Reports)*

Overall score	93	91	90	87	85	74	69
Price (in dollars)	500	300	500	150	250	200	130

12. The table shows the overall scores and the prices per gallon for nine randomly selected interior paints. The overall score represents hiding, surface smoothness, and resistance to staining, scrubbing, gloss change, sticking, mildew, and fading. At $\alpha = 0.10$, can you conclude that there is a correlation between overall score and price? *(Adapted from Consumer Reports)*

Overall score	86	84	82	81	75	74	71	69	62
Price per gallon (in dollars)	33	32	20	45	19	25	25	18	37

SECTION 11.5

In Exercises 13 and 14, use the runs test to (a) identify the claim and state H_0 and H_a, (b) determine the critical values using Table 4 or Table 12 in Appendix B, (c) calculate the test statistic, (d) decide whether to reject or fail to reject the null hypothesis, and (e) interpret the decision in the context of the original claim. Use $a = 0.05$.

13. A highway patrol officer stops speeding vehicles on an interstate highway. The following shows the genders of the last 25 drivers who were stopped, where F represents a female driver and M represents a male driver. Can you conclude that the stops were not random by gender?

F M M M F M F M F F F M M
F F F M M M F M M F F M

14. The following data represent the departure status of the last 18 buses to leave a bus station, where T represents a bus that departed on time and L represents a bus that departed late. Can you conclude that the departure status of the buses is not random?

T T T T L L L L T
L L L T T T T T T

11 CHAPTER QUIZ

Take this quiz as you would take a quiz in class. After you are done, check your work against the answers given in the back of the book.

For each exercise, (a) identify the claim and state H_0 and H_a, (b) decide which test to use, (c) determine the critical value(s), (d) find the test statistic, (e) decide whether to reject or fail to reject the null hypothesis, and (f) interpret the decision in the context of the original claim.

1. A labor organization claims that there is a difference in the hourly earnings of union and nonunion workers in state and local governments. A random sample of 10 union and 10 nonunion workers in state and local governments and their hourly earnings are listed in the tables. At $\alpha = 0.10$, can you support the organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

Union					**Nonunion**				
27.20	25.60	29.75	32.97	30.33	24.80	21.75	19.85	25.60	20.70
25.30	24.80	26.50	25.05	24.20	23.40	21.15	20.90	20.05	19.10

2. An organization claims that the median number of annual volunteer hours is 52. In a random sample of 75 people who volunteered last year, 47 volunteered for less than 52 hours, 23 volunteered for more than 52 hours, and 5 volunteered for 52 hours. At $\alpha = 0.05$, can you reject the organization's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

3. The table shows the sales prices for a random sample of apartment condominiums and cooperatives in four U.S. regions. At $\alpha = 0.01$, can you conclude that the distributions of the sales prices in these four regions are different? *(Adapted from National Association of Realtors)*

Region	**Sales Price (in thousands of dollars)**							
Northeast	252.5	245.5	237.9	270.2	265.9	250.0	259.4	238.6
Midwest	188.9	205.1	200.9	175.9	170.5	191.9	185.3	187.1
South	175.5	150.9	149.8	164.6	169.5	190.5	172.6	161.0
West	218.5	201.9	255.7	230.0	189.9	225.7	220.0	206.3

4. A meteorologist wants to determine whether days with rain occur randomly in April in his home town. To do so, the meteorologist records whether it rains for each day in April. The results are shown, where R represents a day with rain and N represents a day with no rain. At $\alpha = 0.05$, can the meteorologist conclude that days with rain are not random?

$N\ R\ R\ N\ N\ N\ N\ R\ N\ R\ R\ N\ R\ R\ R$
$N\ R\ R\ R\ R\ N\ N\ N\ N\ R\ N\ R\ N\ N\ R$

5. The table shows the number of larceny-thefts (per 100,000 population) and the number of motor vehicle thefts (per 100,000 population) in six randomly selected large U.S. cities. At $\alpha = 0.10$, can you conclude that there is a correlation between the number of larceny-thefts and the number of motor vehicle thefts? *(Source: U.S. Department of Justice)*

Larceny-thefts	1403	1506	2937	3449	2728	3042
Motor vehicle thefts	161	608	659	897	774	945

PUTTING IT ALL TOGETHER

Real Statistics — Real Decisions

In a recent year, according to the Bureau of Labor Statistics, the median number of years that wage and salary workers had been with their current employer (called employee tenure) was 4.1 years. Information on employee tenure has been gathered since 1996 using the *Current Population Survey* (*CPS*), a monthly survey of about 60,000 households that provides information on employment, unemployment, earnings, demographics, and other characteristics of the U.S. population ages 16 and over. With respect to employee tenure, the questions measure how long workers have been with their current employers, not how long they plan to stay with their employers.

www.bls.gov

EXERCISES

1. *How Would You Do It?*

(a) What sampling technique would you use to select the sample for the *CPS*?

(b) Do you think the technique in part (a) will give you a sample that is representative of the U.S. population? Why or why not?

(c) Identify possible flaws or biases in the survey on the basis of the technique you chose in part (a).

2. *Is There a Difference?*

A congressional representative claims that the median tenure for workers from the representative's district is less than the national median tenure of 4.1 years. The claim is based on the representative's data and is shown in the table at the right above. (Assume that the employees were randomly selected.)

(a) Is it possible that the claim is true? What questions should you ask about how the data were collected?

(b) How would you test the representative's claim? Can you use a parametric test, or do you need to use a nonparametric test?

(c) State the null hypothesis and the alternative hypothesis.

(d) Test the claim using $\alpha = 0.05$. What can you conclude?

3. *Comparing Male and Female Employee Tenures*

A congressional representative claims that there is a difference between the median tenures for male workers and female workers. The claim is based on the representative's data and is shown in the table at the right. (Assume that the employees were randomly selected from the representative's district.)

(a) How would you test the representative's claim? Can you use a parametric test, or do you need to use a nonparametric test?

(b) State the null hypothesis and the alternative hypothesis.

(c) Test the claim using $\alpha = 0.05$. What can you conclude?

Employee Tenure of 20 Workers		
4.6	2.6	3.3
2.8	1.5	1.9
4.0	5.0	3.9
5.1	3.7	5.4
3.6	3.9	6.2
1.7	4.6	3.1
4.4	3.6	

TABLE FOR EXERCISE 2

Employee tenure for a sample of male workers	Employee tenure for a sample of female workers
3.9	4.4
4.4	4.9
4.7	5.4
4.3	4.3
4.9	4.0
3.8	1.8
3.6	5.1
4.7	5.1
2.3	3.3
6.5	2.2
0.9	5.2
5.1	3.0
	1.3
	4.0

TABLE FOR EXERCISE 3

TECHNOLOGY

MINITAB | EXCEL | TI-83/84 PLUS

U.S. INCOME AND ECONOMIC RESEARCH

The National Bureau of Economic Research (NBER) is a private, nonprofit, nonpartisan research organization. The NBER provides information for better understanding of how the U.S. economy works. Researchers at the NBER concentrate on four types of empirical research: developing new statistical measurements, estimating quantitative models of economic behavior, assessing the effects of public policies on the U.S. economy, and projecting the effects of alternative policy proposals.

One of the NBER's interests is the median income of people in different regions of the United States. The table at the right shows the annual incomes (in dollars) of a random sample of people (15 years and over) in a recent year in four U.S. regions: Northeast, Midwest, South, and West.

Annual Income of People (in dollars)			
Northeast	**Midwest**	**South**	**West**
47,000	30,035	24,030	41,180
35,145	32,235	37,943	35,298
31,497	31,010	36,280	29,114
27,500	37,660	38,738	36,180
28,500	36,224	22,275	38,558
35,400	35,510	27,975	27,680
33,810	34,535	28,275	33,080
32,500	46,035	35,073	44,930
29,950	23,331	39,730	29,408
25,100	44,213	36,775	26,180
42,700	29,405	25,675	32,956
49,950	31,695	29,875	40,744

EXERCISES

In Exercises 1–5, refer to the annual income of people in the table. Use $\alpha = 0.05$ for all tests.

1. Construct a box-and-whisker plot for each region. Do the median annual incomes appear to differ between regions?

2. Use a technology tool to perform a sign test to test the claim that the median annual income in the Midwest is greater than $30,000.

3. Use a technology tool to perform a Wilcoxon rank sum test to test the claim that the median annual incomes in the Northeast and South are the same.

4. Use a technology tool to perform a Kruskal-Wallis test to test the claim that the distributions of annual incomes for all four regions are the same.

5. Use a technology tool to perform a one-way ANOVA to test the claim that the average annual incomes for all four regions are the same. Assume that the populations of incomes are normally distributed, the samples are independent, and the population variances are equal. How do your results compare with those in Exercise 4?

6. Repeat Exercises 1, 3, 4, and 5 using the data in the following table. The table shows the annual incomes (in dollars) of a random sample of families in a recent year in four U.S. regions: Northeast, Midwest, South, and West.

Annual Income of Families (in dollars)			
Northeast	**Midwest**	**South**	**West**
58,010	57,680	55,200	61,808
107,465	83,260	57,787	70,125
75,800	39,060	49,400	51,982
106,770	55,260	90,200	67,330
65,780	72,216	55,209	53,830
51,500	52,048	35,200	79,220
46,366	67,760	60,300	61,108
66,750	61,860	38,756	86,130
48,800	64,920	64,621	86,955
73,520	57,260	52,800	42,130
44,795	62,260	77,650	66,650
65,650	59,596	51,085	47,910
58,500	70,510	59,200	62,364
72,800	61,460	45,100	66,880
57,799	77,960	55,562	46,923

Extended solutions are given in the *Technology Supplement.*
Technical instruction is provided for MINITAB, Excel, and the TI-83/84 Plus.

CUMULATIVE REVIEW

Chapters 9 – 11

Men, x	Women, y
10.80	12.20
10.38	11.90
10.30	11.50
10.30	12.20
10.79	11.67
10.62	11.82
10.32	11.18
10.06	11.49
9.95	11.08
10.14	11.07
10.06	11.08
10.25	11.06
9.99	10.97
9.92	10.54
9.96	10.82
9.84	10.94
9.87	10.75
9.85	10.93
9.69	10.78

TABLE FOR EXERCISE 1

1. The table at the left shows the winning times (in seconds) for the men's and women's 100-meter runs in the Summer Olympics from 1928 to 2008. *(Source: The International Association of Athletics Federations)*

(a) Display the data in a scatter plot, calculate the correlation coefficient r, and make a conclusion about the type of correlation.

(b) Test the level of significance of the correlation coefficient r found in part (a). Use $\alpha = 0.05$.

(c) Find the equation of the regression line for the data. Draw the regression line on the scatter plot.

(d) Use the regression line to predict the women's 100-meter time when the men's 100-meter time is 9.90 seconds.

2. An employment agency claims that there is a difference in the weekly earnings of workers who are union members and workers who are not union members. A random sample of nine union members and eight nonunion members and their weekly earnings (in dollars) are shown in the table. At $\alpha = 0.05$, can you support the agency's claim? *(Adapted from U.S. Bureau of Labor Statistics)*

Union member	855	994	692	800	884	991	1040	904	930
Not a union member	758	691	862	557	655	814	803	638	

3. An investment company claims that the median age of people with mutual funds is 50 years. The ages (in years) of 20 randomly selected mutual fund owners are listed below. At $\alpha = 0.01$, is there enough evidence to reject the company's claim? *(Adapted from Investment Company Institute)*

46 34 33 27 58 64 54 36 38 42
26 51 49 44 46 50 39 34 51 63

4. The table at the right shows the residential natural gas expenditures (in dollars) in one year for a random sample of households in four regions of the United States. Assume that the populations are normally distributed and the population variances are equal. At $\alpha = 0.10$, can you reject the claim that the mean expenditures are equal for all four regions? *(Adapted from U.S. Energy Information Administration)*

Northeast	Midwest	South	West
1478	393	434	625
649	980	319	538
834	609	694	1045
1173	1157	678	497
1013	865	856	305
1565	1337	499	358
655	870	451	549
648	810	1021	633

5. The equation used to predict sweet potato yield (in pounds per acre) is $\hat{y} = 11{,}182 + 174.53x_1 - 104.41x_2$, where x_1 is the number of acres planted (in thousands) and x_2 is the number of acres harvested (in thousands). Use the multiple regression equation to predict the y-values for the given values of the independent variables. *(Adapted from United States Department of Agriculture)*

(a) $x_1 = 91, x_2 = 88$

(b) $x_1 = 110, x_2 = 98$

6. A school administrator reports that the standard deviations of reading test scores for eighth grade students are the same in Colorado and Utah. A random sample of 16 test scores from Colorado has a standard deviation of 34.6 points and a random sample of 15 test scores from Utah has a standard deviation of 33.2 points. At $\alpha = 0.10$, can you reject the administrator's claim? Assume the samples are independent and each population has a normal distribution. *(Adapted from National Center for Education Statistics)*

7. An employment agency representative wants to determine whether there is a difference in the annual household incomes in four regions of the United States. To do so, the representative randomly selects several households in each region and records the annual household income for each in the table. At $\alpha = 0.01$, can the representative conclude that the distributions of the annual household incomes in these regions are different? *(Adapted from U.S. Census Bureau)*

Region	Household Income (in thousands of dollars)						
Northeast	54.3	47.1	55.7	54.8	50.0	52.5	51.6
Midwest	49.3	54.4	45.2	48.5	50.7	51.8	52.0
South	44.4	45.6	49.2	41.5	46.4	49.2	47.0
West	56.8	54.7	51.4	53.5	52.4	54.0	55.9

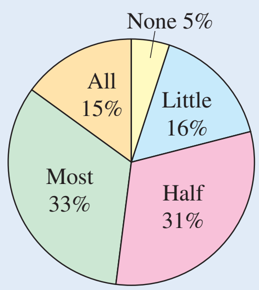


FIGURE FOR EXERCISE 8

8. Results from a previous survey asking U.S. parents how much they intend to contribute to the college costs of their children are shown in the pie chart. To determine whether this distribution is still the same, you randomly select 900 U.S. parents and ask them how much they intend to contribute to the college costs of their children. The results are shown in the table. At $\alpha = 0.05$, are the distributions different? *(Adapted from Sallie Mae, Inc.)*

Survey results

Response	Frequency, f
None	31
Little	164
Half	277
Most	305
All	123

9. The table shows the metacarpal bone lengths (in centimeters) and the heights (in centimeters) of nine adults. The equation of the regression line is $\hat{y} = 1.700x + 94.428$. *(Adapted from the American Journal of Physical Anthropology)*

Metacarpal bone length, x	45	51	39	41	48	49	46	43	47
Height, y	171	178	157	163	172	183	173	175	173

(a) Find the coefficient of determination and interpret the results.

(b) Find the standard error of estimate s_e and interpret the results.

(c) Construct a 95% prediction interval for the height of an adult when his or her metacarpal bone length is 50 centimeters. Interpret the results.

10. The table shows the overall scores and the prices of eight all-season tires. The overall score represents safety-related tests, such as braking, handling, and resistance to hydroplaning. At $\alpha = 0.10$, can you conclude that there is a correlation between the overall score and the price? Use the Spearman rank correlation coefficient. *(Source: Consumer Reports)*

Overall score	74	82	78	84	80	64	70	74
Price (in dollars)	77	96	77	116	98	67	70	81