Key Formulas	
Section	
17.6	**Test Statistic for the Kruskal-Wallis Test**

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

where $R_i = \sum_{j=1}^{n_i} r_{ij}$ and $N = \sum_{i=1}^{k} n_i$.

Additional Exercises

1. A new method for temporarily relieving the lung congestion of cystic fibrosis patients has been introduced. The traditional method of relieving the congestion involves a series of manual techniques where the chest and back area are pounded and massaged. The new method is a mechanical vest which has been designed to perform the manual techniques. A study is conducted to measure the effectiveness of the new vest. Five cystic fibrosis patients are randomly selected and the diameter of the blood vessels in their lungs is measured after using the traditional treatment and after using the vest treatment. The larger the diameter of the blood vessels within the lungs, the better the treatment. If the study provides conclusive evidence that the vest is more effective than the manual method in increasing the diameter of the blood vessels, the hospital will recommend the vest to its patients because the vest allows the patients to be much more independent. The results of the study are as follows.

Diameter of Lung Blood Vessels (in mm)					
Subject	1	2	3	4	5
After Traditional Treatment	0.5	0.4	0.7	0.6	0.2
After Vest Treatment	0.6	0.6	0.7	0.7	0.5

a. What assumption must be made in order to perform the test of hypothesis using the paired difference *t*-test?

b. Using the Sign Test, does the data provide conclusive evidence that the median diameter of blood vessels in the lungs is significantly larger after using the vest treatment than after using the traditional treatment at $\alpha = 0.01$?

c. What assumptions were made in performing the Sign Test?

d. Using the Signed-Rank Test, does the data provide conclusive evidence that the median diameter of blood vessels in the lungs is significantly larger after using the vest treatment than after using the traditional treatment at $\alpha = 0.01$?

e. What assumptions were made in performing the Signed-Rank Test?

f. Which test do you think produces more accurate results? Why?

g. Perform a paired difference *t*-test. How do the results of the Sign Test and the Signed-Rank Test compare with the results of the *t*-test?

2. As private companies prepare to go public, many analysts attempt to predict whether the stock will have a positive or negative return in the first day of trading. One particular analyst believes the return after the first day of trading is positive, on average. Another analyst wishes to test this analyst's claim using 15 recently offered public stocks.[2]

Initial Public Offerings		
Company	**Initial Offer Price ($)**	**Price After 1st Day of Trading ($)**
Enduro Royalty Trust (NDRO)	22.00	21.26
ZELTIQ Aesthetics (ZLTQ)	13.00	15.50
Ubiquiti Networks (UBNT)	15.00	17.50
Tudou Holdings Limited (TUDO)	29.00	25.56
Carbonite (CARB)	10.00	12.35
SandRidge Permian Trust (PER)	18.00	18.00
American Capital Mortgage Investment (MTGE)	20.00	18.41
C&J Energy Services (CJES)	29.00	30.50
Chefs Warehouse Holdings (CHEF)	15.00	17.50
Spirit Airlines (SAVE)	12.00	11.55
Pandora Media (P)	16.00	17.42
Wesco Aircraft Holdings (WAIR)	15.00	14.92
American Midstream Partners (AMID)	21.00	20.95
Dunkin Brands Group (DNKN)	19.00	27.85
Skullcandy (SKUL)	20.00	20.00

a. Use the Sign Test to test the hypothesis that the return after the first day of trading is positive. Test at the 0.05 level.

b. Suppose that an analyst claimed that the median return on a stock in the first day of trading is +$1.00. Perform a test of hypothesis to determine if the median return is different than what the analyst claims. Use $\alpha = 0.05$.

c. Use the Wilcoxon Signed-Rank Test to determine if the price after the first day of trading is generally greater than the initial offer price. Test at $\alpha = 0.05$.

d. Analyze the results of the Sign Test performed in part **a.** and the Signed-Rank Test performed in part **c.** in terms of the problem. Which test do you think yields more accurate results?

e. What concerns do you have with the tests performed in this problem?

3. A weight loss center is trying to determine which of its diets results in higher client satisfaction. The center polled 20 clients (10 were on Diet A and 10 on Diet B) and had them rate their satisfaction in the diets from 1 to 100.

Diet Ratings										
Diet A	84	77	89	98	97	100	75	85	96	78
Diet B	94	81	95	93	97	99	82	92	95	89

a. Which nonparametric test do you think is most appropriate to test the claim that Diet B results in higher client satisfaction than Diet A? Explain why.

b. Write the null and alternative hypotheses for a Rank-Sum Test to determine if Diet B results in higher client satisfaction than Diet A.

c. Using the Rank-Sum Test, does the data provide sufficient evidence at $\alpha = 0.05$ that Diet B results in greater client satisfaction than Diet A?

d. What assumptions were made in the test performed in part **c.**?

4. Are students with higher GPAs more likely to get a higher paying job upon graduation? Consider the following data regarding student GPA and starting salary.

GPA and Starting Salary

GPA	Starting Salary ($)
2.37	47,000
3.20	48,000
3.21	52,000
3.39	50,000
3.55	54,000
3.57	58,000
3.76	60,000
3.77	70,000
3.79	69,000
3.90	65,000

a. Determine the ranks for the *x*-variable, GPA.

b. Determine the ranks for the *y*-variable, Starting Salary.

c. Calculate the Spearman rank correlation coefficient.

d. Interpret the value of the coefficient. Is the relationship between these two variables positive or negative? Is this relationship what you expected? Explain.

e. Comment on the strength of the relationship between these two variables.

f. Is there evidence at $\alpha = 0.10$ that these two variables are related?

5. *Fortune* magazine releases a list of the world's most admired companies. In the survey they ask business people to vote for companies that they admire most from various industries. The table below lists the top 10 most admired companies of 2011 along with each company's Fortune 500 ranking. Note that the Fortune 500 ranks companies based on revenues.[3]

Top 10 Most Admired Companies and Profits

Company	Most Admired Ranking	Fortune 500 Ranking
Apple	1	35
Google	2	92
Berkshire Hathaway	3	7
Southwest Airlines	4	205
Procter & Gamble	5	26

Top 10 Most Admired Companies and Profits		
Company	**Most Admired Ranking**	**Fortune 500 Ranking**
Coca-Cola	6	70
Amazon.com	7	78
FedEx	8	73
Microsoft	9	38
McDonald's	10	111

a. Compute the Spearman rank correlation coefficient. Interpret this value.

b. With 95% confidence, can we conclude that there is an association between company admiration and revenue?

6. The given table shows key dates of the Dow-Jones Industrial Average. Apply the runs test to check for randomness at the 0.05 level. (**Hint:** First find the median of the values, then label each value by A if it is above the median and B if it is below the median.)

Significant Levels on the Dow (December 1974 to February 2009)		
Date	**Dow Jones Industrial Average**	**Significance**
December 6, 1974	577	The last Bear Market bottom
July 12, 1976	1011	Highest point between January, 1973 and October, 1982
August 12, 1982	776	The start of the "Reagan Bull"
August 25, 1987	2722	The 1987 high
October 19, 1987	1738	The (508 point) crash of 1987
February 2, 1994	3975	The top of the post 1987 crash recovery
November 23, 1994	3674	The start of the Clinton "super bull"
March 29, 1999	10,006	The first Dow close above 10,000
January 14, 2000	11,723	The "Clinton bull" high
March 17, 2000	10,630	The biggest one day gain (499 points)
March 20, 2001	9720	Dow closes below previous year low for the first time since 1982
September 11-14, 2001	9605	Terrorist attack closed the Dow for four days
September 17, 2001	8920	The biggest one day fall (685 points)
September 21, 2001	8235	The Dow's second worst week ever (–14.26%)
December 31, 2001	10,021	Dow up 21.7% from September 21 low but down 7.2% on the year
September 30, 2002	7591	New 2002 low – all treasury yields (except 30-year bond) at 2002 lows
October 9, 2002	7286	New 2002 low – Dow down 37.8% from the January, 2000 all-time high
October 31, 2002	8397	Dow up 806 points (10.6%) for October – first positive month since March
November 6, 2002	8771	Federal reserve cuts rates for the first time since December, 2001

Significant Levels on the Dow (December 1974 to February 2009)		
Date	**Dow Jones Industrial Average**	**Significance**
December 31, 2002	8341	Dow down 16.8% for 2002 – first three consecutive year loss since 1939-41
May 23, 2003	8601	Senate passes bill raising the treasury debt limit
June 25, 2003	9011	Federal reserve cuts rates by 0.25%
December 31, 2003	10,453	Dow up 25.32% in 2003
October 3, 2006	11,727	Dow exceeds the previous all-time high in January, 2000
October 9, 2007	14,164	New all-time high on the Dow
July 2, 2008	11,215	Dow closes more than 20% below the October, 2007 high
February 27, 2009	7062	Dow closes more than 50% below the October, 2007 high

7. The irrational number π can be approximated by the rational number $\frac{22}{7}$. Test the randomness of odd and even digits in $\frac{22}{7}$ at the 0.05 level using the first nine digits.

8. A polling agency conducts exit interviews after an election. If R = Republican and D = Democrat, the first 20 voter responses in a random sample are as follows.

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

Test for non-randomness using $\alpha = 0.05$.

9. Consider the following *U.S. News and World Report* college rankings for schools in the Big Ten, the Big 12, and the Atlantic Coast conferences.[4]

College Rankings by Conference		
Atlantic Coast	**Big Ten**	**Big 12**
29	55	101
10	62	45
101	42	58
71	28	143
68	45	90
31	71	94
55	45	101
25	12	75
38	68	101
101	71	132
36	75	160
25		97

With $\alpha = 0.10$, use the Kruskal-Wallis Test to determine if there is a significant difference in *U.S. News and World Report* rankings among the three athletic conferences.

10. Consider the following scores reported by *Condé Nast Traveler* in their annual list of the top cities to visit around the world. The results were determined from more than 8 million votes cast in the Readers' Choice Awards survey.[5,6,7]

Readers' Choice City Rankings, 2011					
Asia		**Europe**		**United States**	
Kyoto	82.3	Florence	85.0	Charleston, SC	84.7
Bangkok	81.6	Barcelona	82.8	San Francisco, CA	83.7
Hong Kong	81.1	Rome	82.4	Santa Fe, NM	83.0
Chiang Mai	80.8	Paris	81.9	Chicago, IL	82.2
Ubud	80.0	Bruges	81.7	Honolulu, HI	80.9
Singapore	78.4	Venice	81.7	New York, NY	80.8
Tokyo	76.8	Salzburg	81.4	Savannah, GA	79.1
Luang Prabang	76.4	Vienna	81.0	Carmel, CA	78.5
Thimphu	75.1	Prague	79.7	Seattle, WA	78.4
Shanghai	74.9	Siena	79.7	Boston, MA	78.0

a. Using the Kruskal-Wallis Test and $\alpha = 0.05$, test to determine if there is a difference in rankings between Asia, Europe, and the United States.

b. What assumptions were made for the test performed in part **a.**?