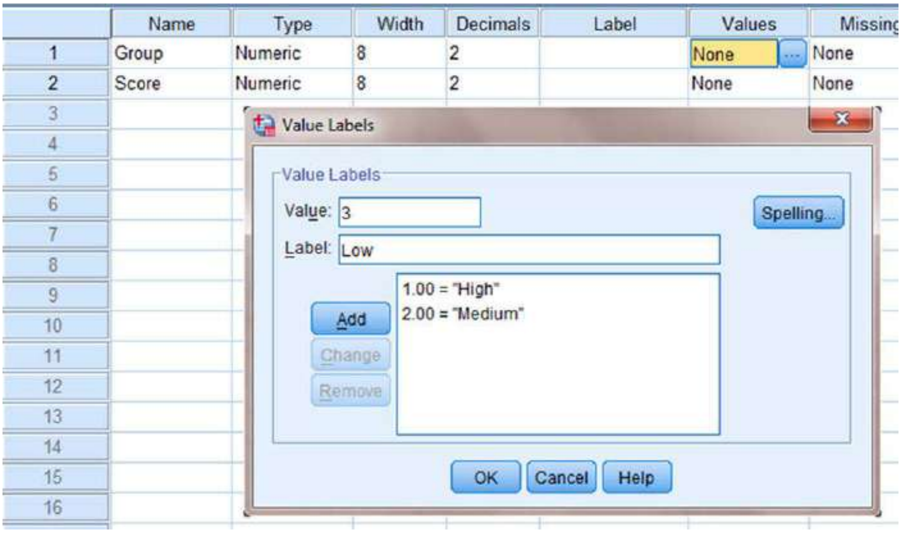


FIGURE 6.2

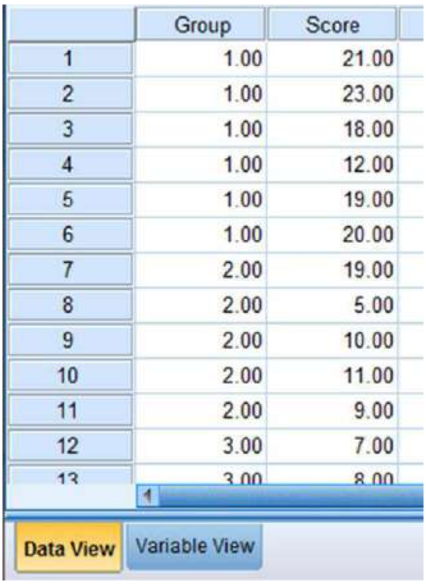


FIGURE 6.3

6.3.2.3 Analyze Your Data As shown in Figure 6.4, use the pull-down menus to choose “Analyze,” “Nonparametric Tests,” “Legacy Dialogs,” and “K Independent Samples. . . .”

Use the top arrow button to place your variable with your data values, or dependent variable (DV), in the box labeled “Test Variable List:.” Then, use the lower arrow button to place your grouping variable, or independent variable (IV), in the box labeled “Grouping Variable.” As shown in Figure 6.5, we have placed the

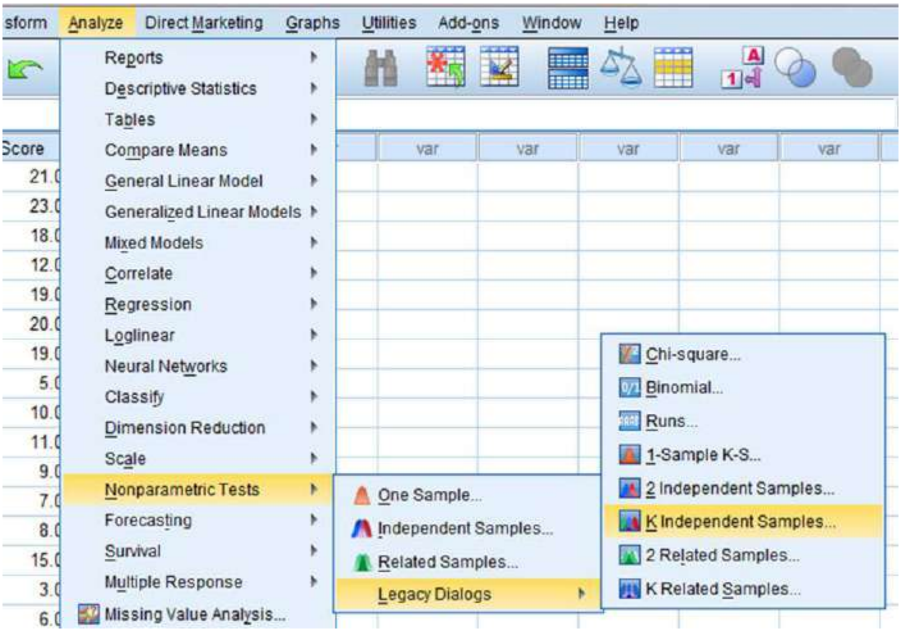


FIGURE 6.4

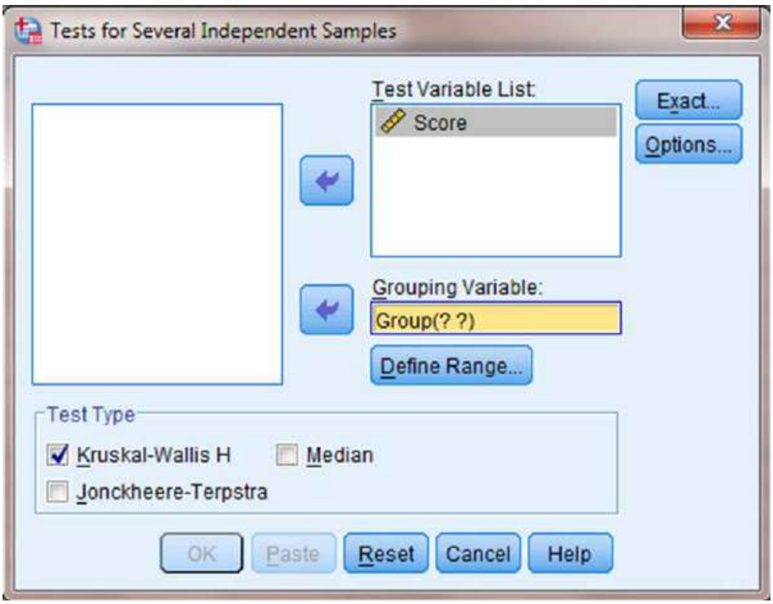


FIGURE 6.5

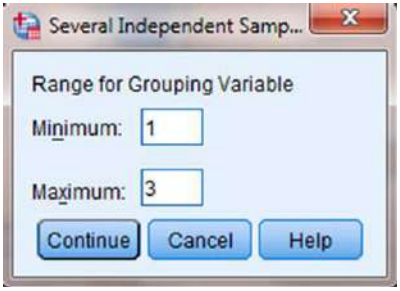


FIGURE 6.6

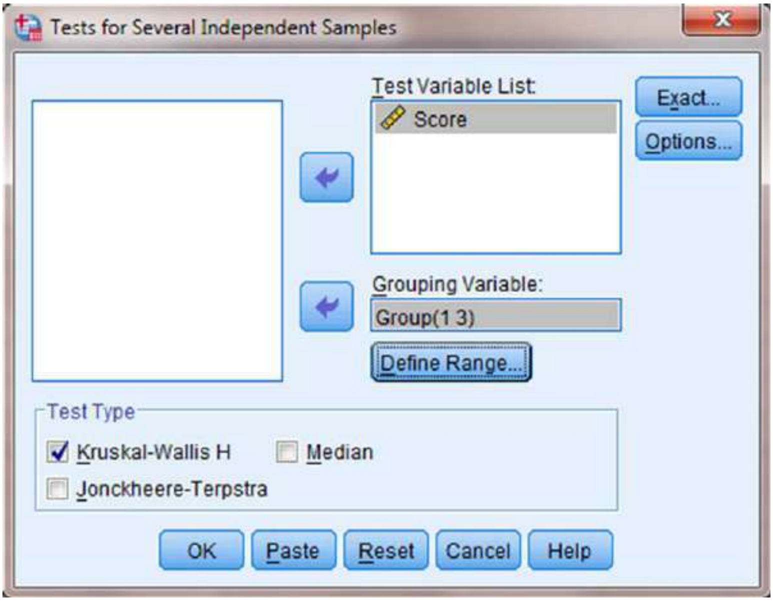


FIGURE 6.7

"Score" variable in the "Test Variable List" and the "Group" variable in the "Grouping Variable" box.

Click on the "Define Range . . ." button to assign a reference value to your independent variable (i.e., "Grouping Variable").

As shown in Figure 6.6, type 1 into the box next to "Minimum" and 3 in the box next to "Maximum." Then, click "Continue." This step references the value labels you defined when you established your grouping variable.

Now that the groups have been assigned (see Fig. 6.7), click "OK" to perform the analysis.

6.3.2.4 Interpret the Results from the SPSS Output Window SPSS Output 6.1 provides the mean ranks of groups and group sizes. The second output table provides the Kruskal–Wallis H-test statistic ($H = 9.944$). Since this test uses a χ^2

Ranks

	Group	N	Mean Rank
Score	High	6	13.92
	Medium	5	8.10
	Low	6	4.83
	Total	17	

Test Statistics[a,b]

	Score
Chi-Square	9.944
df	2
Asymp. Sig.	.007

a. Kruskal Wallis Test

b. Grouping Variable: Group

SPSS OUTPUT 6.1

distribution, SPSS calls the *H* statistic “Chi-Square.” This table also returns the degrees of freedom ($df = 2$) and the significance ($p = 0.007$).

Based on the results from SPSS, three social interaction groups were compared: high ($n_H = 6$), medium ($n_M = 5$), and low ($n_L = 6$). The Kruskal–Wallis *H*-test was significant ($H_{(2)} = 9.94$, $p < 0.05$). In order to compare individual pairs of samples, contrasts must be used.

Note that to perform Mann–Whitney *U*-tests for sample contrasts, simply use the grouping values you established when you defined your variables in step 1. Remember to use your corrected level of risk α_B when examining your significance.

6.3.3 Sample Kruskal–Wallis *H*-Test (Large Data Samples)

Researchers were interested in continuing their study of social interaction. In a new study, they examined the self-confidence of teenagers with respect to social interaction. Three levels of social interaction were based on the following characteristics:

High = constant interaction; talks with many different people; initiates discussion

Medium = interacts with a variety of people; some periods of isolation; tends to focus on fewer people

Low = remains mostly isolated from others; speaks if spoken to, but leaves interaction quickly

The researchers assigned each participant into one of the three social interaction groups. Researchers administered a self-assessment of self-confidence. The assess-

TABLE 6.5

Original self-confidence scores placed within social interaction groups

High	Medium	Low
18	35	37
27	47	24
24	11	7
30	31	19
48	12	20
16	39	14
43	11	38
46	14	16
49	40	12
34	48	31
28	32	15
20	9	20
37	44	25
21	30	10
20	33	36
16	26	45
23	22	48
12	3	42
50	41	42
25	17	21
	8	
	10	
	41	

ment instrument measured self-confidence on a 50-point ordinal scale. Table 6.5 shows the scores obtained by each of the participants, with 50 points indicating high self-confidence.

We want to determine if there is a difference between any of the three groups in Table 6.5. The Kruskal–Wallis H-test will be used to analyze the data.

6.3.3.1 State the Null and Research Hypotheses The null hypothesis states that there is no tendency for teen self-confidence to rank systematically higher or lower for any of the levels of social interaction. The research hypothesis states that there is a tendency for teen self-confidence to rank systematically higher or lower for at least one level of social interaction than at least one of the other levels. We generally use the concept of "systematic differences" in the hypotheses.

The null hypothesis is

$$H_O\colon \theta_L = \theta_M = \theta_H$$

The research hypothesis is

> H_A: There is a tendency for teen self-confidence to rank systematically higher or lower for at least one level of social interaction when compared with the other levels.

6.3.3.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

6.3.3.3 Choose the Appropriate Test Statistic The data are obtained from three independent, or unrelated, samples of teenagers. They were assessed using an instrument with a 50-point ordinal scale. Since we are comparing three independent samples of values based on an ordinal scale instrument, we will use the Kruskal–Wallis H-test.

6.3.3.4 Compute the Test Statistic First, combine and rank the three samples together (see Table 6.6).

TABLE 6.6

Original ordinal score	Participant rank	Social interaction group
3	1	Medium
7	2	Low
8	3	Medium
9	4	Medium
10	5.5	Medium
10	5.5	Low
11	7.5	Medium
11	7.5	Medium
12	10	High
12	10	Medium
12	10	Low
14	12.5	Medium
14	12.5	Low
15	14	Low
16	16	High
16	16	High
16	16	Low
17	18	Medium
18	19	High
19	20	Low
20	22.5	High
20	22.5	High

TABLE 6.6 (*Continued*)

Original ordinal score	Participant rank	Social interaction group
20	22.5	Low
20	22.5	Low
21	25.5	High
21	25.5	Low
22	27	Medium
23	28	High
24	29.5	High
24	29.5	Low
25	31.5	High
25	31.5	Low
26	33	Medium
27	34	High
28	35	High
30	36.5	High
30	36.5	Medium
31	38.5	Medium
31	38.5	Low
32	40	Medium
33	41	Medium
34	42	High
35	43	Medium
36	44	Low
37	45.5	High
37	45.5	Low
38	47	Low
39	48	Medium
40	49	Medium
41	50.5	Medium
41	50.5	Medium
42	52.5	Low
42	52.5	Low
43	54	High
44	55	Medium
45	56	Low
46	57	High
47	58	Medium
48	60	High
48	60	Medium
48	60	Low
49	62	High
50	63	High

TABLE 6.7

Ordinal data ranks		
High	Medium	Low
10	1	2
16	3	5.5
16	4	10
19	5.5	12.5
22.5	7.5	14
22.5	7.5	16
25.5	10	20
28	12.5	22.5
29.5	18	22.5
31.5	27	25.5
34	33	29.5
35	36.5	31.5
36.5	38.5	38.5
42	40	44
45.5	41	45.5
54	43	47
57	48	52.5
60	49	52.5
62	50.5	56
63	50.5	60
	55	
	58	
	60	

Place the participant ranks in their social interaction groups to compute the sum of ranks, R_i, for each group (see Table 6.7).

Next, compute the sum of ranks for each social interaction group. The ranks in each group are added to obtain a total R-value for the group.

For the high group, $R_H = 709.5$ and $n_H = 20$.

For the medium group, $R_M = 699$ and $n_M = 23$.

For the low group, $R_L = 607.5$ and $n_L = 20$.

These R-values are used to compute the Kruskal–Wallis H-test statistic (see Formula 6.1). The number of participants in each group is identified by a lowercase n. The total group size in the study is identified by the uppercase N. In this study, $N = 63$.

Now, using the data from Table 6.7, compute the H statistic using Formula 6.1:

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

$$H = \frac{12}{63(63+1)}\left(\frac{709.5^2}{20} + \frac{699^2}{23} + \frac{607.5^2}{20}\right) - 3(63+1)$$
$$= 0.003(25,169.51 + 21,243.52 + 18,452.81) - 192 = 0.003(64,865.85) - 192$$
$$H = 1.053$$

Since there were ties involved in the ranking, correct the value of *H*. First, compute the tie correction (see Formula 6.2). There were 11 sets of ties with two values, three sets of ties with three values, and one set of ties with four values. Then, divide the original *H* statistic by the tie correction C_H:

$$C_H = 1 - \frac{\sum(T^3 - T)}{N^3 - N} = 1 - \frac{11(2^3-2) + 3(3^3-3) + (4^3-4)}{63^3 - 63}$$
$$= 1 - \frac{189}{249984} = 1 - 0.0008$$
$$C_H = 0.9992$$

Next, we divide to find the corrected *H* statistic:

$$\text{corrected } H = \text{original } H \div C_H = 1.053 \div 0.9992$$

For this set of data, notice that the corrected *H* does not differ greatly from the original *H*. With the correction, $H = 1.054$.

6.3.3.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Since the data have at least one large sample, we will use the χ^2 distribution (see Table B.2 found in Appendix B) to find the critical value for the Kruskal–Wallis *H*-test. In this case, we look for the critical value for $df = 2$ and $\alpha = 0.05$. Using the table, the critical value for rejecting the null hypothesis is 5.99.

6.3.3.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 5.99 and the obtained value is $H = 1.054$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value exceeds the obtained value, we do not reject the null hypothesis. Since the critical value exceeds the obtained value, we do not reject the null hypothesis.

6.3.3.7 Interpret the Results We did not reject the null hypothesis, suggesting that no real difference exists between any of the three groups. In particular, the data suggest that there is no difference in self-confidence between one or more of the three social interaction types.

6.3.3.8 Reporting the Results The reporting of results for the Kruskal–Wallis *H*-test should include such information as sample size for each of the groups, the *H* statistic, degrees of freedom, and *p*-value's relation to α. For this example, three social interaction groups were compared. The three social interaction groups were high ($n_H = 20$), medium ($n_M = 23$), and low ($n_L = 20$). The Kruskal–Wallis *H*-test was not significant ($H_{(2)} = 1.054$, $p < 0.05$).

6.4 EXAMPLES FROM THE LITERATURE

To be shown are varied examples of the nonparametric procedures described in this chapter. We have summarized each study's research problem and researchers' rationale(s) for choosing a nonparametric approach. We encourage you to obtain these studies if you are interested in their results.

Gömleksız and Bulut (2007) examined primary school teachers' views on the implementation and effectiveness of a new primary school mathematics curriculum. When they examined the data, some of the samples were found to be non-normal. For those samples, they used a Kruskal–Wallis *H*-test, followed by Mann–Whitney *U*-tests to compare unrelated samples.

In the study of Finson et al. (2006), the students of nine middle school teachers were asked to draw a scientist. Based on the drawings, students' perceptions of scientists were compared with their teachers' teaching styles using the Kruskal–Wallis *H*-test. Then, the samples were individually compared using Mann–Whitney *U*-test. The researchers used nonparametric statistical analyses because only relatively small sample sizes of subjects were available.

Belanger and Desrochers (2001) investigated the nature of infants' ability to perceive event causality. The researchers noted that they chose nonparametric statistical tests because the data samples lacked a normal distribution based on results from a Shapiro–Wilk test. In addition, they stated that the sample sizes were small. A Kruskal–Wallis *H*-test revealed no significant differences between samples. Therefore, they did not perform any sample contrasts.

Plata and Trusty (2005) investigated the willingness of high school boys' willingness to allow same-sex peers with learning disabilities (LDs) to participate in school activities and out-of-school activities. The authors compared the willingness of 38 educationally successful and 33 educationally at-risk boys. The boys were from varying socioeconomic backgrounds. Due to the data's ordinal nature and small sample sizes among some samples, nonparametric statistics were used for the analysis. The Kruskal–Wallis *H*-test was chosen for multiple comparisons. When sample pairs were compared, the researchers performed a *post-hoc* analysis of the differences between mean rank pairs using a multiple comparison technique.

6.5 SUMMARY

More than two samples that are not related may be compared using a nonparametric procedure called the Kruskal–Wallis *H*-test. The parametric equivalent to this test is known as the one-way analysis of variance (ANOVA). When the Kruskal–Wallis *H*-test produces significant results, it does not identify which nor how many sample pairs are significantly different. The Mann–Whitney *U*-test, with a Bonferroni procedure to avoid type I error rate inflation, is a useful method for comparing individual sample pairs.

In this chapter, we described how to perform and interpret a Kruskal–Wallis H-test followed with sample contrasts. We also explained how to perform the procedures using SPSS. Finally, we offered varied examples of these nonparametric statistics from the literature. The next chapter will involve comparing two variables.

6.6 PRACTICE QUESTIONS

1. A researcher conducted a study with $n = 15$ participants to investigate strength gains from exercise. The participants were divided into three groups and given one of three treatments. Participants' strength gains were measured and ranked. The rankings are presented in Table 6.8.

TABLE 6.8

Treatments		
I	II	III
7	13	12
2	1	5
4	7	16
11	8	9
15	3	14

Use a Kruskal–Wallis H-test with $\alpha = 0.05$ to determine if one or more of the groups are significantly different. If a significant difference exists, use a two-tailed Mann–Whitney U-tests or two-sample Kolmogorov–Smirnov tests to identify which groups are significantly different. Use the Bonferroni procedure to limit the type I error rate. Report your findings.

2. A researcher investigated how physical attraction influences the perception among others of a person's effectiveness with difficult tasks. The photographs of 24 people were shown to a focus group. The group was asked to classify the photos into three groups: very attractive, average, and very unattractive. Then, the group ranked the photographs according to their impression of how capable they were of solving difficult problems. Table 6.9 shows the classification and rankings of the people in the photos (1 = *most effective*, 24 = *least effective*).

Use a Kruskal–Wallis H-test with $\alpha = 0.05$ to determine if one or more of the groups are significantly different. If a significant difference exists, use two-tailed Mann–Whitney U-tests to identify which groups are significantly different. Use the Bonferroni procedure to limit the type I error rate. Report your findings.

TABLE 6.9

Very attractive	Average	Very unattractive
1	3	11
2	4	15
5	8	16
6	9	18
7	13	20
10	14	21
12	19	23
17	22	24

6.7 SOLUTIONS TO PRACTICE QUESTIONS

1. The results from the Kruskal–Wallis *H*-test are displayed in SPSS Output 6.2.

Ranks

	Treatment	N	Mean Rank
RankGain	Treatment 1	5	7.40
	Treatment 2	5	6.00
	Treatment 3	5	10.60
	Total	15	

Test Statistics[a,b]

	RankGain
Chi-Square	2.800
df	2
Asymp. Sig.	.247

a. Kruskal Wallis Test

b. Grouping Variable: Treatment

SPSS OUTPUT 6.2

According to the data, the results from the Kruskal–Wallis *H*-test indicated that the three groups are not significantly different ($H_{(2)} = 2.800, p > 0.05$). Therefore, no follow-up contrasts are needed.

2. The results from the Kruskal–Wallis *H*-test are displayed in SPSS Output 6.3. According to the data, the results from the Kruskal–Wallis *H*-test indicated that one or more of the three groups are significantly different ($H_{(2)} = 9.920, p < 0.05$). Therefore, we must examine each set of samples with follow-up contrasts to find the differences between groups.

Based on the significance from the Kruskal–Wallis *H*-test, we compare the samples with Mann–Whitney *U*-tests. Since there are $k = 3$ groups, use

Ranks

	Classification	N	Mean Rank
Ranking	Very Attractive	8	7.50
	Average	8	11.50
	Very Unattractive	8	18.50
	Total	24	

Test Statistics[a,b]

	Ranking
Chi-Square	9.920
df	2
Asymp. Sig.	.007

a. Kruskal Wallis Test

b. Grouping Variable: Classification

SPSS OUTPUT 6.3

Mann-Whitney Test

Ranks

	Classification	N	Mean Rank	Sum of Ranks
Ranking	Very Attractive	8	7.00	56.00
	Average	8	10.00	80.00
	Total	16		

Test Statistics[a]

	Ranking
Mann-Whitney U	20.000
Wilcoxon W	56.000
Z	-1.260
Asymp. Sig. (2-tailed)	.208
Exact Sig. [2*(1-tailed Sig.)]	.234[b]

a. Grouping Variable: Classification

b. Not corrected for ties.

SPSS OUTPUT 6.4

$\alpha_B = 0.0167$ to avoid type I error rate inflation. The results from the Mann–Whitney *U*-tests are displayed in the remaining SPSS Output 6.4, SPSS Output 6.5, and SPSS Output 6.6.

a. *Very Attractive–Attractive Comparison.*

The results from the Mann–Whitney *U*-test ($U = 20.0$, $n_1 = 8$, $n_2 = 8$, $p > 0.0167$) indicated that the two samples were not significantly different.

Mann-Whitney Test

Ranks

	Classification	N	Mean Rank	Sum of Ranks
Ranking	Average	8	6.00	48.00
	Very Unattractive	8	11.00	88.00
	Total	16		

Test Statistics[a]

	Ranking
Mann-Whitney U	12.000
Wilcoxon W	48.000
Z	-2.100
Asymp. Sig. (2-tailed)	.036
Exact Sig. [2*(1-tailed Sig.)]	.038[b]

a. Grouping Variable: Classification

b. Not corrected for ties.

SPSS OUTPUT 6.5

Mann-Whitney Test

Ranks

	Classification	N	Mean Rank	Sum of Ranks
Ranking	Very Attractive	8	5.00	40.00
	Very Unattractive	8	12.00	96.00
	Total	16		

Test Statistics[a]

	Ranking
Mann-Whitney U	4.000
Wilcoxon W	40.000
Z	-2.941
Asymp. Sig. (2-tailed)	.003
Exact Sig. [2*(1-tailed Sig.)]	.002[b]

a. Grouping Variable: Classification

b. Not corrected for ties.

SPSS OUTPUT 6.6

b. *Attractive–Very Unattractive Comparison.*
The results from the Mann–Whitney U-test ($U = 12.0$, $n_1 = 8$, $n_2 = 8$, $p > 0.0167$) indicated that the two samples were not significantly different.

c. *Very Attractive–Very Unattractive Comparison.*
The results from the Mann–Whitney U-test ($U = 4.0$, $n_1 = 8$, $n_2 = 8$, $p < 0.0167$) indicated that the two samples were significantly different.

CHAPTER 7

COMPARING VARIABLES OF ORDINAL OR DICHOTOMOUS SCALES: SPEARMAN RANK-ORDER, POINT-BISERIAL, AND BISERIAL CORRELATIONS

7.1 OBJECTIVES

In this chapter, you will learn the following items:

- How to compute the Spearman rank-order correlation coefficient.
- How to perform the Spearman rank-order correlation using SPSS®.
- How to compute the point-biserial correlation coefficient.
- How to perform the point-biserial correlation using SPSS.
- How to compute the biserial correlation coefficient.

7.2 INTRODUCTION

The statistical procedures in this chapter are quite different from those in the last several chapters. Unlike this chapter, we had compared samples of data. This chapter, however, examines the relationship between two variables. In other words, this chapter will address how one variable changes with respect to another.

The relationship between two variables can be compared with a correlation analysis. If any of the variables are ordinal or dichotomous, we can use a nonparametric correlation. The Spearman rank-order correlation, also called the Spearman's ρ, is used to compare the relationship between ordinal, or rank-ordered, variables. The point-biserial and biserial correlations are used to compare the relationship between two variables if one of the variables is dichotomous. The parametric equivalent to these correlations is the Pearson product-moment correlation.

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

In this chapter, we will describe how to perform and interpret a Spearman rank-order, point-biserial, and biserial correlations. We will also explain how to perform the procedures using SPSS. Finally, we offer varied examples of these nonparametric statistics from the literature.

7.3 THE CORRELATION COEFFICIENT

When comparing two variables, we use an obtained value called a correlation coefficient. A population's correlation coefficient is represented by the Greek letter rho, ρ. A sample's correlation coefficient is represented by the letter r.

We will describe two types of relationships between variables. A direct relationship is a positive correlation with an obtained value ranging from 0 to 1.0. As one variable increases, the other variable also increases. An indirect, or inverse, relationship is a negative correlation with an obtained value ranging from 0 to −1.0. In this case, one variable increases as the other variable decreases.

In general, a significant correlation coefficient also communicates the relative strength of a relationship between the two variables. A value close to 1.0 or −1.0 indicates a nearly perfect relationship, while a value close to 0 indicates an especially weak or trivial relationship. Cohen (1988, 1992) presented a more detailed description of a correlation coefficient's relative strength. Table 7.1 summarizes his findings.

TABLE 7.1

Correlation coefficient for a direct relationship	Correlation coefficient for an indirect relationship	Relationship strength of the variables
0.0	0.0	None/trivial
0.1	−0.1	Weak/small
0.3	−0.3	Moderate/medium
0.5	−0.5	Strong/large
1.0	−1.0	Perfect

There are three important caveats to consider when assigning relative strength to correlation coefficients, however. First, Cohen's work was largely based on behavioral science research. Therefore, these values may be inappropriate in fields such as engineering or the natural sciences. Second, the correlation strength assignments vary for different types of statistical tests. Third, r-values are not based on a linear scale. For example, $r = 0.6$ is not twice as strong as $r = 0.3$.

7.4 COMPUTING THE SPEARMAN RANK-ORDER CORRELATION COEFFICIENT

The Spearman rank-order correlation is a statistical procedure that is designed to measure the relationship between two variables on an ordinal scale of measurement if the sample size is $n \geq 4$. Use Formula 7.1 to determine a Spearman rank-order correlation coefficient r_s if none of the ranked values are ties. Sometimes, the symbol r_s is represented by the Greek symbol rho, or ρ:

$$r_s = 1 - \frac{6\sum D_i^2}{n(n^2 - 1)} \tag{7.1}$$

where n is the number of rank pairs and D_i is the difference between a ranked pair.

If ties are present in the values, use Formula 7.2, Formula 7.3, and Formula 7.4 to determine r_s:

$$r_s = \frac{(n^3 - n) - 6\sum D_i^2 - (T_x + T_y)/2}{\sqrt{(n^3 - n)^2 - (T_x + T_y)(n^3 - n) + T_x T_y}} \tag{7.2}$$

where

$$T_x = \sum_{i=1}^{g} (t_i^3 - t_i) \tag{7.3}$$

and

$$T_y = \sum_{i=1}^{g} (t_i^3 - t_i) \tag{7.4}$$

g is the number of tied groups in that variable and t_i is the number of tied values in a tie group.

If there are no ties in a variable, then $T = 0$.

Use Formula 7.5 to determine the degrees of freedom for the correlation:

$$df = n - 2 \tag{7.5}$$

where n is the number of paired values.

After r_s is determined, it must be examined for significance. Small samples allow one to reference a table of critical values, such as Table B.7 found in Appendix B. However, if the sample size n exceeds those available from the table, then a large sample approximation may be performed. For large samples, compute a z-score and use a table with the normal distribution (see Table B.1 in Appendix B) to obtain a critical region of z-scores. Formula 7.6 may be used to find the z-score of a correlation coefficient for large samples:

$$z^* = r\left[\sqrt{n-1}\right] \tag{7.6}$$

where n is the number of paired values and r is the correlation coefficient.

Note that the method for determining a z-score given a correlation coefficient and examining it for significance is the same for each type of correlation. We will illustrate a large sample approximation with a sample problem when we address the point-biserial correlation.

Although we will use Formula 7.6 to determine the significance of the correlation coefficient, some statisticians recommend using the formula based on the Student's t-distribution, as shown in Formula 7.7:

$$t = r_s \sqrt{\frac{n-2}{1-r_s^2}} \tag{7.7}$$

According to Siegel and Castellan (1988), the advantage of using the Student's *t*-distribution over the *z*-score is small with larger sample sizes *n*.

7.4.1 Sample Spearman Rank-Order Correlation (Small Data Samples without Ties)

Eight men were involved in a study to examine the resting heart rate regarding frequency of visits to the gym. The assumption is that the person who visits the gym more frequently for a workout will have a slower heart rate. Table 7.2 shows the number of visits each participant made to the gym during the month the study was conducted. It also provides the mean heart rate measured at the end of the week during the final 3 weeks of the month.

TABLE 7.2

Participant	Number of visits	Mean heart rate
1	5	100
2	12	89
3	7	78
4	14	66
5	2	77
6	8	103
7	15	67
8	17	63

The values in this study do not possess characteristics of a strong interval scale. For instance, the number of visits to the gym does not necessarily communicate duration and intensity of physical activity. In addition, heart rate has several factors that can result in differences from one person to another. Ordinal measures offer a clearer relationship to compare these values from one individual to the next. Therefore, we will convert these values to ranks and use a Spearman rank-order correlation.

7.4.1.1 State the Null and Research Hypothesis The null hypothesis states that there is no correlation between number of visits to the gym in a month and mean resting heart rate. The research hypothesis states that there is a correlation between the number of visits to the gym and the mean resting heart rate.

The null hypothesis is

H_O: $\rho_s = 0$

The research hypothesis is

H_A: $\rho_s \neq 0$

7.4.1.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

7.4.1.3 Choose the Appropriate Test Statistic As stated earlier, we decided to analyze the variables using an ordinal, or rank, procedure. Therefore, we will convert the values in each variable to ordinal data. In addition, we will be comparing the two variables, the number of visits to the gym in a month and the mean resting heart rate. Since we are comparing two variables in which one or both are measured on an ordinal scale, we will use the Spearman rank-order correlation.

7.4.1.4 Compute the Test Statistic First, rank the scores for each variable separately as shown in Table 7.3. Rank them from the lowest score to the highest score to form an ordinal distribution for each variable.

TABLE 7.3

	Original scores		Ranked scores	
Participant	Number of visits	Mean heart rate	Number of visits	Mean heart rate
1	5	100	2	7
2	12	89	5	6
3	7	78	3	5
4	14	66	6	2
5	2	77	1	4
6	8	103	4	8
7	15	67	7	3
8	17	63	8	1

To calculate the Spearman rank-order correlation coefficient, we need to calculate the differences between rank pairs and their subsequent squares where D = rank (mean heart rate) − rank (number of visits). It is helpful to organize the data to manage the summation in the formula (see Table 7.4).

Next, compute the Spearman rank-order correlation coefficient:

$$
\begin{aligned}
r_s &= 1 - \frac{6\sum D_i^2}{n(n^2-1)} \\
&= 1 - \frac{6(136)}{8(8^2-1)} = 1 - \frac{816}{8(64-1)} \\
&= 1 - \frac{816}{8(63)} = 1 - \frac{816}{504} \\
&= 1 - 1.619 \\
r_s &= -0.619
\end{aligned}
$$

TABLE 7.4

Ranked scores		Rank differences	
Number of visits	Mean heart rate	D	D^2
2	7	5	25
5	6	1	1
3	5	2	4
6	2	−4	16
1	4	3	9
4	8	4	16
7	3	−4	16
8	1	−7	49
			$\sum D_i^2 = 136$

7.4.1.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Table B.7 in Appendix B lists critical values for the Spearman rank-order correlation coefficient. In this study, the critical value is found for $n = 8$ and $df = 6$. Since we are conducting a two-tailed test and $\alpha = 0.05$, the critical value is 0.738. If the obtained value exceeds or is equal to the critical value, 0.738, we will reject the null hypothesis. If the critical value exceeds the absolute value of the obtained value, we will not reject the null hypothesis.

7.4.1.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 0.738 and the obtained value is $|r_s| = 0.619$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value is greater than the obtained value, we must not reject the null hypothesis. Since the critical value exceeds the absolute value of the obtained value, we do not reject the null hypothesis.

7.4.1.7 Interpret the Results We did not reject the null hypothesis, suggesting that there is no significant correlation between the number of visits the males made to the gym in a month and their mean resting heart rates.

7.4.1.8 Reporting the Results The reporting of results for the Spearman rank-order correlation should include such information as the number of participants (n), two variables that are being correlated, correlation coefficient (r_s), degrees of freedom (df), and p-value's relation to α.

For this example, eight men ($n = 8$) were observed for 1 month. Their number of visits to the gym was documented (variable 1) and their mean resting heart rate was recorded during the last 3 weeks of the month (variable 2). These data were put in ordinal form for purposes of the analysis. The Spearman rank-order correlation coefficient was not significant ($r_{s(6)} = -0.619, p > 0.05$). Based on this data, we can state that there is no clear relationship between adult male resting heart rate and the frequency of visits to the gym.

7.4.2 Sample Spearman Rank-Order Correlation (Small Data Samples with Ties)

The researcher repeated the experiment in the previous example using females. Table 7.5 shows the number of visits each participant made to the gym during the month of the study and their subsequent mean heart rates.

TABLE 7.5

Participant	Number of visits	Mean heart rate
1	5	96
2	12	63
3	7	78
4	14	66
5	3	79
6	8	95
7	15	67
8	12	64
9	2	99
10	16	62
11	12	65
12	7	76
13	17	61

As with the previous example, the values in this study do not possess characteristics of a strong interval scale, so we will use ordinal measures. We will convert these values to ranks and use a Spearman rank-order correlation.

Steps 1–3 are the same as the previous example. Therefore, we will begin with step 4.

7.4.2.1 Compute the Test Statistic First, rank the scores for each variable as shown in Table 7.6. Rank the scores from the lowest score to the highest score to form an ordinal distribution for each variable.

To calculate the Spearman rank-order correlation coefficient, we need to calculate the differences between rank pairs and their subsequent squares where $D =$ rank (mean heart rate) $-$ rank (number of visits). It is helpful to organize the data to manage the summation in the formula (see Table 7.7).

Next, compute the Spearman rank-order correlation coefficient. Since there are ties present in the ranks, we will use formulas that account for the ties. First, use Formula 7.3 and Formula 7.4. For the number of visits, there are two groups of ties.

TABLE 7.6

	Original scores		Rank scores	
Participant	Number of visits	Mean heart rate	Number of visits	Mean heart rate
1	5	96	3	12
2	12	63	8	3
3	7	78	4.5	9
4	14	66	10	6
5	3	79	2	10
6	8	95	6	11
7	15	67	11	7
8	12	64	8	4
9	2	99	1	13
10	16	62	12	2
11	12	65	8	5
12	7	76	4.5	8
13	17	61	13	1

TABLE 7.7

	Rank scores		Rank differences	
Participant	Number of visits	Mean heart rate	D	D^2
1	3	12	9	81
2	8	3	−5	25
3	4.5	9	4.5	20.25
4	10	6	−4	16
5	2	10	8	64
6	6	11	5	25
7	11	7	−4	16
8	8	4	−4	16
9	1	13	12	144
10	12	2	−10	100
11	8	5	−3	9
12	4.5	8	3.5	12.25
13	13	1	−12	144
				$\sum D_i^2 = 672.5$

The first group has two tied values (rank = 4.5 and $t = 2$) and the second group has three tied values (rank = 8 and $t = 3$):

$$
\begin{aligned}
T_x &= \sum_{i=1}^{g}(t_i^3 - t_i) \\
&= (2^3 - 2) + (3^3 - 3) = (8 - 2) + (27 - 3) \\
&= 6 + 24 \\
T_x &= 30
\end{aligned}
$$

For the mean resting heart rate, there are no ties. Therefore, $T_y = 0$. Now, calculate the Spearman rank-order correlation coefficient using Formula 7.2:

$$
\begin{aligned}
r_s &= \frac{(n^3 - n) - 6\sum D_i^2 - (T_x + T_y)/2}{\sqrt{(n^3 - n)^2 - (T_x + T_y)(n^3 - n) + T_x T_y}} \\
&= \frac{(13^3 - 13) - 6(672.5) - (30 + 0)/2}{\sqrt{(13^3 - 13)^2 - (30 + 0)(13^3 - 13) + (30)(0)}} \\
&= \frac{2184 - 4035 - 15}{\sqrt{(2184)^2 - (30)(2184) + 0}} \\
&= \frac{-1866}{\sqrt{4,704,336}} = \frac{-1866}{2169} \\
r_s &= -0.860
\end{aligned}
$$

7.4.2.2 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Table B.7 in Appendix B lists critical values for the Spearman rank-order correlation coefficient. To be significant, the absolute value of the obtained value, $|r_s|$, must be greater than or equal to the critical value on the table. In this study, the critical value is found for $n = 13$ and $df = 11$. Since we are conducting a two-tailed test and $\alpha = 0.05$, the critical value is 0.560.

7.4.2.3 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 0.560 and the obtained value is $|r_s| = 0.860$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value is greater than the obtained value, we must not reject the null hypothesis. Since the critical value is less than the absolute value of the obtained value, we reject the null hypothesis.

7.4.2.4 Interpret the Results We rejected the null hypothesis, suggesting that there is a significant correlation between the number of visits the females made to the gym in a month and their mean resting heart rates.

7.4.2.5 Reporting the Results The reporting of results for the Spearman rank-order correlation should include such information as the number of participants (n), two variables that are being correlated, correlation coefficient (r_s), degrees of freedom (df), and p-value's relation to α.

For this example, 13 women ($n = 13$) were observed for 1 month. Their number of visits to the gym was documented (variable 1) and their mean resting heart rate was recorded during the last 3 weeks of the month (variable 2). These data were put in ordinal form for purposes of the analysis. The Spearman rank-order correlation coefficient was significant ($r_{s(11)} = -0.860$, $p < 0.05$). Based on this data, we can state that there is a very strong inverse relationship between adult female resting heart rate and the frequency of visits to the gym.

7.4.3 Performing the Spearman Rank-Order Correlation Using SPSS

We will analyze the data from the previous example using SPSS.

7.4.3.1 Define Your Variables First, click the "Variable View" tab at the bottom of your screen. Then, type the names of your variables in the "Name" column. As shown in Figure 7.1, the first variable is called "Number_of_Visits" and the second variable is called "Mean_Heart_Rate."

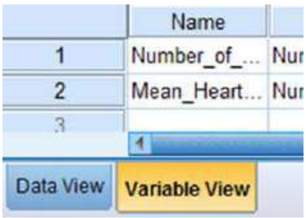
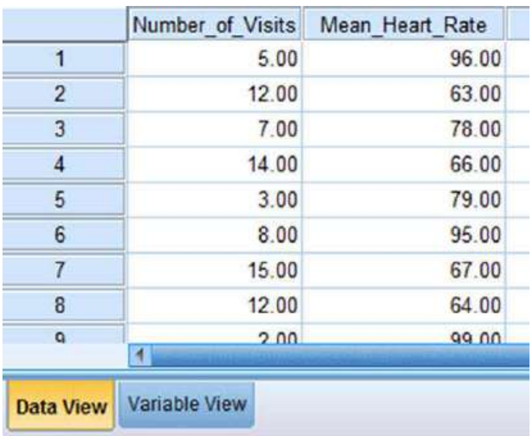


FIGURE 7.1

7.4.3.2 Type in Your Values Click the "Data View" tab at the bottom of your screen as shown in Figure 7.2. Type the values in the respective columns.



	Number_of_Visits	Mean_Heart_Rate
1	5.00	96.00
2	12.00	63.00
3	7.00	78.00
4	14.00	66.00
5	3.00	79.00
6	8.00	95.00
7	15.00	67.00
8	12.00	64.00
9	2.00	99.00

FIGURE 7.2

7.4.3.3 Analyze Your Data As shown in Figure 7.3, use the pull-down menus to choose "Analyze," "Correlate," and "Bivariate. . . ."

Use the arrow button to place both variables with your data values in the box labeled "Variables:" as shown in Figure 7.4. Then, in the "Correlation Coefficients" box, uncheck "Pearson" and check "Spearman." Finally, click "OK" to perform the analysis.

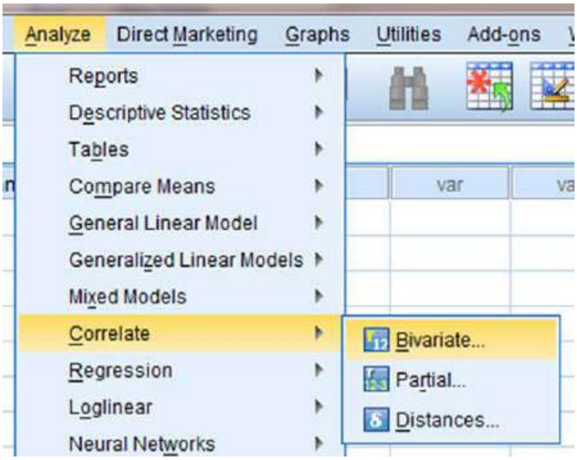


FIGURE 7.3

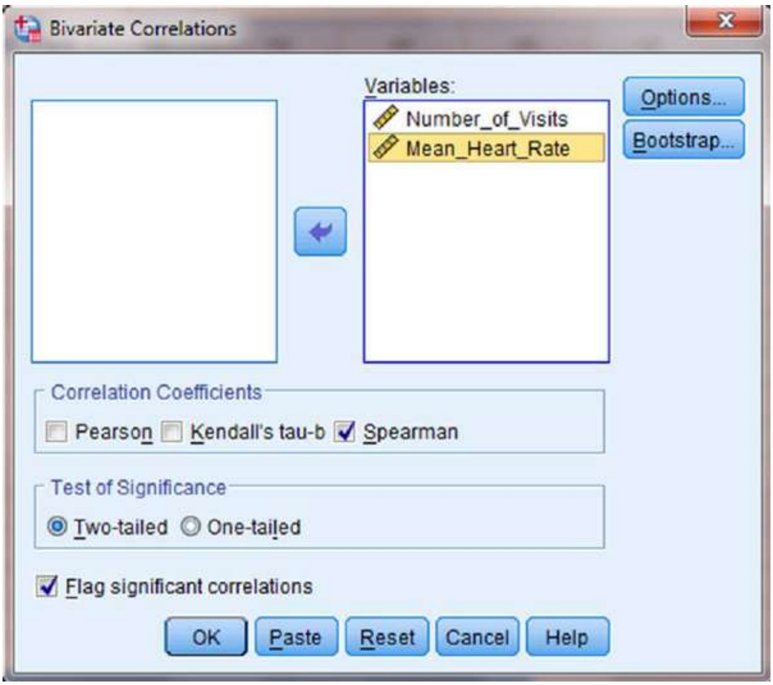


FIGURE 7.4

7.4.3.4 Interpret the Results from the SPSS Output Window The output table (see SPSS Output 7.1) provides the Spearman rank-order correlation coefficient ($r_s = -0.860$) labeled Spearman's rho. It also returns the number of pairs ($n = 13$) and the two-tailed significance ($p \approx 0.000$). In this example, the significance is not actually zero. The reported value does not return enough digits to show the significance's actual precision.

Correlations

			Number_of_Visits	Mean_Heart_Rate
Spearman's rho	Number_of_Visits	Correlation Coefficient	1.000	-.860**
		Sig. (2-tailed)	.	.000
		N	13	13
	Mean_Heart_Rate	Correlation Coefficient	-.860**	1.000
		Sig. (2-tailed)	.000	.
		N	13	13

**. Correlation is significant at the 0.01 level (2-tailed).

SPSS OUTPUT 7.1

Based on the results from SPSS, the Spearman rank-order correlation coefficient was significant ($r_{s(11)} = -0.860$, $p < 0.05$). Based on these data, we can state that there is a very strong inverse relationship between adult female resting heart rate and the frequency of visits to the gym.

7.5 COMPUTING THE POINT-BISERIAL AND BISERIAL CORRELATION COEFFICIENTS

The point-biserial and biserial correlations are statistical procedures for use with dichotomous variables. A dichotomous variable is simply a measure of two conditions. A dichotomous variable is either discrete or continuous. A discrete dichotomous variable has no particular order and might include such examples as gender (male vs. female) or a coin toss (heads vs. tails). A continuous dichotomous variable has some type of order to the two conditions and might include measurements such as pass/fail or young/old. Finally, since the point-biserial and biserial correlations each involves an interval scale analysis, they are special cases of the Pearson product-moment correlation.

7.5.1 Correlation of a Dichotomous Variable and an Interval Scale Variable

The point-biserial correlation is a statistical procedure to measure the relationship between a discrete dichotomous variable and an interval scale variable. Use Formula 7.8 to determine the point-biserial correlation coefficient r_{pb}:

$$r_{pb} = \frac{\overline{x}_p - \overline{x}_q}{s}\sqrt{P_p P_q} \tag{7.8}$$

where $\overline{x}_p$ is the mean of the interval variable's values associated with the dichotomous variable's first category, $\overline{x}_q$ is the mean of the interval variable's values associated with the dichotomous variable's second category, s is the standard deviation of the variable on the interval scale, P_p is the proportion of the interval variable values associated with the dichotomous variable's first category, and P_q is the proportion of the interval variable values associated with the dichotomous variable's second category.

Recall the formulas for mean (Formula 7.9) and standard deviation (Formula 7.10):

$$\overline{x}_q = \sum x_i \div n \tag{7.9}$$

and

$$s = \sqrt{\frac{\sum (x_i - \overline{x})^2}{n-1}} \tag{7.10}$$

where $\sum x_j$ is the sum of the values in the sample and n is the number of values in the sample.

The biserial correlation is a statistical procedure to measure the relationship between a continuous dichotomous variable and an interval scale variable. Use Formula 7.11 to determine the biserial correlation coefficient r_b:

$$r_b = \left[\frac{\overline{x}_p - \overline{x}_q}{s_x}\right]\frac{P_p P_q}{y} \tag{7.11}$$

where $\overline{x}_p$ is the mean of the interval variable's values associated with the dichotomous variable's first category, $\overline{x}_q$ is the mean of the interval variable's values associated with the dichotomous variable's second category, s_x is the standard deviation of the variable on the interval scale, P_p is the proportion of the interval variable values associated with the dichotomous variable's first category, P_q is the proportion of the interval variable values associated with the dichotomous variable's second category, and y is the height of the unit normal curve ordinate at the point dividing P_p and P_q (see Fig. 7.5).

You may use Table B.1 in Appendix B or Formula 7.12 to find the height of the unit normal curve ordinate, y:

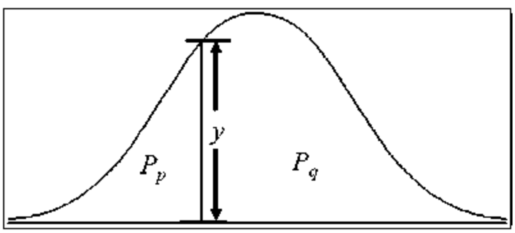


FIGURE 7.5

$$y = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \tag{7.12}$$

where e is the natural log base and approximately equal to 2.718282 and z is the z-score at the point dividing P_p and P_q.

Formula 7.13 is the relationship between the point-biserial and the biserial correlation coefficients. This formula is necessary to find the biserial correlation coefficient because SPSS only determines the point-biserial correlation coefficient:

$$r_b = r_{bp} \frac{\sqrt{P_p P_q}}{y} \tag{7.13}$$

After the correlation coefficient is determined, it must be examined for significance. Small samples allow one to reference a table of critical values, such as Table B.8 found in Appendix B. However, if the sample size n exceeds those available from the table, then a large sample approximation may be performed. For large samples, compute a z-score and use a table with the normal distribution (see Table B.1 in Appendix B) to obtain a critical region of z-scores. As described earlier in this chapter, Formula 7.6 may be used to find the z-score of a correlation coefficient for large samples.

7.5.2 Correlation of a Dichotomous Variable and a Rank-Order Variable

As explained earlier, the point-biserial and biserial correlation procedures earlier involve a dichotomous variable and an interval scale variable. If the correlation was a dichotomous variable and a rank-order variable, a slightly different approach is needed.

To find the point-biserial correlation coefficient for a discrete dichotomous variable and a rank-order variable, simply use the Spearman rank-order described earlier and assign arbitrary values to the dichotomous variable such as 0 and 1. To find the biserial correlation coefficient for a continuous dichotomous variable and a rank-order variable, use the same procedure and then apply Formula 7.13 given earlier.

7.5.3 Sample Point-Biserial Correlation (Small Data Samples)

A researcher in a psychological lab investigated gender differences. She wished to compare male and female ability to recognize and remember visual details. She used 17 participants (8 males and 9 females) who were initially unaware of the actual experiment. First, she placed each one of them alone in a room with various objects and asked them to wait. After 10 min, she asked each of the participants to complete a 30 question posttest relating to several details in the room. Table 7.8 shows the participants' genders and posttest scores.

The researcher wishes to determine if a relationship exists between the two variables and the relative strength of the relationship. Gender is a discrete dichotomous

TABLE 7.8

Participant	Gender	Posttest score
1	M	7
2	M	19
3	M	8
4	M	10
5	M	7
6	M	15
7	M	6
8	M	13
9	F	14
10	F	11
11	F	18
12	F	23
13	F	17
14	F	20
15	F	14
16	F	24
17	F	22

variable and visual detail recognition is an interval scale variable. Therefore, we will use a point-biserial correlation.

7.5.3.1 State the Null and Research Hypothesis The null hypothesis states that there is no correlation between gender and visual detail recognition. The research hypothesis states that there is a correlation between gender and visual detail recognition.

The null hypothesis is

$$H_O\colon \rho_{pb} = 0$$

The research hypothesis is

$$H_A\colon \rho_{pb} \neq 0$$

7.5.3.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

7.5.3.3 Choose the Appropriate Test Statistic As stated earlier, we decided to analyze the relationship between the two variables. A correlation will provide the relative strength of the relationship between the two variables. Gender is a discrete dichotomous variable and visual detail recognition is an interval scale variable. Therefore, we will use a point-biserial correlation.

7.5.3.4 Compute the Test Statistic First, compute the standard deviation of all values from the interval scale data. It is helpful to organize the data as shown in Table 7.9.

TABLE 7.9

Participant	Gender	Posttest score	$x_i - x$	$(x_i - \bar{x})^2$
1	M	7	−7.59	57.58
2	M	19	4.41	19.46
3	M	8	−6.59	43.40
4	M	10	−4.59	21.05
5	M	7	−7.59	57.58
6	M	15	0.41	0.17
7	M	6	−8.59	73.76
8	M	13	−1.59	2.52
9	F	14	−0.59	0.35
10	F	11	−3.59	12.88
11	F	18	3.41	11.64
12	F	23	8.41	70.76
13	F	17	2.41	5.82
14	F	20	5.41	29.29
15	F	14	−0.59	0.35
16	F	24	9.41	88.58
17	F	22	7.41	54.93
		$\sum x_i = 248$		$\sum(x_i - \bar{x})^2 = 550.12$

Using the summations from Table 7.9, calculate the mean and the standard deviation for the interval data:

$$\bar{x} = \sum x_i \div n$$

$$\bar{x} = 248 \div 17$$

$$\bar{x} = 14.59$$

$$s_x = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{550.12}{17-1}} = \sqrt{34.38}$$

$$s_x = 5.86$$

Next, compute the means and proportions of the values associated with each item from the dichotomous variable. The mean males' posttest score was

$$\bar{x}_p = \sum x_p \div n_M$$

$$= (7 + 19 + 8 + 10 + 7 + 15 + 6 + 13) \div 8$$

$$\bar{x}_p = 10.63$$

The mean females' posttest score was

$$\begin{aligned}\overline{x}_q &= \sum x_q \div n_F \\ &= (14+11+18+23+17+20+14+24+22) \div 9 \\ \overline{x}_q &= 18.11\end{aligned}$$

The males' proportion was

$$\begin{aligned}P_p &= n_M \div n \\ &= 8 \div 17 \\ P_p &= 0.47\end{aligned}$$

The females' proportion was

$$\begin{aligned}P_q &= n_F \div n \\ &= 9 \div 17 \\ P_q &= 0.53\end{aligned}$$

Now, compute the point-biserial correlation coefficient using the values computed earlier:

$$\begin{aligned}r_{pb} &= \frac{\overline{x}_p - \overline{x}_q}{s_x}\sqrt{P_p P_q} \\ &= \frac{10.63 - 18.11}{5.86}\sqrt{(0.47)(0.53)} \\ &= \frac{-7.49}{5.86}\sqrt{0.25} = (-1.28)(0.50) \\ r_{pb} &= -0.637\end{aligned}$$

The sign on the correlation coefficient is dependent on the order we managed our dichotomous variable. Since that was arbitrary, the sign is irrelevant. Therefore, we use the absolute value of the point-biserial correlation coefficient:

$$r_{pb} = 0.637$$

7.5.3.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Table B.8 in Appendix B lists critical values for the Pearson product-moment correlation coefficient. Using the critical values, table requires that the degrees of freedom be known. Since $df = n - 2$ and $n = 17$, then $df = 17 - 2$. Therefore, $df = 15$. Since we are conducting a two-tailed test and $\alpha = 0.05$, the critical value is 0.482.

7.5.3.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 0.482 and the obtained value is $|r_{pb}| = 0.637$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value is greater than the obtained value, we must not reject the null hypothesis. Since the critical value is less than the absolute value of the obtained value, we reject the null hypothesis.

7.5.3.7 Interpret the Results We rejected the null hypothesis, suggesting that there is a significant and moderately strong correlation between gender and visual detail recognition.

7.5.3.8 Reporting the Results The reporting of results for the point-biserial correlation should include such information as the number of participants (n), two variables that are being correlated, correlation coefficient (r_{pb}), degrees of freedom (df), p-value's relation to α, and the mean values of each dichotomous variable.

For this example, a researcher compared male and female ability to recognize and remember visual details. Eight males ($n_M = 8$) and nine females ($n_F = 9$) participated in the experiment. The researcher measured participants' visual detail recognition with a 30 question test requiring participants to recall details in a room they had occupied. A point-biserial correlation produced significant results ($r_{pb(15)} = 0.637$, $p < 0.05$). These data suggest that there is a strong relationship between gender and visual detail recognition. Moreover, the mean scores on the detail recognition test indicate that males ($\overline{x}_M = 10.63$) recalled fewer details, while females ($\overline{x}_F = 18.11$) recalled more details.

7.5.4 Performing the Point-Biserial Correlation Using SPSS

We will analyze the data from the previous example using SPSS.

7.5.4.1 Define Your Variables First, click the "Variable View" tab at the bottom of your screen. Then, type the names of your variables in the "Name" column. As shown in Figure 7.6, the first variable is called "Gender" and the second variable is called "Posttest_Score."

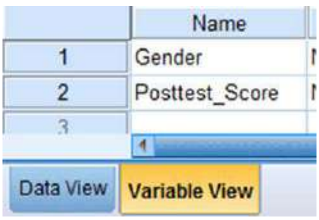


FIGURE 7.6

7.5.4.2 Type in Your Values Click the "Data View" tab at the bottom of your screen as shown in Figure 7.7. Type in the values in the respective columns. Gender is a discrete dichotomous variable and SPSS needs a code to reference the values. We code male values with 0 and female values with 1. Any two values can be chosen for coding the data.

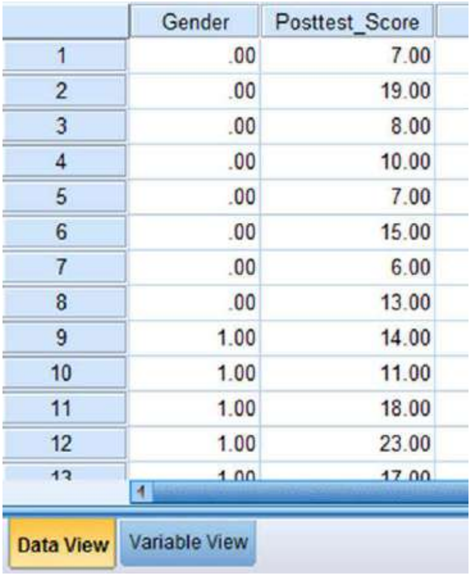


FIGURE 7.7

7.5.4.3 Analyze Your Data As shown in Figure 7.8, use the pull-down menus to choose "Analyze," "Correlate," and "Bivariate. . . ."

Use the arrow button near the middle of the window to place both variables with your data values in the box labeled "Variables:" as shown in Figure 7.9. In the "Correlation Coefficients" box, "Pearson" should remain checked since the Pearson product-moment correlation will perform an approximate point-biserial correlation. Finally, click "OK" to perform the analysis.

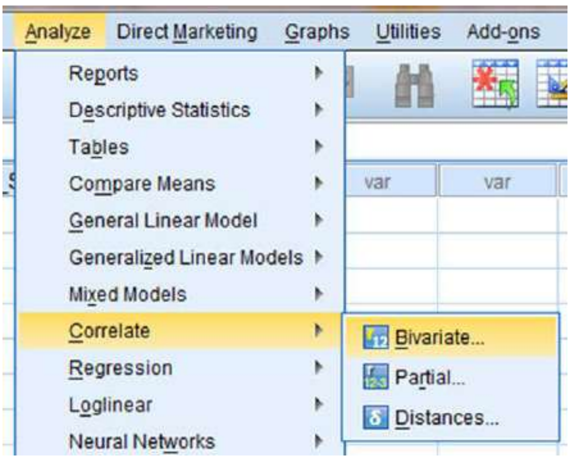


FIGURE 7.8

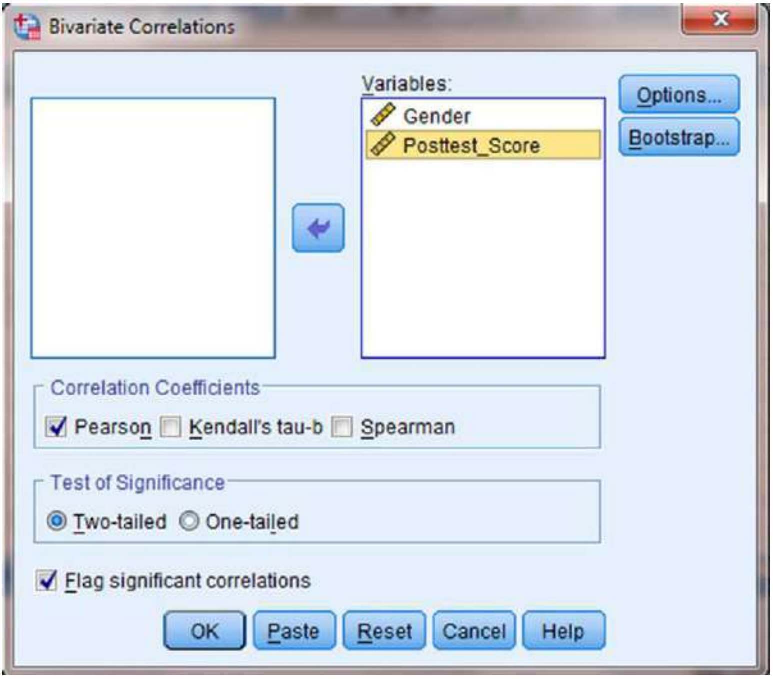


FIGURE 7.9

7.5.4.4 Interpret the Results from the SPSS Output Window The output table (see SPSS Output 7.2) provides the Pearson product-moment correlation coefficient ($r = 0.657$). This correlation coefficient is approximately equal to the point-biserial correlation coefficient. It also returns the number of pairs ($n = 17$) and the two-tailed significance ($p = 0.004$).

Correlations

		Gender	Posttest_Score
Gender	Pearson Correlation	1	.657**
	Sig. (2-tailed)		.004
	N	17	17
Posttest_Score	Pearson Correlation	.657**	1
	Sig. (2-tailed)	.004	
	N	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

SPSS OUTPUT 7.2

Based on the results from SPSS, the point-biserial correlation coefficient was significant ($r_{pb(15)} = 0.657$, $p < 0.05$). Based on these data, we can state that there is a strong relationship between gender and visual detail recognition (as measured by the posttest).

7.5.5 Sample Point-Biserial Correlation (Large Data Samples)

A colleague of the researcher from the previous example wished to replicate the study investigating gender differences. As before, he compared male and female ability to recognize and remember visual details. He used 26 participants (14 males and 12 females) who were initially unaware of the actual experiment. Table 7.10 shows the participants' genders and posttest scores.

TABLE 7.10

Participant	Gender	Posttest score
1	M	6
2	M	15
3	M	8
4	M	10
5	M	6
6	M	12
7	M	7
8	M	13
9	M	13
10	M	10
11	M	18
12	M	23
13	M	17
14	M	20
15	F	14
16	F	26
17	F	14
18	F	11
19	F	29
20	F	20
21	F	15
22	F	18
23	F	9
24	F	14
25	F	21
26	F	22

We will once again use a point-biserial correlation. However, we will use a large sample approximation to examine the results for significance since the sample size is large.

7.5.5.1 State the Null and Research Hypothesis The null hypothesis states that there is no correlation between gender and visual detail recognition. The research hypothesis states that there is a correlation between gender and visual detail recognition.

The null hypothesis is

H_O: $\rho_{pb} = 0$

The research hypothesis is

H_A: $\rho_{pb} \neq 0$

7.5.5.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

7.5.5.3 Choose the Appropriate Test Statistic As stated earlier, we decided to analyze the relationship between the two variables. A correlation will provide the relative strength of the relationship between the two variables. Gender is a discrete dichotomous variable and visual detail recognition is an interval scale variable. Therefore, we will use a point-biserial correlation.

7.5.5.4 Compute the Test Statistic First, compute the standard deviation of all values from the interval scale data. Organize the data to manage the summations (see Table 7.11):

$$\bar{x} = \sum x_i \div n$$

$$\bar{x} = 391 \div 26$$

$$\bar{x} = 15.04$$

$$s_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{934.96}{26-1}} = \sqrt{37.40}$$

$$S_x = 6.115$$

Next, compute the means and proportions of the values associated with each item from the dichotomous variable. The mean males' posttest score was

$$\bar{x}_p = \sum x_p \div n_M$$
$$= (6+15+8+10+6+12+7+13+13+10+18+23+17+20) \div 14$$
$$\bar{x}_p = 12.71$$

The mean females' posttest score was

$$\bar{x}_q = \sum x_q \div n_F$$
$$= (14+26+14+11+29+20+15+18+9+14+21+22) \div 12$$
$$\bar{x}_q = 17.75$$

The males' proportion was

TABLE 7.11

Participant	Gender	Posttest score	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
1	M	6	−9.04	81.69
2	M	15	−0.04	0.00
3	M	8	−7.04	49.54
4	M	10	−5.04	25.39
5	M	6	−9.04	81.69
6	M	12	−3.04	9.23
7	M	7	−8.04	64.62
8	M	13	−2.04	4.16
9	M	13	−2.04	4.16
10	M	10	−5.04	25.39
11	M	18	2.96	8.77
12	M	23	7.96	63.39
13	M	17	1.96	3.85
14	M	20	4.96	24.62
15	F	14	−1.04	1.08
16	F	26	10.96	120.16
17	F	14	−1.04	1.08
18	F	11	−4.04	16.31
19	F	29	13.96	194.92
20	F	20	4.96	24.62
21	F	15	0.04	0.00
22	F	18	2.96	8.77
23	F	9	−6.04	36.46
24	F	14	−1.04	1.08
25	F	21	5.96	35.54
26	F	22	6.96	48.46
		$\sum x_i = 391$		$\sum(x_i - \overline{x})^2 = 934.96$

$$P_p = n_M \div n$$
$$= 14 \div 26$$
$$P_p = 0.54$$

The females' proportion was

$$P_q = n_F \div n$$
$$= 12 \div 26$$
$$P_q = 0.46$$

Now, compute the point-biserial correlation coefficient using the values computed earlier:

$$r_{pb} = \frac{\overline{x}_p - \overline{x}_q}{s_x}\sqrt{P_pP_q} = \frac{12.71-17.75}{6.115}\sqrt{(0.54)(0.46)}$$
$$= \frac{-5.04}{6.115}\sqrt{0.25} = (-0.823)(0.50)$$
$$r_{pb} = -0.411$$

The sign on the correlation coefficient is dependent on the order we managed our dichotomous variable. Since that was arbitrary, the sign is irrelevant. Therefore, we use the absolute value of the point-biserial correlation coefficient:

$$r_{pb} = 0.411$$

Since our number of values is large, we will use a large sample approximation to examine the obtained value for significance. We will find a z-score for our data using an approximation to the normal distribution:

$$z^* = r_{pb}\left[\sqrt{n-1}\right]$$
$$z^* = 0.411\left[\sqrt{26-1}\right]$$
$$z^* = 2.055$$

7.5.5.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Table B.1 in Appendix B is used to establish the critical region of z-scores. For a two-tailed test with $\alpha = 0.05$, we must not reject the null hypothesis if $-1.96 \leq z^* \leq 1.96$.

7.5.5.6 Compare the Obtained Value with the Critical Value Notice that z^* is in the positive tail of the distribution ($2.055 > 1.96$). Therefore, we reject the null hypothesis. This suggests that the correlation between gender and visual detail recognition is real.

7.5.5.7 Interpret the Results We rejected the null hypothesis, suggesting that there is a significant and moderately weak correlation between gender and visual detail recognition.

7.5.5.8 Reporting the Results The reporting of results for the point-biserial correlation should include such information as the number of participants (n), two variables that are being correlated, correlation coefficient (r_{pb}), degrees of freedom (df), p-value's relation to α, and the mean values of each dichotomous variable.

For this example, a researcher replicated a study that compared male and female ability to recognize and remember visual details. Fourteen males ($n_M = 14$) and 12 females ($n_F = 12$) participated in the experiment. The researcher measured participants' visual detail recognition with a 30 question test requiring participants to recall details in a room they had occupied. A point-biserial correlation produced significant results ($r_{pb(24)} = 0.411, p < 0.05$). These data suggest that there is a moderate relationship between gender and visual detail recognition. Moreover, the mean scores on the detail recognition test indicate that males ($\bar{x}_M = 12.71$) recalled fewer details, while females ($\bar{x}_F = 17.75$) recalled more details.

7.5.6 Sample Biserial Correlation (Small Data Samples)

A graduate anthropology department at a university wished to determine if its students' grade point averages (GPAs) can be used to predict performance on the department's comprehensive exam required for graduation. The comprehensive exam is graded on a pass/fail basis. Sixteen students participated in the comprehensive exam last year. Five of the students failed the exam. The GPAs and the exam performance of the students are displayed in Table 7.12.

TABLE 7.12

Participant	Exam performance	GPA
1	F	3.5
2	F	3.4
3	F	3.3
4	F	3.2
5	F	3.6
6	P	4.0
7	P	3.6
8	P	4.0
9	P	4.0
10	P	3.8
11	P	3.9
12	P	3.9
13	P	4.0
14	P	3.8
15	P	3.5
16	P	3.6

Exam performance is a continuous dichotomous variable and GPA is an interval scale variable. Therefore, we will use a biserial correlation.

7.5.6.1 State the Null and Research Hypothesis The null hypothesis states that there is no correlation between student GPA and comprehensive exam performance. The research hypothesis states that there is a correlation between student GPA and comprehensive exam performance.

The null hypothesis is

$$H_O\colon \rho_b = 0$$

The research hypothesis is

H_A: $\rho_b \neq 0$

7.5.6.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

7.5.6.3 Choose the Appropriate Test Statistic As stated earlier, we decided to analyze the relationship between the two variables. A correlation will provide the relative strength of the relationship between the two variables. Exam performance is a continuous dichotomous variable and GPA is an interval scale variable. Therefore, we will use a biserial correlation.

7.5.6.4 Compute the Test Statistic First, compute the standard deviation of all values from the interval scale data. Organize the data to manage the summations (see Table 7.13):

$$\bar{x} = \sum x_i \div n$$

$$\bar{x} = 59.1 \div 16$$

TABLE 7.13

Participant	Exam performance	GPA	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	F	3.5	−0.19	0.04
2	F	3.4	−0.29	0.09
3	F	3.3	−0.39	0.16
4	F	3.2	−0.49	0.24
5	F	3.6	−0.09	0.01
6	P	4.0	0.31	0.09
7	P	3.6	−0.09	0.01
8	P	4.0	0.31	0.09
9	P	4.0	0.31	0.09
10	P	3.8	0.11	0.01
11	P	3.9	0.21	0.04
12	P	3.9	0.21	0.04
13	P	4.0	0.31	0.09
14	P	3.8	0.11	0.01
15	P	3.5	−0.19	0.04
16	P	3.6	−0.09	0.01
		$\sum x_i = 59.1$		$\sum(x_i - \bar{x})^2 = 1.07$

$$\bar{x} = 3.69$$

$$s_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{1.07}{16-1}} = \sqrt{0.071}$$

$$s_x = 0.267$$

Next, compute the means and proportions of the values associated with each item from the dichotomous variable. The mean GPA of the exam failures was

$$\begin{aligned} \bar{x}_p &= \sum x_p \div n_F \\ &= (3.5 + 3.4 + 3.3 + 3.2 + 3.6) \div 5 \\ \bar{x}_p &= 3.4 \end{aligned}$$

The mean GPA of the ones who passed the exam was

$$\begin{aligned} \bar{x}_q &= \sum x_q \div n_P \\ &= (4.0 + 3.6 + 4.0 + 4.0 + 3.8 + 3.9 + 3.9 + 4.0 + 3.8 + 3.5 + 3.6) \div 11 \\ \bar{x}_q &= 3.8 \end{aligned}$$

The proportion of exam failures was

$$\begin{aligned} P_p &= n_F \div n \\ &= 5 \div 16 \\ P_p &= 0.3125 \end{aligned}$$

The proportion of the ones who passed the exam was

$$\begin{aligned} P_q &= n_P \div n \\ &= 11 \div 16 \\ P_q &= 0.6875 \end{aligned}$$

Now, determine the height of the unit normal curve ordinate, y, at the point dividing P_p and P_q. We could reference the table of values for the normal distribution, such as Table B.1 in Appendix B, to find y. However, we will compute the value. Using Table B.1 also provides the z-score at the point dividing P_p and P_q, $z = 0.49$:

$$\begin{aligned} y &= \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \\ &= \frac{1}{\sqrt{2\pi}} e^{-0.49^2/2} = \frac{1}{2.51} e^{-0.12} = (0.40)(0.89) \\ y &= 0.3538 \end{aligned}$$

Now, compute the biserial correlation coefficient using the values computed earlier:

$$r_b = \left[\frac{\overline{x}_p - \overline{x}_q}{s_x}\right]\frac{P_p P_q}{y}$$
$$= \left[\frac{3.4 - 3.8}{0.267}\right]\frac{(0.3125)(0.6875)}{0.3538} = (-1.60)(0.6072)$$
$$r_b = -0.972$$

The sign on the correlation coefficient is dependent on the order we managed our dichotomous variable. A quick inspection of the variable means indicates that the GPA of the failures was smaller than the GPA of the ones who passed. Therefore, we should convert the biserial correlation coefficient to a positive value:

$$r_b = 0.972$$

7.5.6.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Table B.8 in Appendix B lists critical values for the Pearson product-moment correlation coefficient. The table requires the degrees of freedom and $df = n - 2$. In this study, $n = 16$ and $df = 16 - 2$. Therefore, $df = 14$. Since we are conducting a two-tailed test and $\alpha = 0.05$, the critical value is 0.497.

7.5.6.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 0.497 and the obtained value is $|r_b| = 0.972$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value is greater than the obtained value, we must not reject the null hypothesis. Since the critical value is less than the absolute value of the obtained value, we reject the null hypothesis.

7.5.6.7 Interpret the Results We rejected the null hypothesis, suggesting that there is a significant and very strong correlation between student GPA and comprehensive exam performance.

7.5.6.8 Reporting the Results The reporting of results for the biserial correlation should include such information as the number of participants (n), two variables that are being correlated, correlation coefficient (r_b), degrees of freedom (df), p-value's relation to α, and the mean values of each dichotomous variable.

For this example, a researcher compared the GPAs of graduate anthropology students who passed their comprehensive exam with students who failed the exam. Five students failed the exam ($n_F = 5$) and 11 students passed it ($n_P = 11$). The researcher compared student GPA and comprehensive exam performance. A biserial correlation produced significant results ($r_{b(14)} = 0.972$, $p < 0.05$). The data suggest that there is an especially strong relationship between student GPA and comprehensive exam performance. Moreover, the mean GPA of the failing students ($\overline{x}_{\text{failure}} = 3.4$) and passing students ($\overline{x}_{\text{passing}} = 3.8$) indicates that the relationship is a direct correlation.

7.5.7 Performing the Biserial Correlation Using SPSS

SPSS does not compute the biserial correlation coefficient. To do so, Field (2005) has suggested using SPSS to perform a Pearson product-moment correlation (as described earlier) and then applying Formula 7.13. However, this procedure will only produce an approximation of the biserial correlation coefficient and we recommend you use a spreadsheet with the procedure we described for the sample biserial correlation.

7.6 EXAMPLES FROM THE LITERATURE

Listed are varied examples of the nonparametric procedures described in this chapter. We have summarized each study's research problem and researchers' rationale(s) for choosing a nonparametric approach. We encourage you to obtain these studies if you are interested in their results.

Greiner and Smith (2006) investigated factors that might affect teacher retention. When they examined the relationship between the Texas state-mandated teacher certification examination and teacher retention, they used a point-biserial correlation. The researchers used the point biserial since teacher retention was measured as a discrete dichotomous variable.

Blumberg and Sokol (2004) examined gender differences in the cognitive strategies that 2nd- and 5th-grade children use when they learn how to play a video game. In part of the study, participants were classified as frequent players or infrequent players. That classification was correlated with game performance. Since player frequency was a discrete dichotomy, the researchers chose a point-biserial correlation.

McMillian et al. (2006) investigated the attitudes of female registered nurses toward male registered nurses. The researchers performed several analyses with a variety of statistical tests. In one analysis, they used a Spearman rank-order correlation to examine the relationship between town population and the participants' responses on an attitude inventory. The attitude inventory was a modified instrument to measure level of sexist attitude. Participants indicated agreement or disagreement with statements using a four-point Likert scale. The Spearman rank-order correlation was chosen because the attitude inventory resembled an ordinal scale.

Fitzgerald et al (2007) examined the validity of an instrument designed to measure the performance of physical therapy interns. They used a correlation analysis to examine the relationship between two measures of clinical competence. Since one of the measures was ordinal, the researchers used a Spearman rank-order correlation.

Flannelly et al. (2005) reviewed the research literature of studies that investigated the effects of religion on adolescent tobacco use. The authors used a biserial correlation to compare studies' effect (no effect vs. effect) with sample size.

7.7 SUMMARY

The relationship between two variables can be compared with a correlation analysis. If any of the variables are ordinal or dichotomous, a nonparametric correlation is useful. The Spearman rank-order correlation, also called the Spearman's ρ, is used to compare the relationship involving ordinal, or rank-ordered, variables. The point-biserial and biserial correlations are used to compare the relationship between two variables if one of the variables is dichotomous. The parametric equivalent to these correlations is the Pearson product-moment correlation.

In this chapter, we described how to perform and interpret a Spearman rank-order, point-biserial, and biserial correlations. We also explained how to perform the procedures using SPSS. Finally, we offered varied examples of these nonparametric statistics from the literature. The next chapter will involve comparing nominal scale data.

7.8 PRACTICE QUESTIONS

1. The business department at a small college wanted to compare the relative class rank of its MBA graduates with their fifth-year salaries. The data collected by the department are presented in Table 7.14. Compare the graduates' class rank with their fifth-year salaries.

TABLE 7.14

Relative class rank	Fifth-year salary ($)
1	83,450
2	67,900
3	89,000
4	80,500
5	91,000
6	55,440
7	101,300
8	50,560
9	76,050

Use a two-tailed Spearman rank-order correlation with $\alpha = 0.05$ to determine if a relationship exists between the two variables. Report your findings.

2. A researcher was contracted by the military to assess soldiers' perception of a new training program's effectiveness. Fifteen soldiers participated in the program. The researcher used a survey to measure the soldiers' perceptions of the program's effectiveness. The survey used a Likert-type scale that ranged from 5 = *strongly agree* to 1 = *strongly disagree*. Using the data presented in Table 7.15, compare the soldiers' average survey scores with the total number of years the soldiers had been serving.

Use a two-tailed Spearman rank-order correlation with $\alpha = 0.05$ to determine if a relationship exists between the two variables. Report your findings.

TABLE 7.15

Average survey score	Years of service
4.0	18
4.0	15
2.4	2
4.2	13
3.4	4
4.0	10
5.0	24
1.8	4
3.2	9
2.5	5
2.5	3
3.0	8
3.6	16
4.6	14
4.8	12

3. A middle school history teacher wished to determine if there is a connection between gender and history knowledge among 8th-grade gifted students. The teacher administered a 50 item test at the beginning of the school year to 16 gifted 8th-grade students. The scores from the test are presented in Table 7.16.

Use a two-tailed point-biserial correlation with $\alpha = 0.05$ to determine if a relationship exists between the two variables. Report your findings.

TABLE 7.16

Participant	Gender	Posttest score
1	M	44
2	M	30
3	M	50
4	M	33
5	M	37
6	M	35
7	M	36
8	F	29
9	F	39
10	F	33
11	F	50
12	F	45
13	F	37
14	F	30
15	F	34
16	F	50

4. A researcher wished to determine if there is a connection between poverty and self-esteem. Income level was used to classify 18 participants as either below poverty or above poverty. Participants completed a 20 item survey to measure self-esteem. The scores from the survey are reported in Table 7.17.

Use a two-tailed biserial correlation with $\alpha = 0.05$ to determine if a relationship exists between the two variables. Report your findings.

TABLE 7.17

Participant	Poverty level	Survey score
1	Above	15
2	Above	19
3	Above	15
4	Above	20
5	Above	7
6	Above	12
7	Above	3
8	Above	15
9	Below	9
10	Below	5
11	Below	13
12	Below	13
13	Below	11
14	Below	10
15	Below	8
16	Below	9
17	Below	10
18	Below	17

7.9 SOLUTIONS TO PRACTICE QUESTIONS

1. The results from the analysis are displayed in SPSS Output 7.3.
The results from the Spearman rank-order correlation ($r_s = -0.217$, $p > 0.05$) did not produce significant results. Based on these data, we can state that there is no clear relationship between graduates' relative class rank and fifth-year salary.

2. The results from the analysis are displayed in SPSS Output 7.4.
The results from the Spearman rank-order correlation ($r_s = 0.806$, $p < 0.05$) produced significant results. Based on these data, we can state that there is a very strong correlation between soldiers' survey scores concerning the new program's effectiveness and their total years of military service.

3. The results from the point-biserial correlation ($r_{pb} = 0.047$, $p > 0.05$) did not produce significant results. Based on these data, we can state that there is no clear

Correlations

			Class_Rank	Fifth_Yr_Salary
Spearman's rho	Class_Rank	Correlation Coefficient	1.000	-.217
		Sig. (2-tailed)	.	.576
		N	9	9
	Fifth_Yr_Salary	Correlation Coefficient	-.217	1.000
		Sig. (2-tailed)	.576	.
		N	9	9

SPSS OUTPUT 7.3

Correlations

			Survey_Score	Years_of_Service
Spearman's rho	Survey_Score	Correlation Coefficient	1.000	.806**
		Sig. (2-tailed)	.	.000
		N	15	15
	Years_of_Service	Correlation Coefficient	.806**	1.000
		Sig. (2-tailed)	.000	.
		N	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

SPSS OUTPUT 7.4

relationship between 8th-grade gifted students' gender and their score on the history knowledge test administered by the teacher.

Note that the results obtained from using SPSS is $r_{pb} = 0.049$, $p > 0.05$.

4. The results from the biserial correlation ($r_b = 0.372$, $p > 0.05$) did not produce significant results. Based on these data, we can state that there is no clear relationship between poverty level and self-esteem.

CHAPTER 8

TESTS FOR NOMINAL SCALE DATA: CHI-SQUARE AND FISHER EXACT TESTS

8.1 OBJECTIVES

In this chapter, you will learn the following items:

- How to perform a chi-square goodness-of-fit test.
- How to perform a chi-square goodness-of-fit test using SPSS®.
- How to perform a chi-square test for independence.
- How to perform a chi-square test for independence using SPSS.
- How to perform the Fisher exact test.
- How to perform Fisher exact test using SPSS.

8.2 INTRODUCTION

Sometimes, data are best collected or conveyed nominally or categorically. These data are represented by counting the number of times a particular event or condition occurs. In such cases, you may be seeking to determine if a given set of counts, or frequencies, statistically matches some known, or expected, set. Or, you may wish to determine if two or more categories are statistically independent. In either case, we can use a nonparametric procedure to analyze nominal data.

In this chapter, we present three procedures for examining nominal data: chi-square (χ^2) goodness of fit, χ^2-test for independence, and the Fisher exact test. We will also explain how to perform the procedures using SPSS. Finally, we offer varied examples of these nonparametric statistics from the literature.

8.3 THE χ^2 GOODNESS-OF-FIT TEST

Some situations in research involve investigations and questions about relative frequencies and proportions for a distribution. Some examples might include a

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

comparison of the number of women pediatricians with the number of men pediatricians, a search for significant changes in the proportion of students entering a writing contest over 5 years, or an analysis of customer preference of three candy bar choices. Each of these examples asks a question about a proportion in the population.

When comparing proportions, we are not measuring a numerical score for each individual. Instead, we classify each individual into a category. We then find out what proportion of the population is classified into each category. The χ^2 goodness-of-fit test is designed to answer this type of question.

The χ^2 goodness-of-fit test uses sample data to test the hypothesis about the proportions of the population distribution. The test determines how well the sample proportions fit the proportions specified in the null hypothesis.

8.3.1 Computing the χ^2 Goodness-of-Fit Test Statistic

The χ^2 goodness-of-fit test is used to determine how well the obtained sample proportions or frequencies for a distribution fit the population proportions or frequencies specified in the null hypothesis. The χ^2 statistic can be used when two or more categories are involved in the comparison. Formula 8.1 is referred to as Pearson's χ^2 and is used to determine the χ^2 statistic:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} \tag{8.1}$$

where f_o is the observed frequency (the data) and f_e is the expected frequency (the hypothesis).

Use Formula 8.2 to determine the expected frequency f_e:

$$f_e = P_i n \tag{8.2}$$

where P_i is a category's frequency proportion with respect to the other categories and n is the sample size of all categories and $\Sigma f_o = n$.

Use Formula 8.3 to determine the degrees of freedom for the χ^2-test:

$$df = C - 1 \tag{8.3}$$

where C is the number of categories.

8.3.2 Sample χ^2 Goodness-of-Fit Test (Category Frequencies Equal)

A marketing firm is conducting a study to determine if there is a significant preference for a type of chicken that is served as a fast food. The target group is college students. It is assumed that there is no preference when the study is started. The types of food that are being compared are chicken sandwich, chicken strips, chicken nuggets, and chicken taco.

The sample size for this study was $n = 60$. The data in Table 8.1 represent the observed frequencies for the 60 participants who were surveyed at the fast food restaurants.

TABLE 8.1

Chicken sandwich	Chicken strips	Chicken nuggets	Chicken taco
10	25	18	7

We want to determine if there is any preference for one of the four chicken fast foods that were purchased to eat by the college students. Since the data only need to be classified into categories, and no sample mean nor sum of squares needs to be calculated, the χ^2 statistic goodness-of-fit test can be used to test the nonparametric data.

8.3.2.1 State the Null and Research Hypotheses The null hypothesis states that there is no preference among the different categories. There is an equal proportion or frequency of participants selecting each type of fast food that uses chicken. The research hypothesis states that one or more of the chicken fast foods is preferred over the others by the college student.

The null hypothesis is

H_O: In the population of college students, there is no preference of one chicken fast food over any other. Thus, the four fast food types are selected equally often and the population distribution has the proportions shown in Table 8.2.

TABLE 8.2

Chicken sandwich	Chicken strips	Chicken nuggets	Chicken taco
25%	25%	25%	25%

H_A: In the population of college students, there is at least one chicken fast food preferred over the others.

8.3.2.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

8.3.2.3 Choose the Appropriate Test Statistic The data are obtained from 60 college students who eat fast food chicken. Each student was asked which of the four chicken types of food he or she purchased to eat and the result was tallied under the corresponding category type. The final data consisted of frequencies for each of the four types of chicken fast foods. These categorical data, which are represented by frequencies or proportions, are analyzed using the χ^2 goodness-of-fit test.

8.3.2.4 Compute the Test Statistic First, tally the observed frequencies, f_o, for the 60 students who were in the study. Use these data to create the observed frequency table shown in Table 8.3.

TABLE 8.3

	Chicken sandwich	Chicken strips	Chicken nuggets	Chicken taco
Observed frequencies	10	25	18	7

Next, calculate the expected frequency for each category. In this case, the expected frequency, f_e, will be the same for all four categories since our research problem assumes that all categories are equal:

$$f_e = P_i n = \frac{1}{4}(60)$$

$$f_e = 15$$

Table 8.4 presents the expected frequencies for each category.

TABLE 8.4

	Chicken sandwich	Chicken strips	Chicken nuggets	Chicken taco
Expected frequencies	15	15	15	15

Using the values for the observed and expected frequencies, the χ^2 statistic may be calculated:

$$\begin{aligned}
\chi^2 &= \sum \frac{(f_o - f_e)^2}{f_e} \\
&= \frac{(10-15)^2}{15} + \frac{(25-15)^2}{15} + \frac{(18-15)^2}{15} + \frac{(7-15)^2}{15} \\
&= \frac{(-5)^2}{15} + \frac{(10)^2}{15} + \frac{(3)^2}{15} + \frac{(-8)^2}{15} \\
&= \frac{25}{15} + \frac{100}{15} + \frac{9}{15} + \frac{64}{15} \\
&= 1.67 + 6.67 + 0.60 + 4.27 \\
\chi^2 &= 13.21
\end{aligned}$$

8.3.2.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic

Before we go to the table of critical values, we must determine the degrees of freedom, *df*. In this example, there are four categories, $C = 4$. To find the degrees of freedom, use $df = C - 1 = 4 - 1$. Therefore, $df = 3$.

Now, we use Table B.2 in Appendix B, which lists the critical values for the χ^2. The critical value is found in the χ^2 table for three degrees of freedom, $df = 3$.

Since we set the $\alpha = 0.05$, the critical value is 7.81. A calculated value that is greater than or equal to 7.81 will lead us to reject the null hypothesis.

8.3.2.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 7.81 and the obtained value is $\chi^2 = 13.21$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value exceeds the obtained value, we do not reject the null hypothesis. Since the critical value is less than our obtained value, we must reject the null hypothesis.

Note that the critical value for $\alpha = 0.01$ is 11.34. Since the obtained value is 13.21, a value greater than 11.34, the data indicate that the results are highly significant.

8.3.2.7 Interpret the Results We rejected the null hypothesis, suggesting that there is a real difference among chicken fast food choices preferred by college students. In particular, the data show that a larger portion of the students preferred the chicken strips and only a few of them preferred the chicken taco.

8.3.2.8 Reporting the Results The reporting of results for the χ^2 goodness of fit should include such information as the total number of participants in the sample and the number that were classified in each category. In some cases, bar graphs are good methods of presenting the data. In addition, include the χ^2 statistic, degrees of freedom, and p-value's relation to α. For this study, the number of students who ate each type of chicken fast food should be noted either in a table or plotted on a bar graph. The probability, $p < 0.01$, should also be indicated with the data to show the degree of significance of the χ^2.

For this example, 60 college students were surveyed to determine which fast food type of chicken they purchased to eat. The four choices were chicken sandwich, chicken strips, chicken nuggets, and chicken taco. Student choices were 10, 25, 18, and 7, respectively. The χ^2 goodness-of-fit test was significant ($\chi^2_{(3)} = 13.21$, $p < 0.01$). Based on these results, a larger portion of the students preferred the chicken strips while only a few students preferred the chicken taco.

8.3.3 Sample χ^2 Goodness-of-Fit Test (Category Frequencies Not Equal)

Sometimes, research is being conducted in an area where there is a basis for different expected frequencies in each category. In this case, the null hypothesis will indicate different frequencies for each of the categories according to the expected values. These values are usually obtained from previous data that were collected in similar studies.

In this study, a school system uses three different physical fitness programs because of scheduling requirements. A researcher is studying the effect of the programs on 10th-grade students' 1-mile run performance. Three different physical fitness programs were used by the school system and will be described later.

- *Program 1.* Delivers health and physical education in 9-week segments with an alternating rotation of nine straight weeks of health education and then nine straight weeks of physical education.
- *Program 2.* Delivers health and physical education everyday with 30 min for health, 10 min for dress-out time, and 50 min of actual physical activity
- *Program 3.* Delivers health and physical education in 1-week segments with an alternating rotation of 1 week of health education and then 1 week of physical education

Using students who participated in all three programs, the researcher is comparing these programs based on student performances on the 1-mile run. The researcher recorded the program in which each student received the most benefit. Two hundred fifty students had participated in all three programs. The results for all of the students are recorded in Table 8.5.

TABLE 8.5

Program 1	Program 2	Program 3
110	55	85

We want to determine if the frequency distribution in the case earlier is different from previous studies. Since the data only need to be classified into categories, and no sample mean or sum of squares needs to be calculated, the χ^2 goodness-of-fit test can be used to test the nonparametric data.

8.3.3.1 State the Null and Research Hypotheses

The null hypothesis states the proportion of students who benefited most from one of the three programs based on a previous study. As shown in Table 8.6, there are unequal expected frequencies for the null hypothesis. The research hypothesis states that there is at least one of the three categories that will have a different proportion or frequency than those identified in the null hypothesis.

The null hypothesis is

H_O: The proportions do not differ from the previously determined proportions shown in Table 8.6.

H_A: The population distribution has a different shape than that specified in the null hypothesis.

TABLE 8.6

Program 1	Program 2	Program 3
32%	22%	45%

8.3.3.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

8.3.3.3 Choose the Appropriate Test Statistic The data are being obtained from the 1-mile run performance of 250 10th-grade students who participated in a school system's three health and physical education programs. Each student was categorized based on the plan in which he or she benefited most. The final data consisted of frequencies for each of the three plans. These categorical data which are represented by frequencies or proportions are analyzed using the χ^2 goodness-of-fit test.

8.3.3.4 Compute the Test Statistic First, tally the observed frequencies for the 250 students who were in the study. This was performed by the researcher. Use the data to create the observed frequency table shown in Table 8.7.

TABLE 8.7

	Program 1	Program 2	Program 3
Observed frequencies	110	55	85

Next, calculate the expected frequencies for each category. In this case, the expected frequency will be different for each category. Each one will be based on proportions stated in the null hypothesis:

$$f_{ei} = P_i n$$

$$f_e \text{ for program } 1 = 0.32(250) = 80$$

$$f_e \text{ for program } 2 = 0.22(250) = 55$$

$$f_e \text{ for program } 3 = 0.46(250) = 115$$

Table 8.8 presents the expected frequencies for each category.

Use the values for the observed and expected frequencies calculated earlier to calculate the χ^2 statistic:

$$\begin{aligned}\chi^2 &= \sum \frac{(f_o - f_e)^2}{f_e} \\ &= \frac{(110-80)^2}{80} + \frac{(55-55)^2}{55} + \frac{(85-115)^2}{115} \\ &= \frac{30^2}{80} + \frac{0^2}{55} + \frac{-30^2}{115} \\ &= 11.25 + 0 + 7.83 \\ \chi^2 &= 19.08\end{aligned}$$

TABLE 8.8

	Program 1	Program 2	Program 3
Expected frequencies	80	55	115

8.3.3.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Before we go to the table of critical values, we must determine the degrees of freedom, *df*. In this example, there are four categories, $C = 3$. To find the degrees of freedom, use $df = C - 1 = 3 - 1$. Therefore, $df = 2$.

Now, we use Table B.2 in Appendix B, which lists the critical values for the χ^2. The critical value is found in the χ^2 table for two degrees of freedom, $df = 2$. Since we set $\alpha = 0.05$, the critical value is 5.99. A calculated value that is greater than 5.99 will lead us to reject the null hypothesis.

8.3.3.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 5.99 and the obtained value is $\chi^2 = 19.08$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value exceeds the obtained value, we do not reject the null hypothesis. Since the critical value is less than our obtained value, we must reject the null hypothesis.

Note that the critical value for $\alpha = 0.01$ is 9.21. Since the obtained value is 19.08, which is greater than the critical value, the data indicate that the results are highly significant.

8.3.3.7 Interpret the Results We rejected the null hypothesis, suggesting that there is a real difference in how the health and physical education program affects the performance of students on the 1-mile run as compared with the existing research. By comparing the expected frequencies of the past study and those obtained in the current study, it can be noted that the results from program 2 did not change. Program 2 was least effective in both cases, with no difference between the two. Program 1 became more effective and program 3 became less effective.

8.3.3.8 Reporting the Results The reporting of results for the χ^2 goodness-of-fit should include such information as the total number of participants in the sample, the number that were classified in each category, and the expected frequencies that are being used for comparison. It is important also to cite a source for the expected frequencies so that the decisions made from the study can be supported. In addition, include the χ^2 statistic, degrees of freedom, and *p*-value's relation to α. It is often a good idea to present a bar graph to display the observed and expected frequencies from the study. For this study, the probability, $p < 0.01$, should also be indicated with the data to show the degree of significance of the χ^2.

For this example, 250 10th-grade students participated in three different health and physical education programs. Using 1-mile run performance, students' program of greatest benefit was compared with the results from past research. The χ^2 goodness-of-fit test was significant ($\chi^2_{(2)} = 19.08$, $p < 0.01$). Based on these results, program 2 was least effective in both cases, with no difference between the two. Program 1 became more effective and program 3 became less effective.

8.3.4 Performing the χ^2 Goodness-of-Fit Test Using SPSS

We will analyze the data from the example earlier using SPSS.

8.3.4.1 Define Your Variables First, click the "Variable View" tab at the bottom of your screen (see Fig. 8.1). The χ^2 goodness-of-fit test requires two variables: one variable to identify the categories and a second variable to identify the observed frequencies. Type the names of these variables in the "Name" column. In our example, we define the variables as "Program" and "count."

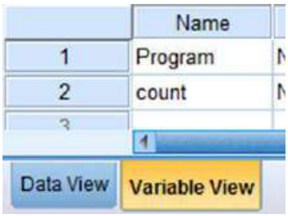


FIGURE 8.1

You must assign values to serve as a reference for each category in the observed frequency variable. It is often easiest to assign each category a whole number value. As shown in Figure 8.2, our categories are "Program 1," "Program 2," and "Program 3." First, we selected the "count" variable and clicked the gray square in the "Values" field. Then, we set a value of 1 to equal "Program 1," a value of 2 to equal "Program 2," and a value of 3 to equal "Program 3." We use the "Add" button to move the variable labels to the box below. Repeat this procedure for the "Program" variable so that the output tables will display these labels.

8.3.4.2 Type in Your Values Click the "Data View" tab at the bottom of your screen. First, enter the data for each category using the whole numbers you assigned to represent the categories. As shown in Figure 8.3, we entered the values "1," "2," and "3" in the "Program" variable. Second, enter the observed frequencies next to the corresponding category values. In our example, we entered the observed frequencies "110," "55," and "85."

8.3.4.3 Analyze Your Data First, use the "Weight Cases" command to allow the observed frequency variable to reference the category variable. As shown in Figure 8.4, use the pull-down menus to choose "Data" and "Weight Cases . . .".

The default setting is "Do not weight cases." Click the circle next to "Weight cases by" as shown in Figure 8.5. Select the variable with the observed frequencies.

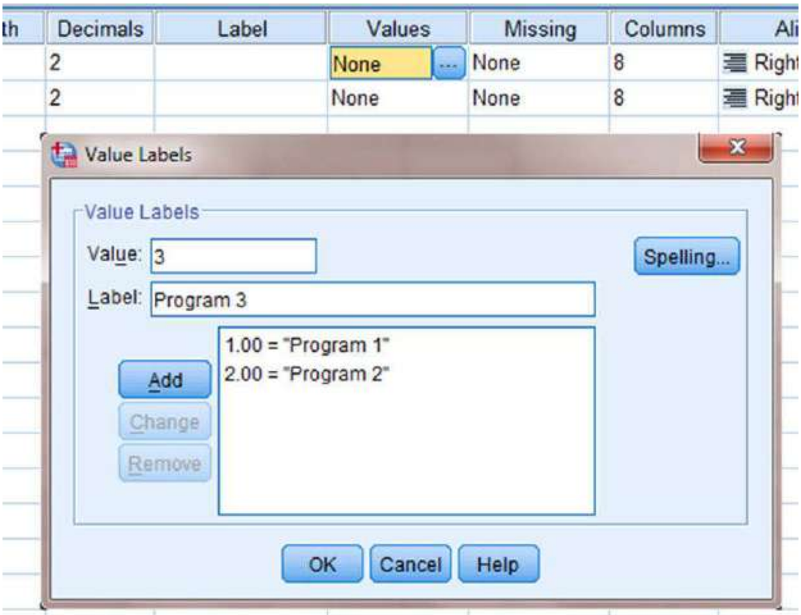


FIGURE 8.2

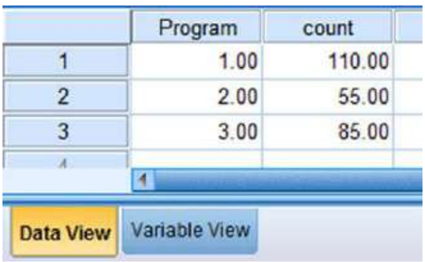


FIGURE 8.3

Move that variable to the "Frequency Variable:" box by clicking the small arrow button. In our example, we have moved the "count" variable. Finally, click "OK."

As shown in Figure 8.6, use the pull-down menus to choose "Analyze," "Nonparametric Tests," "Legacy Dialogs," and "Chi-square. . .".

First, move the category variable to the "Test Variable List:" box by selecting that variable and clicking the small arrow button near the center of the window. As shown in Figure 8.7, we have chosen the "Program" variable. Then, enter your "Expected Values." Notice that the option "All categories equal" is the default setting. Since this example does not have equal categories, we must select the "Values:" option to set the expected values. Enter the expected frequencies for each category in the order that they are listed in the SPSS Data View. After you type in an expected frequency in the "Values:" field, click "Add."

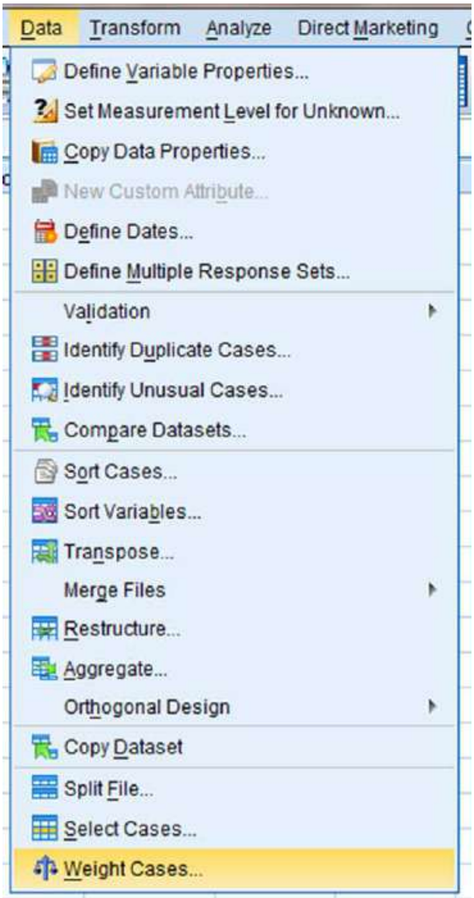


FIGURE 8.4

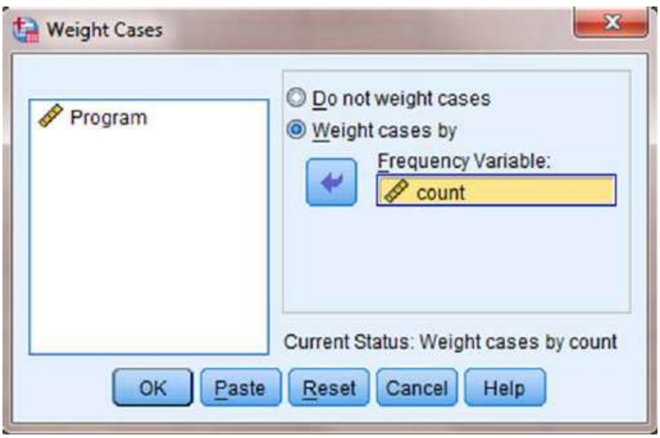


FIGURE 8.5

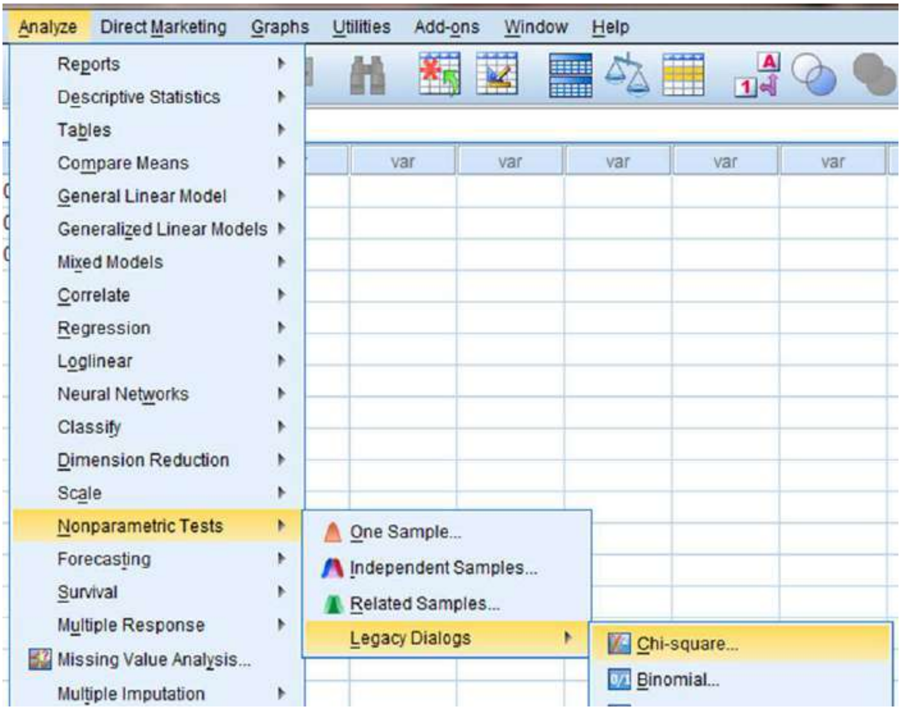


FIGURE 8.6

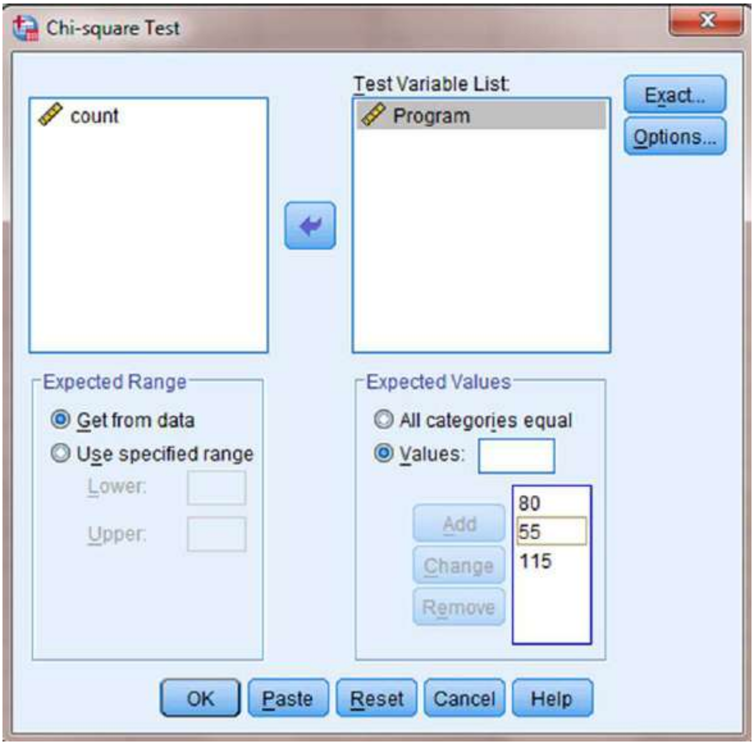


FIGURE 8.7

For our example, we have entered 80, 55, and 115, respectively. Finally, click "OK" to perform the analysis.

8.3.4.4 Interpret the Results from the SPSS Output Window The first output table (see SPSS Output 8.1a) provides the observed and expected frequencies for each category and the total count.

Program

	Observed N	Expected N	Residual
Program 1	110	80.0	30.0
Program 2	55	55.0	.0
Program 3	85	115.0	-30.0
Total	250		

SPSS OUTPUT 8.1A

The second output table (see SPSS Output 8.1b) provides the χ^2 statistic ($\chi^2 = 19.076$), the degrees of freedom ($df = 2$), and the significance ($p \approx 0.000$).

Test Statistics

	Program
Chi-Square	19.076[a]
df	2
Asymp. Sig.	.000

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 55.0.

SPSS OUTPUT 8.1B

Based on the results from SPSS, three programs were compared with unequal expected frequencies. The χ^2 goodness-of-fit test was significant ($\chi^2_{(2)} = 19.08$, $p < 0.01$). Based on these results, program 2 was least effective in both cases, with no difference between the two. Program 1 became more effective and program 3 became less effective.

8.4 THE χ^2 TEST FOR INDEPENDENCE

Some research involves investigations of frequencies of statistical associations of two categorical attributes. Examples might include a sample of men and women who bought a pair of shoes or a shirt. The first attribute, A, is the gender of the shopper with two possible categories:

$$\text{men} = A_1$$

$$\text{women} = A_2$$

The second attribute, B, is the clothing type purchased by each individual:

$$\text{pair of shoes} = B_1$$
$$\text{shirt} = B_2$$

We will assume that each person purchased only one item, either a pair of shoes or a shirt. The entire set of data is then arranged into a joint-frequency distribution table. Each individual is classified into one category which is identified by a pair of categorical attributes (see Table 8.9).

TABLE 8.9

	A_1	A_2
B_1	(A_1, B_1)	(A_2, B_1)
B_2	(A_1, B_2)	(A_2, B_2)

The χ^2-test for independence uses sample data to test the hypothesis that there is no statistical association between two categories. In this case, whether there is a significant association between the gender of the purchaser and the type of clothing purchased. The test determines how well the sample proportions fit the proportions specified in the null hypothesis.

8.4.1 Computing the χ^2 Test for Independence

The χ^2-test for independence is used to determine whether there is a statistical association between two categorical attributes. The χ^2 statistic can be used when two or more categories are involved for two attributes. Formula 8.4 is referred to as Pearson's χ^2 and is used to determine the χ^2 statistic:

$$\chi^2 = \sum_j \sum_k \frac{(f_{ojk} - f_{ejk})^2}{f_{ejk}} \tag{8.4}$$

where f_{ojk} is the observed frequency for cell A_jB_k and f_{ejk} is the expected frequency for cell A_jB_k.

In tests for independence, the expected frequency f_{ejk} in any cell is found by multiplying the row total times the column total and dividing the product by the grand total N. Use Formula 8.5 to determine the expected frequency f_{ejk}:

$$f_{ejk} = \frac{(\text{freq.}A_j)(\text{freq.}B_k)}{N} \tag{8.5}$$

The degrees of freedom df for the χ^2 is found using Formula 8.6:

$$df = (R-1)(C-1) \tag{8.6}$$

where R is the number of rows and C is the number of columns.

It is important to note that Pearson's χ^2 formula returns a value that is too small when data form a 2 × 2 contingency table. This increases the chance of a type I error. In such a circumstance, one might use the Yates's continuity correction shown in Formula 8.7:

$$\chi^2 = \sum_j \sum_k \frac{(f_{ojk} - f_{ejk} - 0.5)^2}{f_{ejk}} \tag{8.7}$$

Daniel (1990) has cited a number of criticisms to the Yates's continuity correction. While he recognizes that the procedure has been frequently used, he also observes a decline in its popularity. Toward the end of this chapter, we present an alternative for analyzing a 2 × 2 contingency table using the Fisher exact test.

At this point, the analysis is limited to identifying an association's presence or absence. In other words, the χ^2-test's level of significance does not describe the strength of its association. We can use the *effect size* to analyze the degree of association. For the χ^2-test for independence, the effect size between the nominal variables of a 2 × 2 contingency table can be calculated and represented with the phi (ϕ) coefficient (see Formula 8.8):

$$\phi = \sqrt{\frac{\chi^2}{n}} \tag{8.8}$$

where χ^2 is the chi-square test statistic and n is the number in the entire sample.

The ϕ coefficient ranges from 0 to 1. Cohen (1988) defined the conventions for effect size as *small* = 0.10, *medium* = 0.30, and *large* = 0.50. (Correlation coefficient and effect size are both measures of association. See Chapter 7 concerning correlation for more information on Cohen's assignment of effect size's relative strength.)

When the χ^2 contingency table is larger than 2 × 2, Cramer's V statistic may be used to express effect size. The formula for Cramer's V is shown in Formula 8.9:

$$V = \sqrt{\frac{\chi^2}{(n)(L-1)}} \tag{8.9}$$

where χ^2 is the chi-square test statistic, n is the total number in the sample, and L is the minimum value of the row totals and column totals from the contingency table.

8.4.2 Sample χ^2 Test for Independence

A counseling department for a school system is conducting a study to investigate the association between children's attendance in public and private preschool and their behavior in the kindergarten classroom. It is the researcher's desire to see if there is any positive association between early exposure to learning and behavior in the classroom.

The sample size for this study was $n = 100$. The following data in Table 8.10 represent the observed frequencies for the 100 children whose behavior was observed

TABLE 8.10

	Behavior in kindergarten			
	Poor	Average	Good	Row totals
Public preschool	12	25	10	47
Private preschool	6	12	0	18
No preschool	2	23	10	35
Column totals	20	60	20	100

during their first 6 weeks of school. The students who were in the study were identified by the type of preschool educational exposure they received.

We want to determine if there is any association between type of preschool experience and behavior in kindergarten in the first 6 weeks of school. Since the data only need to be classified into categories, and no sample mean nor sum of squares needs to be calculated, the χ^2 statistic for independence can be used to test the nonparametric data.

8.4.2.1 State the Null and Research Hypotheses The null hypothesis states that there is no association between the two categories. The behavior of the children in kindergarten is independent of the type of preschool experience they had. The research hypothesis states that there is a significant association between the preschool experience of the children and their behavior in kindergarten.

The null hypothesis is

> H_o: In the general population, there is no association between type of preschool experience a child has and the child's behavior in kindergarten.

The research hypothesis states

> H_A: In the general population, there is a predictable relationship between the preschool experience and the child's behavior in kindergarten.

8.4.2.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

8.4.2.3 Choose the Appropriate Test Statistic The data are obtained from 100 children in kindergarten who experienced differing preschool preparation prior to entering formal education. The kindergarten teachers for the children were ask to rate students' behaviors using three broad levels of ratings obtained from a survey. The students were then divided up into three groups according to preschool experience (no preschool, private preschool, and public preschool). These data are organized into a two-dimensional categorical distribution that can be analyzed using an independent χ^2-test.

8.4.2.4 Compute the Test Statistic First, tally the observed frequencies f_{ojk} for the 100 students who were in the study. Use these data to create the observed frequency table shown in Table 8.11.

TABLE 8.11

	Behavior in kindergarten (observed)			
	Poor	Average	Good	Row totals
Public preschool	12	25	10	47
Private preschool	6	12	0	18
No preschool	2	23	10	35
Column totals	20	60	20	100

Next, calculate the expected frequency f_{ejk} for each category:

$$f_{e11} = \frac{(47)(20)}{100} = 940/100 = 9.4$$
$$f_{e12} = \frac{(47)(60)}{100} = 2820/100 = 28.2$$
$$f_{e13} = \frac{(47)(20)}{100} = 940/100 = 9.4$$
$$f_{e21} = \frac{(18)(20)}{100} = 360/100 = 3.6$$
$$f_{e22} = \frac{(18)(60)}{100} = 1080/100 = 10.8$$
$$f_{e23} = \frac{(18)(20)}{100} = 360/100 = 3.6$$
$$f_{e31} = \frac{(35)(20)}{100} = 700/100 = 7.0$$
$$f_{e32} = \frac{(35)(60)}{100} = 2100/100 = 21.0$$
$$f_{e33} = \frac{(35)(20)}{100} = 700/100 = 7.0$$

Place these values in an expected frequency table (see Table 8.12).

TABLE 8.12

	Behavior in kindergarten (expected)			
	Poor	Average	Good	Row totals
Public preschool	9.4	28.2	9.4	47
Private preschool	3.6	10.8	3.6	18
No preschool	7.0	21.0	7.0	35
Column totals	20	60	20	100

Using the values for the observed and expected frequencies in Tables 8.11 and 8.12, the χ^2 statistic can be calculated:

$$\begin{aligned}\chi^2 &= \sum_j \sum_k \frac{(f_{ojk} - f_{ejk})^2}{f_{ejk}} \\ &= \frac{(12-9.4)^2}{9.4} + \frac{(25-28.2)^2}{28.2} + \frac{(10-9.4)^2}{9.4} + \frac{(6-3.6)^2}{3.6} + \frac{(12-10.8)^2}{10.8} \\ &\quad + \frac{(0-3.6)^2}{3.6} + \frac{(2-7)^2}{7} + \frac{(23-21)^2}{21} + \frac{(10-7)^2}{7} \\ &= \frac{(2.6)^2}{9.4} + \frac{(-3.2)^2}{28.2} + \frac{(0.6)^2}{9.4} + \frac{(2.4)^2}{3.6} + \frac{(1.2)^2}{10.8} \\ &\quad + \frac{(-3.6)}{3.6} + \frac{(-5)^2}{7} + \frac{(2)^2}{21} + \frac{(3)^2}{7} \\ &= 0.72 + 0.36 + 0.04 + 1.60 + 0.13 + 3.60 + 3.57 + 0.19 + 1.29 \\ \chi^2 &= 11.50\end{aligned}$$

8.4.2.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Before we go to the table of critical values, we need to determine the degrees of freedom, *df*. In this example, there are three categories in the preschool experience dimension, $R = 3$, and three categories in the behavior dimension, $C = 3$. To find the degrees of freedom, use $df = (R - 1)(C - 1) = (3 - 1)(3 - 1)$. Therefore, $df = 4$.

Now, we use Table B.2 in Appendix B, which lists the critical values for the χ^2. The critical value is found in the χ^2 table for four degrees of freedom, $df = 4$. Since we set $\alpha = 0.05$, the critical value is 9.49. A calculated value that is greater than or equal to 9.49 will lead us to reject the null hypothesis.

8.4.2.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is 9.49 and the obtained value is $\chi^2 = 11.50$. If the critical value is less than or equal to the obtained value, we must reject the null hypothesis. If instead, the critical value exceeds the obtained value, we do not reject the null hypothesis. Since the critical value is less than our obtained value, we must reject the null hypothesis.

8.4.2.7 Interpret the Results We rejected the null hypothesis, suggesting that there is a real association between type of preschool experience children obtained and their behavior in the kindergarten classroom during their first few weeks in school. In particular, data tend to show that children who have private schooling do not tend to get good behavior ratings in school. The other area that tends to show some significant association is between poor behavior and no preschool experience. The students who had no preschool had very few poor behavior ratings in comparison with the other two groups.

At this point, the analysis is limited to identifying an association's presence or absence. In other words, the χ^2-test's level of significance does not describe the strength of its association. The American Psychological Association (2001), however, has called for a measure of the degree of association called the effect size. For the χ^2-test for independence with a 3×3 contingency table, we determine the strength of association, or the effect size, using Cramer's V.

From Table 8.11, we find that $L = 3$. For $n = 100$ and $\chi^2 = 11.50$, we use Formula 8.9 to determine V:

$$V = \sqrt{\frac{11.50}{(100)(3-1)}}$$
$$= \sqrt{0.06}$$
$$V = 0.24$$

Our effect size, Cramer's V, is 0.24. This value indicates a medium level of association between type of preschool experience children obtained and their behavior in the kindergarten classroom during their first few weeks in school.

8.4.2.8 Reporting the Results The reporting of results for the χ^2-test for independence should include such information as the total number of participants in the sample and the number of participants classified in each of the categories. In addition, include the χ^2 statistic, degrees of freedom, and the p-value's relation to α. For this study, the number of children who were in each category (including preschool experience and behavior rating) should be presented in the two-dimensional table (see Table 8.10).

For this example, the records of 100 kindergarten students were examined to determine whether there was an association between preschool experience and behavior in kindergarten. The three preschool experiences were no preschool, private preschool, and public preschool. The three behavior ratings were poor, average, and good. The χ^2-test was significant ($\chi^2_{(4)} = 11.50, p < 0.05$). Moreover, our effect size, using Cramer's V, was 0.24. Based on the results, there was a tendency shown for students with private preschool to not have good behavior and those with no preschool to not have poor behavior. It also indicated that average behavior was strong for all three preschool experiences.

8.4.3 Performing the χ^2 Test for Independence Using SPSS

We will analyze the data from the example earlier using SPSS.

8.4.3.1 Define Your Variables First, click the "Variable View" tab at the bottom of your screen, as shown in Figure 8.8. The χ^2-test for independence requires variables to identify the conditions in the rows: one variable to identify the conditions of the rows and a second variable to identify the conditions of the columns. According to the previous example, the "Behavior" variable will represent the columns. "School_Type" will represent the rows. Finally, we need a variable to represent the observed frequencies. "Frequency" represents the observed frequencies.

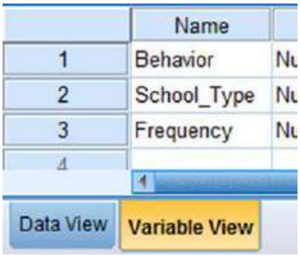


FIGURE 8.8

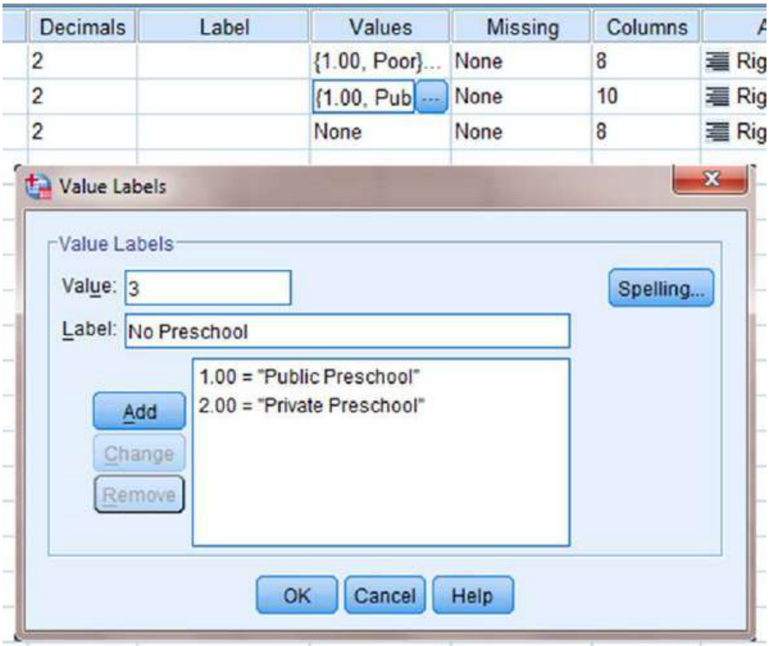
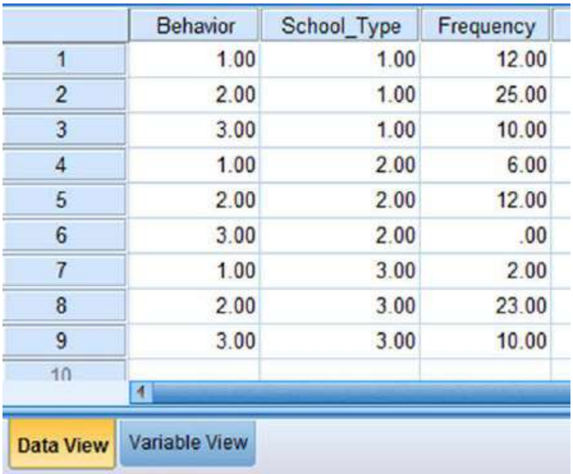


FIGURE 8.9

You must assign values to serve as a reference for the column and row variables. It is often easiest to assign each category a whole number value. First, click the gray square in the "Values" field to set the desired values. As shown in Figure 8.9, we have already assigned the value labels for the "Behavior" variable. For the "School_Type" variable, we set a value of 1 to equal "Public Preschool," a value of 2 to equal "Private Preschool," and a value of 3 to equal "No Preschool." Clicking the "Add" button moves each of the value labels to the list below. Finally, click "OK" to return to the SPSS Variable View screen.

8.4.3.2 Type in Your Values Click the "Data View" tab at the bottom of your screen, as shown in Figure 8.10. Use the whole number references you set earlier



	Behavior	School_Type	Frequency
1	1.00	1.00	12.00
2	2.00	1.00	25.00
3	3.00	1.00	10.00
4	1.00	2.00	6.00
5	2.00	2.00	12.00
6	3.00	2.00	.00
7	1.00	3.00	2.00
8	2.00	3.00	23.00
9	3.00	3.00	10.00
10			

FIGURE 8.10

for the row and column variables. Each possible combination of conditions should exist. Then, enter the corresponding observed frequencies. In our example, row 1 represents a "Behavior" of 1 which is "Poor" and a "School_Type" of 1 which is "Public School." The observed frequency for this condition is 12.

8.4.3.3 Analyze Your Data First, use the "Weight Cases" command to allow the observed frequency variable to reference the category variable. As shown in Figure 8.11, use the pull-down menus to choose "Data" and "Weight Cases . . .".

The default setting is "Do not weight cases." Click the circle next to "Weight cases by" as shown in Figure 8.12. Select the variable with the observed frequencies. Move that variable to the "Frequency Variable:" box by clicking the small arrow button. In our example, we have moved the "Frequency" variable. Finally, click "OK."

As shown in Figure 8.13, use the pull-down menus to choose "Analyze," "Descriptive Statistics," and "Crosstabs . . .".

When the Crosstabs window is open, move the variable that represents the rows to the "Row(s):" box by selecting that variable and clicking the small arrow button next to that box. As shown in Figure 8.14, we have chosen the "School_Type" variable. Then, move the variable that represents the column to the "Column(s):" box. In our example, we have chosen the "Behavior" variable. Next, click the "Statistics . . ." button.

As shown in Figure 8.15, check the box next to "Chi-square" and the box next to "Phi and Cramer's V." Once those boxes are checked, click "Continue" to return to the Crosstabs window. Now, click the "Cells . . ." button.

As shown in Figure 8.16, check the boxes next to "Observed" and "Expected." Then, click "Continue" to return to the Crosstabs window. Finally, click "OK" to perform the analysis.

8.4.3.4 Interpret the Results from the SPSS Output Window The second to fourth output tables from SPSS are of interest in this procedure.

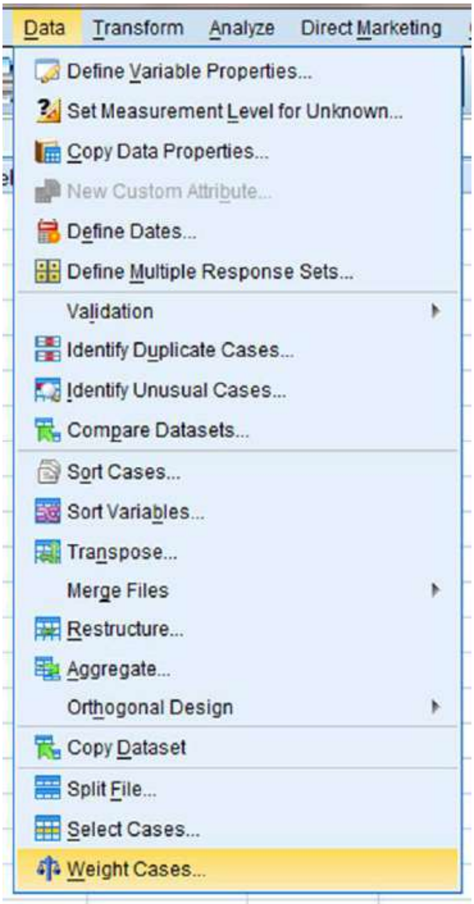


FIGURE 8.11

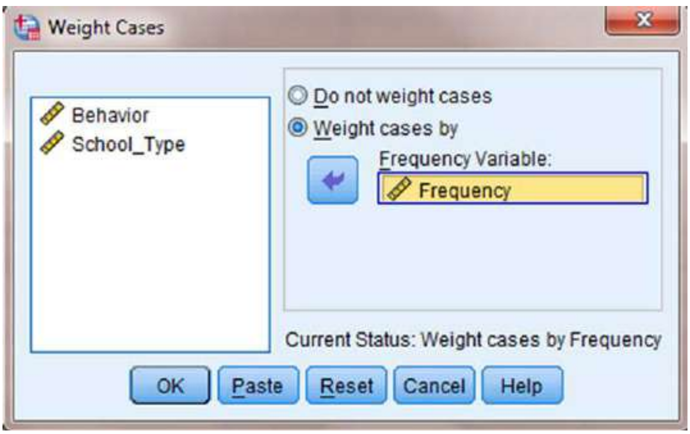


FIGURE 8.12

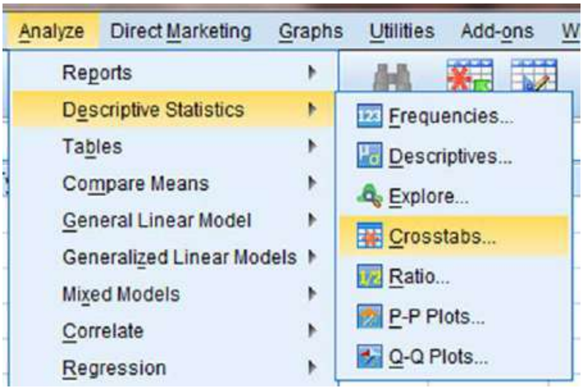


FIGURE 8.13

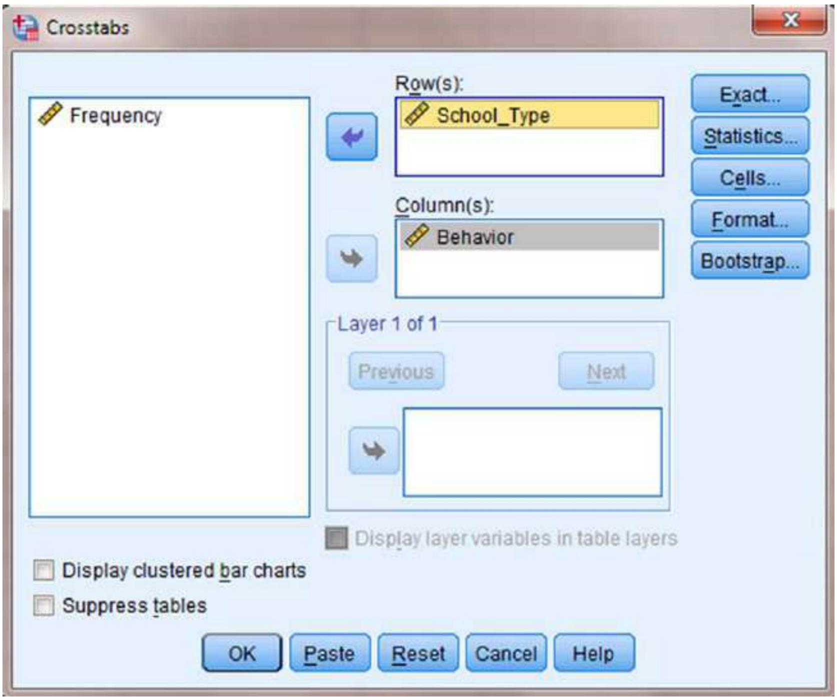


FIGURE 8.14

The second output table (see SPSS Output 8.2a) provides the observed and expected frequencies for each category and the total counts.

The third output table (see SPSS Output 8.2b) provides the χ^2 statistic ($\chi^2 = 11.502$), the degrees of freedom ($df = 4$), and the significance ($p = 0.021$).

The fourth output table (see SPSS Output 8.2c) provides the Cramer's V statistic ($V = 0.240$) to determine the level of association or effect size.

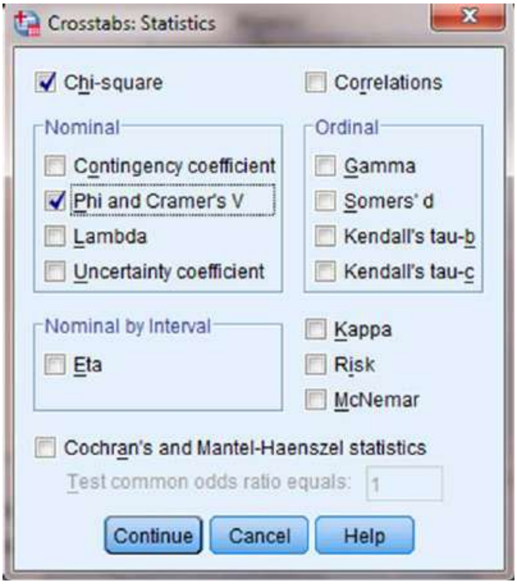


FIGURE 8.15

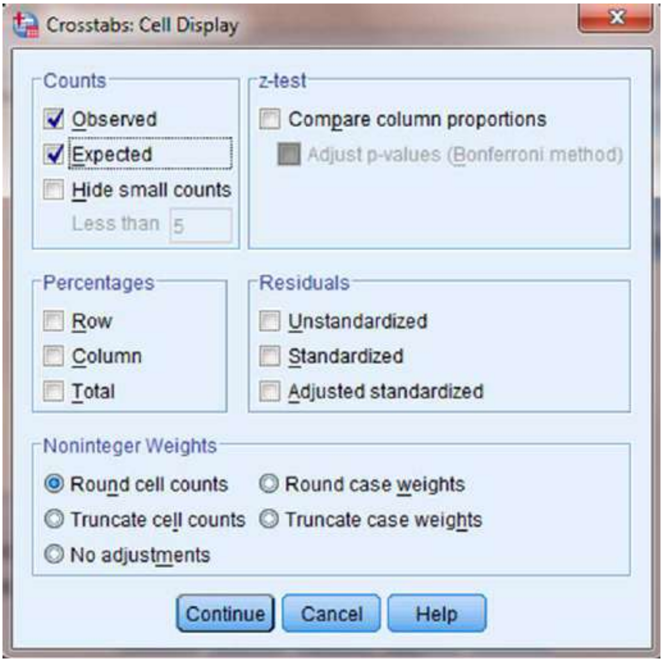


FIGURE 8.16

School_Type * Behavior Crosstabulation

			Behavior			Total
			Poor	Average	Good	
School_Type	Public Preschool	Count	12	25	10	47
		Expected Count	9.4	28.2	9.4	47.0
	Private Preschool	Count	6	12	0	18
		Expected Count	3.6	10.8	3.6	18.0
	No Preschool	Count	2	23	10	35
		Expected Count	7.0	21.0	7.0	35.0
Total		Count	20	60	20	100
		Expected Count	20.0	60.0	20.0	100.0

SPSS OUTPUT 8.2A

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	11.502[a]	4	.021
Likelihood Ratio	16.042	4	.003
Linear-by-Linear Association	3.072	1	.080
N of Valid Cases	100		

a. 2 cells (22.2%) have expected count less than 5. The minimum expected count is 3.60.

SPSS OUTPUT 8.2B

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.339	.021
	Cramer's V	.240	.021
N of Valid Cases		100	

SPSS OUTPUT 8.2C

Based on the results from SPSS, three programs were compared with unequal expected frequencies. The χ^2 goodness-of-fit test was significant ($\chi^2_{(4)} = 11.502$, $p < 0.05$). Based on these results, there is a real association between type of preschool experience children obtained and their behavior in the kindergarten classroom during their first few weeks in school. In addition, the measured effect size presented a medium level of association ($V = 0.240$).

8.5 THE FISHER EXACT TEST

A special case arises if a contingency table's size is 2×2 and at least one expected cell count is less than 5. In this circumstance, SPSS will calculate a Fisher exact test instead of a χ^2-test for independence.

The Fisher exact test is useful for analyzing discrete data obtained from small, independent samples. They can be either nominal or ordinal. It is used when the scores of two independent random samples fall into one of two mutually exclusive classes or obtains one of two possible scores. The results form a 2 × 2 contingency table, as noted earlier.

In this chapter, we will describe how to perform and interpret the Fisher exact test for different samples.

8.5.1 Computing the Fisher Exact Test for 2 × 2 Tables

Compare the 2 × 2 contingency table's one-sided significance with the level of risk, α. Table 8.13 is the 2 × 2 contingency table that is used as the basis for computing Fisher exact test's one-sided significance.

TABLE 8.13

	Group		
Variable	I	II	Combined
+	A	B	A + B
−	C	D	C + D
Total	A + C	B + D	N

The formula for computing the one-sided significance for the Fisher exact test is shown in Formula 8.10. Table B.9 in Appendix B lists the factorials for $n = 0$ to $n = 15$:

$$p = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!} \tag{8.10}$$

If all cell counts are equal to or larger than 5 ($n_i \geq 5$), Daniel (1990) suggested that one use a large sample approximation with the χ^2-test instead of the Fisher exact test.

8.5.2 Sample Fisher Exact Test

A small medical center administered a survey to determine its nurses' attitude of readiness to care for patients. The survey was a 15-item Likert scale with two points positive, two points negative, and a neutral point. The study was conducted to compare the feelings between men and women. Each person was classified according to a total attitude determined by summing the item values on the survey. A maximum positive attitude would be +33 and a maximum negative attitude would be −33.

Table 8.14 and Table 8.15 show the number of men and women who had positive and negative attitudes about how they were prepared. There were four men and six women. Three of the men had positive survey results and only one of the women had a positive survey result.

TABLE 8.14

Participant	Gender	Score	Attitude
1	Male	+30	+
2	Male	+14	+
3	Male	−21	−
4	Male	+22	+
5	Male	+9	+
6	Female	−22	−
7	Female	−13	−
8	Female	−20	−
9	Female	−7	−
10	Female	+19	+
11	Female	−31	−

TABLE 8.15

	Group		
	Men	Women	
Positive	4	1	5
Negative	1	5	6
	5	6	11

We want to determine if there is a difference in attitude between men and women toward their preparation to care for patients. Since the data form a 2 × 2 contingency table and at least one cell has an expected count (see Formula 8.2) of less than 5, the Fisher exact test is a useful procedure to analyze the data and test the hypothesis.

8.5.2.1 State the Null and Research Hypotheses The null hypothesis states that there are no differences between men and women on the attitude survey that measures feelings about the program that teaches care for patients. The alternative hypothesis is that the proportion of men with positive attitudes, P_M, exceed the proportion of women with positive attitudes, P_W.

The hypotheses can be written as follows.

H_O: $P_M = P_W$

H_A: $P_M > P_W$

8.5.2.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α) level, is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

8.5.2.3 Choose the Appropriate Test Statistic The data are obtained from a 2×2 contingency table. Two independent groups were measured on a survey and classified according to two criteria. The classifications were (+) for a positive attitude and (−) for negative attitude. The samples are small, thus requiring nonparametric statistics. We are analyzing data in a 2×2 contingency table and at least one cell has an expected count (see Formula 8.2) of less than 5. Therefore, we will use the Fisher exact test.

8.5.2.4 Compute the Test Statistic First, construct the 2×2 contingency tables for the data in the study and for data that represent a more extreme occurrence than that which was obtained. In this example, there were five men and six women who were in the training program for nurses. Four of the men responded positively to the survey and only one of the women responded positively. The remainder of the people in the study responded negatively.

If we wish to test the null hypothesis statistically, we must consider the possibility of the occurrence of the more extreme outcome that is shown in Table 8.16b. In that table, none of the men responded negatively and none of the women responded positively.

To be shown later are the tables for the statistic. Table 8.16a show the results that occurred in the data collected and Table 8.16b shows the more extreme outcome that could occur.

To test the hypothesis, we first use Formula 8.10 to compute the probability of each possible outcome shown earlier.

TABLE 8.16A

	Group		
	Men	Women	
Positive	4	1	5
Negative	1	5	6
	5	6	11

TABLE 8.16B

	Group		
	Men	Women	
Positive	5	0	5
Negative	0	6	6
	5	6	11

For Table 8.16a,

$$
\begin{aligned}
p_1 &= \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!} \\
&= \frac{(4+1)!(1+5)!(4+1)!(1+5)!}{(11!)(4!)(1!)(1!)(5!)} \\
&= \frac{(5!)(6!)(5!)(6!)}{(11!)(4!)(1!)(1!)(5!)} \\
&= \frac{(120)(720)(120)(720)}{(39{,}916{,}800)(24)(1)(1)(120)} \\
p_1 &= 0.065
\end{aligned}
$$

For Table 8.16b,

$$
\begin{aligned}
p_2 &= \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!} \\
&= \frac{(5+0)!(0+6)!(5+0)!(0+6)!}{(11!)(5!)(0!)(0!)(6!)} \\
&= \frac{(5!)(6!)(5!)(6!)}{(11!)(5!)(0!)(0!)(6!)} \\
&= \frac{(120)(720)(120)(720)}{(39{,}916{,}800)(120)(1)(1)(720)} \\
p_2 &= 0.002
\end{aligned}
$$

The probability is found by adding the two results that were computed earlier:

$$
\begin{aligned}
p &= p_1 + p_2 = 0.065 + 0.002 \\
p &= 0.067
\end{aligned}
$$

8.5.2.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic In the example in this chapter, the probability was computed and compared with the level of risk specified earlier, $\alpha = 0.05$. This computational process involves very large numbers and is aided by the table values. It is recommended that a table of critical values be used when possible.

8.5.2.6 Compare the Obtained Value with the Critical Value The critical value for rejecting the null hypothesis is $\alpha = 0.05$ and the obtained p-value is $p = 0.067$. If the critical value is greater than the obtained value, we must reject the null hypothesis. If the critical value is less than the obtained value, we do not reject the null hypothesis. Since the critical value is less than the obtained value, we do not reject the null hypothesis.

8.5.2.7 Interpret the Results We did not reject the null hypothesis, suggesting that no real difference existed between the attitudes of men and women about their readiness to care for patients. There was, however, a strong trend toward positive feelings on the part of the men and negative feelings on the part of the women. The probability was small, although not significant. This is the type of study that would call for further investigation with other samples to see if this trend was more pronounced. Our analysis does provide some evidence that there is some difference, and if analyzed with a more liberal critical value such as $\alpha = 0.10$, this statistical test would show significance.

Since the Fisher exact test was not statistically significant ($p > \alpha$), we may not have an interest in the strength of the association between the two variables. However, a researcher wishing to replicate the study may wish to know that strength of association.

The effect size is a measure of association between two variables. For the Fisher exact test, which has a 2×2 contingency table, we determine the effect size using the phi (ϕ) coefficient (see Formula 8.8). For $n = 11$ and $\chi^2 = 4.412$ (calculation for χ^2 not shown), we use Formula 8.8 to determine ϕ:

$$\phi = \sqrt{\frac{4.412}{11}}$$
$$= \sqrt{0.401}$$
$$\phi = 0.633$$

Our effect size, the phi (ϕ) coefficient, is 0.633. This value indicates a strong level of association between the two variables. What is more, a replication of the study may be worth the effort.

8.5.2.8 Reporting the Results When reporting the findings, include the table that shows the actual reported frequencies, including all marginal frequencies. In addition, report the p-value and its relationship to the critical value.

For this example, Table 8.15 would be reported. The obtained significance, $p = 0.067$, was greater than the critical value, $\alpha = 0.05$. Therefore, we did not reject the null hypothesis, suggesting that there was no difference between men and women on the attitude survey that measures feelings about the program that teaches care for patients.

8.5.3 Performing the Fisher Exact Test Using SPSS

As noted earlier, SPSS performs a Fisher exact test instead of a χ^2-test for independence if the contingency table's size is 2×2 and at least one expected cell count is less than 5. In other words, to perform a Fisher exact test, use the same method you used for a χ^2-test for independence.

SPSS Outputs 8.3a and 8.3b provide the SPSS output for the sample Fisher exact test computed earlier. Note that all four expected counts were less than 5. In addition, the one-sided significance is $p = 0.067$.

SPSS Output 8.3c provides the effect size for the association. Since the association was not statistically significant ($p > \alpha$), the effect size ($\phi = 0.633$) was not of interest to this study.

Attitude * Gender Crosstabulation

			Gender		Total
			male	female	
Attitude	positive attitude	Count	4	1	5
		Expected Count	2.3	2.7	5.0
	negative attitude	Count	1	5	6
		Expected Count	2.7	3.3	6.0
Total		Count	5	6	11
		Expected Count	5.0	6.0	11.0

SPSS OUTPUT 8.3A

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.412[b]	1	.036		
Continuity Correction[a]	2.228	1	.136		
Likelihood Ratio	4.747	1	.029		
Fisher's Exact Test				.080	.067
Linear-by-Linear Association	4.011	1	.045		
N of Valid Cases	11				

a. Computed only for a 2x2 table

b. 4 cells (100.0%) have expected count less than 5. The minimum expected count is 2.27.

SPSS OUTPUT 8.3B

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.633	.036
	Cramer's V	.633	.036
N of Valid Cases		11	

SPSS OUTPUT 8.3C

8.6 EXAMPLES FROM THE LITERATURE

Listed are varied examples of the nonparametric procedures described in this chapter. We have summarized each study's research problem and researchers' rationale(s) for choosing a nonparametric approach. We encourage you to obtain these studies if you are interested in their results.

Duffy and Sedlacek (2007) examined the surveys of 3570 1st-year college students regarding the factors they deemed most important to their long-term career choice. χ^2 analyses were used to assess work value differences (social, extrinsic, prestige, and intrinsic) in gender, parental income, race, and educational aspirations. The researchers used χ^2-tests for independence since these data were frequencies of nominal items.

Ferrari et al. (2006) studied the effects of leadership experience in an academic honor society on later employment and education. Drawing from honor society alumni, the researchers compared leaders with nonleaders on various aspects of their graduate education or employment. Since most data were frequencies of nominal items, the researchers used χ^2-test for independence.

Helsen et al. (2006) analyzed the correctness of assistant referees' offside judgments during the final round of the FIFA 2002 World Cup. Specifically, they use digital video technology to examine situations involving the viewing angle and special position of a moving object. They used a χ^2 goodness-of-fit test to determine if the ratio of correct to incorrect decisions and the total number of offside decisions were uniformly distributed throughout six 15-min intervals. They also used a χ^2 goodness-of-fit test to determine if flag errors versus nonflag errors led to a judgment bias.

Shorten et al. (2005) analyzed a survey of 34 academic libraries in the United States and Canada that use the Dewey decimal classification (DDC). They wished to determine why these libraries continue using DDC and if they have considered reclassification. Some of the survey questions asked participants to respond to a reason by selecting "more important," "less important," or "not a reason at all." Responses were analyzed with a χ^2 goodness-of-fit test to examine responses for consensus among libraries.

Rimm-Kaufman and Zhang (2005) studied the communication between fathers of "at-risk" children with their preschool and kindergarten schools. Specifically, they examined frequency, characteristics, and predictors of communication based on family sociodemographic characteristics. When analyzing frequencies, they used χ^2-tests. However, when cells contained frequencies of zero, they used a Fisher exact test.

Enander et al. (2007) investigated a newly employed inspection method for self-certification of environmental and health qualities in automotive refinishing facilities. They focused on occupational health and safety, air pollution control, hazardous waste management, and wastewater discharge. A Fisher exact test was used to analyze 2×2 tables with relatively small observed cell frequencies.

To examine the clinical problems of sexual abuse, Mansell et al. (1998) compared children with developmental disabilities to children without developmental disabilities. Categorical data were analyzed with χ^2-tests of independence; however, the Yates' continuity corrections were used when cell counts exhibited less than the minimum expected count needed for the χ^2-test.

8.7 SUMMARY

Nominal, or categorical, data sometimes need analyses. In such cases, you may be seeking to determine if the data statistically match some known or expected set of frequencies. Or, you may wish to determine if two or more categories are statistically independent. In either case, nominal data can be analyzed with a nonparametric procedure.

In this chapter, we presented three procedures for examining nominal data: chi-square (χ^2) goodness of fit, χ^2-test for independence, and the Fisher exact test. We also explained how to perform the procedures using SPSS. Finally, we offered varied examples of these nonparametric statistics from the literature. In the next chapter, we will describe how to determine if a series of events occurred randomly.

8.8 PRACTICE QUESTIONS

1. A police department wishes to compare the average number of monthly robberies at four locations in their town. Use equal categories in order to identify one or more concentrations of robberies. The data are presented in Table 8.17.
Use a χ^2 goodness-of-fit test with $\alpha = 0.05$ to determine if the robberies are concentrated in one or more of the locations. Report your findings.

TABLE 8.17

	Average monthly robberies
Location 1	15
Location 2	10
Location 3	19
Location 4	16

2. The χ^2 goodness-of-fit test serves as a useful tool to ensure that statistical samples approximately match the desired stratification proportions of the population from which they are drawn.

A researcher wishes to determine if her randomly drawn sample matches the racial stratification of school age children. She used the most recent U.S. Census data, which was from 2001. The racial composition of her sample and the 2001 U.S. Census proportions are displayed in Table 8.18.

Use a χ^2 goodness-of-fit test with $\alpha = 0.05$ to determine if the researcher's sample matches the proportions reported by the U.S. Census. Report your findings.

TABLE 8.18

Race	Frequency of race from the researcher's randomly drawn sample	Racial percentage of U.S. school children based on the 2001 U.S. Census (%)
White	57	72
Black	21	20
Asian, Hispanic, or Pacific Islander	14	8

3. A researcher wishes to determine if there is an association between the level of a teacher's education and his/her job satisfaction. He surveyed 158 teachers. The frequencies of the corresponding results are displayed in Table 8.19.

TABLE 8.19

	Teacher education level (observed)			
	Bachelor's degree	Master's degree	Post-Master's degree	Row totals
Satisfied	60	41	19	120
Unsatisfied	10	13	15	38
Column totals	70	54	34	158

First, use a χ^2-test for independence with $\alpha = 0.05$ to determine if there is an association between level of education and job satisfaction. Then, determine the effect size for the association. Report your findings.

4. A professor gave her class a 10-item survey to determine the students' satisfaction with the course. Survey question responses were measured using a five-point Likert scale. The survey had a score range from $+20$ to -20. Table 8.20 shows the scores of the students in a class of 13 students who rated the professor.

TABLE 8.20

Participant	Gender	Score	Satisfaction
1	Male	+12	+
2	Male	+6	+
3	Male	−5	−
4	Male	−10	−
5	Male	+17	+
6	Male	+4	+
7	Female	−2	−
8	Female	−13	−
9	Female	+10	+
10	Female	−8	−
11	Female	−11	−
12	Female	−4	−
13	Female	−14	−

Use a Fisher exact test with $\alpha = 0.05$ to determine if there is an association between gender and course satisfaction of the professor's class. Then, determine the effect size for the association. Report your findings.

8.9 SOLUTIONS TO PRACTICE QUESTIONS

1. The results from the analysis are displayed in SPSS Outputs 8.4a and 8.4b.

Location

	Observed N	Expected N	Residual
Location 1	15	15.0	.0
Location 2	10	15.0	-5.0
Location 3	19	15.0	4.0
Location 4	16	15.0	1.0
Total	60		

SPSS OUTPUT 8.4A

Test Statistics

	Location
Chi-Square	2.800[a]
df	3
Asymp. Sig.	.423

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 15.0.

SPSS OUTPUT 8.4B

According to the data, the results from the χ^2 goodness-of-fit test were not significant ($\chi^2_{(3)} = 2.800$, $p > 0.05$). Therefore, no particular location displayed a significantly higher or lower number of robberies.

2. The results from the analysis are displayed in SPSS Outputs 8.5a and 8.5b.

Race

	Observed N	Expected N	Residual
White	57	66.2	-9.2
Black	21	18.4	2.6
Asian, Hispanic, or Pacific Islander	14	7.4	6.6
Total	92		

SPSS OUTPUT 8.5A

According to the data, the results from the χ^2 goodness-of-fit test were significant ($\chi^2_{(2)} = 7.647$, $p < 0.05$). Therefore, the sample's racial stratification approximately matches the U.S. Census racial composition of school aged children in 2001.

Test Statistics

	Race
Chi-Square	7.647[a]
df	2
Asymp. Sig.	.022

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 7.4.

SPSS OUTPUT 8.5B

3. The results from the analysis are displayed in SPSS Outputs 8.6a–8.6c.

Job_Satisfaction * Education_Level Crosstabulation

			Education_Level			
			Bachelor Degree	Master Degree	Post-Master Degree	Total
Job_Satisfaction	Satisfied	Count	60	41	19	120
		Expected Count	53.2	41.0	25.8	120.0
	Unsatisfied	Count	10	13	15	38
		Expected Count	16.8	13.0	8.2	38.0
Total		Count	70	54	34	158
		Expected Count	70.0	54.0	34.0	158.0

SPSS OUTPUT 8.6A

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	11.150[a]	2	.004
Likelihood Ratio	10.638	2	.005
Linear-by-Linear Association	10.593	1	.001
N of Valid Cases	158		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 8.18.

SPSS OUTPUT 8.6B

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.266	.004
	Cramer's V	.266	.004
N of Valid Cases		158	

SPSS OUTPUT 8.6C

As seen in SPSS Output 8.6a, none of the cells had an expected count of less than 5. Therefore, the χ^2-test was indeed an appropriate analysis. Concerning effect size, the size of the contingency table was larger than 2 × 2. Therefore, a Cramer's *V* was appropriate.

According to the data, the results from the χ^2-test for independence were significant ($\chi^2_{(2)} = 11.150$, $p < 0.05$). Therefore, the analysis provides evidence that teacher education level differentiates between individuals based on job satisfaction. In addition, the effect size ($V = 0.266$) indicated a medium level of association between the variables.

4. The results from the analysis are displayed in SPSS Outputs 8.7a–8.7c.

Satisfaction * Gender Crosstabulation

			Gender		Total
			male	female	
Satisfaction	positive	Count	4	1	5
		Expected Count	2.3	2.7	5.0
	negative	Count	2	6	8
		Expected Count	3.7	4.3	8.0
Total		Count	6	7	13
		Expected Count	6.0	7.0	13.0

SPSS OUTPUT 8.7A

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.745[a]	1	.053		
Continuity Correction[b]	1.859	1	.173		
Likelihood Ratio	3.943	1	.047		
Fisher's Exact Test				.103	.086
Linear-by-Linear Association	3.457	1	.063		
N of Valid Cases	13				

a. 4 cells (100.0%) have expected count less than 5. The minimum expected count is 2.31.

b. Computed only for a 2x2 table

SPSS OUTPUT 8.7B

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.537	.053
	Cramer's V	.537	.053
N of Valid Cases		13	

SPSS OUTPUT 8.7C

As seen in SPSS Output 8.7a, all of the cells had an expected count of less than 5. Therefore, the Fisher exact test was an appropriate analysis. Concerning effect size, the size of the contingency table was 2 × 2. Therefore, a phi (ϕ) coefficient was appropriate.

According to the data, the results from the Fisher exact test were not significant ($p = 0.086$) based on $\alpha = 0.05$. Therefore, the analysis provides evidence that no association exists between gender and course satisfaction of the professor's class. In addition, the effect size ($\phi = 0.537$) was not of interest to this study due to the lack of significant association between variables.

CHAPTER 9

TEST FOR RANDOMNESS: THE RUNS TEST

9.1 OBJECTIVES

In this chapter, you will learn the following items:

- How to use a runs test to analyze a series of events for randomness.
- How to perform a runs test using SPSS®.

9.2 INTRODUCTION

Every investor wishes he or she could predict the behavior of a stock's performance. Is there a pattern to a stock's gain/loss cycle or are the events random? One could make a defensible argument to that question with an analysis of randomness.

The runs test (sometimes called a Wald–Wolfowitz runs test) is a statistical procedure for examining a series of events for randomness. This nonparametric test has no parametric equivalent. In this chapter, we will describe how to perform and interpret a runs test for both small samples and large samples. We will also explain how to perform the procedure using SPSS. Finally, we offer varied examples of these nonparametric statistics from the literature.

9.3 THE RUNS TEST FOR RANDOMNESS

The runs test seeks to determine if a series of events occur randomly or are merely due to chance. To understand a run, consider a sequence represented by two symbols, A and B. One simple example might be several tosses of a coin where A = heads and B = tails. Another example might be whether an animal chooses to eat first or drink first. Use A = eat and B = drink.

The first steps are to list the events in sequential order and count the number of runs. A *run* is a sequence of the same event written one or more times. For example, compare two event sequences. The first sequence is written AAAAAABBBBBB.

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

Then, separate the sequence into same groups as shown in Figure 9.1. There are two runs in this example, $R = 2$. This is a trend pattern in which events are clustered and it does not represent random behavior.

AAAAAA	BBBBBB
1	2

FIGURE 9.1

Consider a second event sequence written ABABABABABAB. Again, separate the events into same groups (see Fig. 9.2) to determine the number of runs. There are 12 runs in this example, $R = 12$. This is a cyclical pattern and does not represent random behavior either. As illustrated in the two examples earlier, too few or too many runs lack randomness.

A	B	A	B	A	B	A	B	A	B	A	B
1	2	3	4	5	6	7	8	9	10	11	12

FIGURE 9.2

A run can also describe how a sequence of events occurs in relation to a custom value. Use two symbols, such as A and B, to define whether an event exceeds or falls below the custom value. A simple example may reference the freezing point of water where A = temperatures above 0°C and B = temperatures below 0°C. In this example, simply list the events in order and determine the number of runs as described earlier.

After the number of runs is determined, it must be examined for significance. We may use a table of critical values (see Table B.10 in Appendix B). However, if the numbers of values in each sample, n_1 or n_2, exceed those available from the table, then a large sample approximation may be performed. For large samples, compute a z-score and use a table with the normal distribution (see Table B.1 in Appendix B) to obtain a critical region of z-scores. Formula 9.1, Fomula 9.2, Formula 9.3, Formula 9.4, and Formula 9.5 are used to find the z-score of a runs test for large samples:

$$\bar{x}_R = \frac{2n_1n_2}{n_1 + n_2} + 1 \tag{9.1}$$

where $\bar{x}_R$ is the mean value of runs, n_1 is the number of times the first event occurred, and n_2 is the number of times the second event occurred;

$$s_R = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \tag{9.2}$$

where s_R is the standard deviation of runs;

$$z^* = \frac{R + h - \bar{x}_R}{s_R} \tag{9.3}$$

where z^* is the z-score for a normal approximation of the data, R is the number of runs, and h is the correction for continuity, ±0.5, where

$$h = +0.5 \text{ if } R < (2n_1n_2/(n_1+n_2)+1) \tag{9.4}$$

and

$$h = -0.5 \text{ if } R > (2n_1n_2/(n_1+n_2)+1) \tag{9.5}$$

9.3.1 Sample Runs Test (Small Data Samples)

The following study seeks to examine gender bias in science instruction. A male science teacher was observed during a typical class discussion. The observer noted the gender of the student that the teacher called on to answer a question. In the course of 15 min, the teacher called on 10 males and 10 females. The observer noticed that the science teacher called on equal numbers of males and females, but he wanted to examine the data for a pattern. To determine if the teacher used a random order to call on students with regard to gender, he used a runs test for randomness. Using M for male and F for female, the sequence of student recognition by the teacher is MFFMFMFMFFMFFFMMFMMM.

9.3.1.1 State the Null and Research Hypotheses The null hypothesis states that the sequence of events is random. The research hypothesis states that the sequence of events is not random.

The null hypothesis is

H_O: The sequence in which the teacher calls on males and females is random.

The research hypothesis is

H_A: The sequence in which the teacher calls on males and females is not random.

9.3.1.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

9.3.1.3 Choose the Appropriate Test Statistic The observer is examining the data for randomness. Therefore, he is using a runs test for randomness.

9.3.1.4 Compute the Test Statistic First, determine the number of runs, R. It is helpful to separate the events as shown in Figure 9.3. The number of runs in the sequence is $R = 13$.

M	FF	M	F	M	F	M	FF	M	FFF	MM	F	MMM
1	2	3	4	5	6	7	8	9	10	11	12	13

FIGURE 9.3

9.3.1.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Since the sample sizes are small, we refer to Table B.10 in Appendix B, which lists the critical values for the runs test. There were 10 males (n_1) and 10 females (n_2). The critical values are found on the table at the point for $n_1 = 10$ and $n_2 = 10$. We set $\alpha = 0.05$. The critical region for the runs test is $6 < R < 16$. If the number of runs, R, is 6 or less, or 16 or greater, we reject our null hypothesis.

9.3.1.6 Compare the Obtained Value with the Critical Value We found that $R = 13$. This value is within our critical region ($6 < R < 16$). Therefore, we do not reject the null hypothesis.

9.3.1.7 Interpret the Results We did not reject the null hypothesis, suggesting that the sequence of events is random. Therefore, we can state that the order in which the science teacher calls on males and females is random.

9.3.1.8 Reporting the Results The reporting of results for the runs test should include such information as the sample sizes for each group, the number of runs, and the p-value with respect to α.

For this example, the runs test indicated that the sequence was random ($R = 13$, $n_1 = 10$, $n_2 = 10$, $p > 0.05$). Therefore, the study provides evidence that the science teacher was demonstrating no gender bias.

9.3.2 Performing the Runs Test Using SPSS

We will analyze the data from the example earlier using SPSS.

9.3.2.1 Define Your Variables First, click the "Variable View" tab at the bottom of your screen. Then, type the names of your variables in the "Name" column. As seen in Figure 9.4, we call our variable "gender."

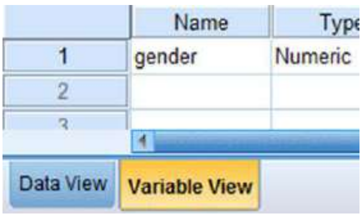


FIGURE 9.4

Next, we establish a grouping variable to differentiate between males and females. When establishing a grouping variable, it is often easiest to assign each group a whole number value. As shown in Figure 9.5, our groups are "male" and "female." First, we select the "Values" column and click the gray square. Then, we set a value of 0 to represent "male" and a value of 2 to represent "female." We use the "Add" button to move each of the value labels to the list. We did not choose the value of 1 since we will use it in step 3 as a reference (custom value) to compare the events. Once we finish, we click the "OK" button to return to the SPSS Variable View.

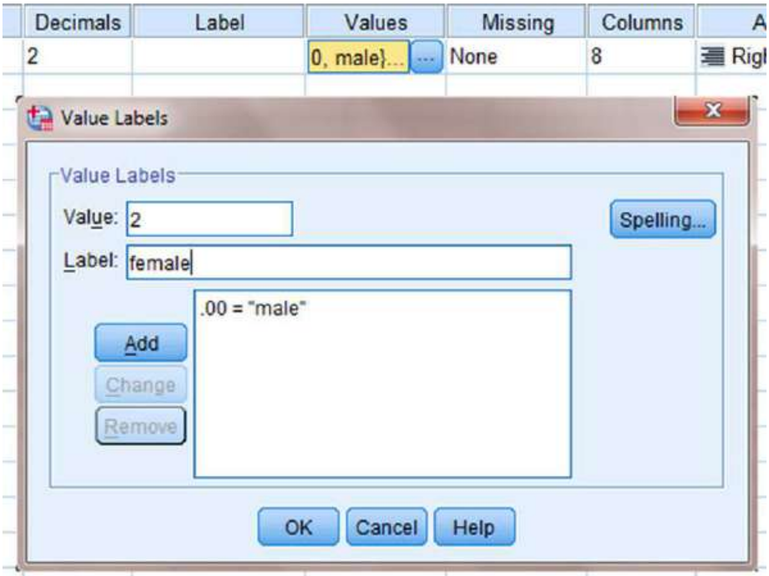


FIGURE 9.5

9.3.2.2 Type in Your Values Click the “Data View” tab at the bottom of your screen (see Figure 9.6). Type the values into the column in the same order they occurred. Remember that we type 0 for “male” and 2 for “female.”

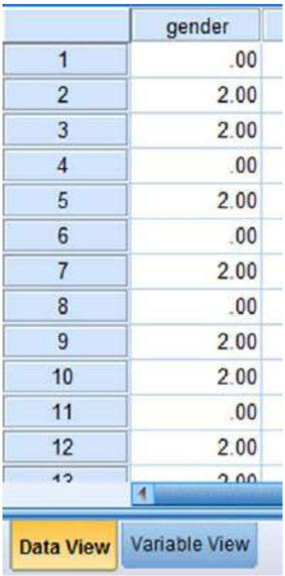


FIGURE 9.6

9.3.2.3 Analyze Your Data As shown in Figure 9.7, use the pull-down menus to choose “Analyze,” “Nonparametric Tests,” “Legacy Dialogs,” and “Runs. . . .”

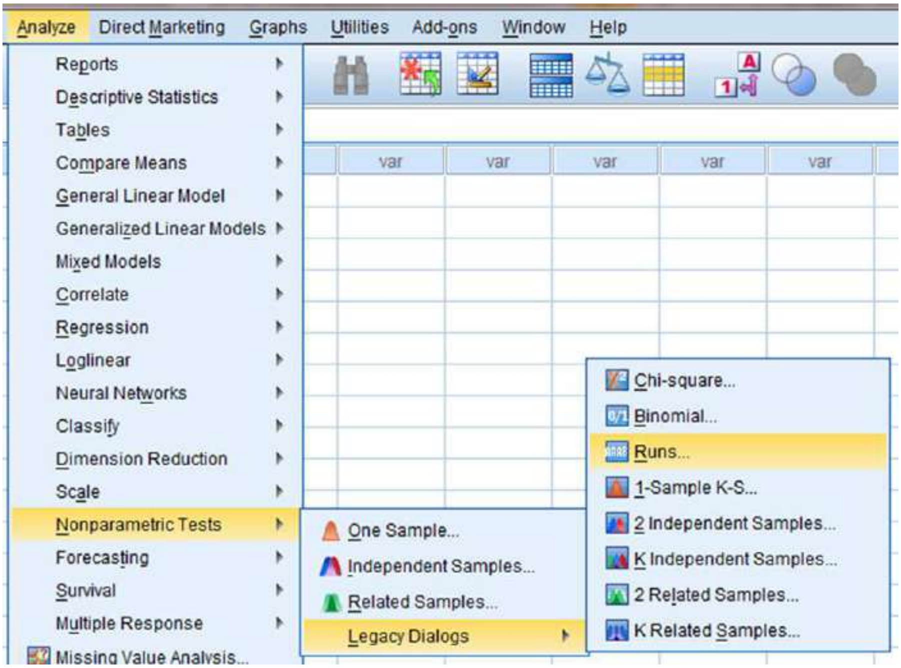


FIGURE 9.7

The runs test required a reference point to compare the events. As shown in Figure 9.8 under "Cut Point," uncheck "Median" and check the box next to "Custom:." Type a value in the box that is between the events' assigned values. For our example, we used 0 and 2 for the events' values, so type a custom value of 1. Next, select the

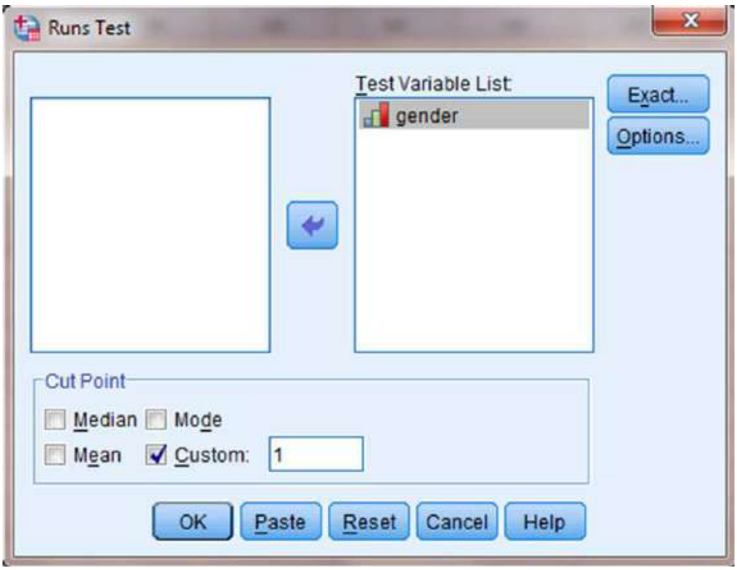


FIGURE 9.8

variable and use the arrow button to place it with your data values in the box labeled "Test Variable List." In our example, we choose the variable "gender." Finally, click "OK" to perform the analysis.

9.3.2.4 ***Interpret the Results from the SPSS Output Window*** The runs test output table (see SPSS Output 9.1) returns the total number of observations ($N = 20$) and the number of runs ($R = 13$). SPSS also calculates the z-score ($z^* = 0.689$) and the two-tailed significance ($p = 0.491$).

Runs Test

	gender
Test Value[a]	1.0000
Total Cases	20
Number of Runs	13
Z	.689
Asymp. Sig. (2-tailed)	.491

a. User-specified.

SPSS OUTPUT 9.1

9.3.2.5 ***Determine the Observation Frequencies for Each Event*** In order to determine the number of observations for each event, an additional set of steps is required. As shown in Figure 9.9, use the pull-down menus to choose "Analyze," "Descriptive Statistics," and "Frequencies. . . ."

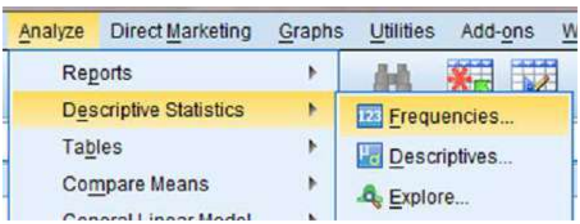


FIGURE 9.9

Next, use the arrow button to place the variable with your data values in the box labeled "Variable(s):" as shown in Figure 9.10. Like before, we choose the variable "gender." Finally, click "OK" to perform the analysis.

The second output table (see SPSS Output 9.2) displays the frequencies for each event. Based on the results from SPSS, the runs test indicated that the sequence was random ($R = 13$, $n_1 = 10$, $n_2 = 10$, $p > 0.05$). Therefore, the science teacher randomly chose between males and females when calling on students.

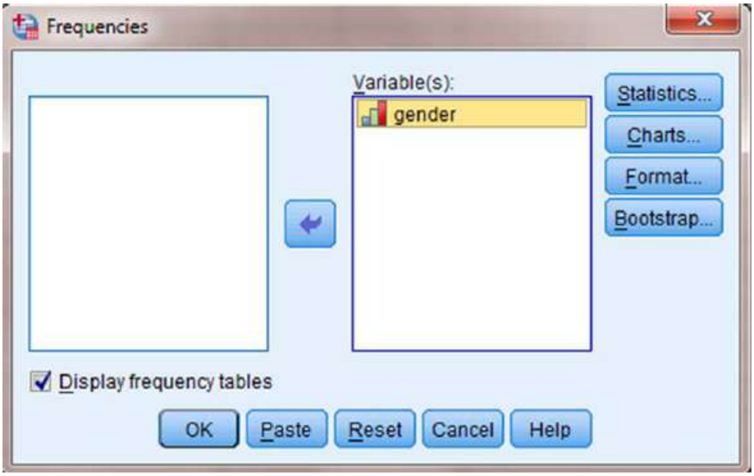


FIGURE 9.10

gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	10	50.0	50.0	50.0
	female	10	50.0	50.0	100.0
	Total	20	100.0	100.0	

SPSS OUTPUT 9.2

9.3.3 Sample Runs Test (Large Data Samples)

The previous study investigating gender bias was replicated. This time, however, a different male teacher was observed and the observation occurred over a longer period of time. As before, the observer noted the gender of the student that the teacher called on to answer a question. In the course of 30 min, the teacher called on 23 males and 14 females. We will once again examine the data for a pattern and use a runs test to examine student recognition with respect to gender. This time, however, we will use a large sample approximation since at least one sample size is large. Using M for male and F for female, the sequence of student recognition by the teacher is FFMMFFFMFFFMMFMMMMFMMMMMMFMFMMFMMMMMF.

9.3.3.1 State the Null and Research Hypotheses The null hypothesis states that the sequence of events is random. The research hypothesis states that the sequence of events is not random.

The null hypothesis is

H_O: The sequence in which the teacher calls on males and females is random.

The research hypothesis is

H_A: The sequence in which the teacher calls on males and females is not random.

9.3.3.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

9.3.3.3 Choose the Appropriate Test Statistic The observer is examining the data for randomness. Therefore, he is using a runs test for randomness.

9.3.3.4 Compute the Test Statistic First, determine the number of runs, R. It is helpful to separate the events as shown in Figure 9.11. The number of runs in the sequence is $R = 17$.

FF	MM	FFF	M	FFF	MM	F	MMMM	F
1	2	3	4	5	6	7	8	9

MMMMMMM	F	M	F	MM	F	MMMMM	F
10	11	12	13	14	15	16	17

FIGURE 9.11

Our number of values exceeds those available from our critical value table for the runs test (Table B.10 in Appendix B is limited to $n_1 \leq 20$ and $n_2 \leq 20$). Therefore, we will find a z-score for our data using a normal approximation. First, we must find the mean $\bar{x}_R$ and the standard deviation s_R for the data:

$$\bar{x}_R = \frac{2n_1n_2}{n_1+n_2}+1 = \frac{2(23)(14)}{23+14}+1$$
$$= \frac{644}{37}+1 = 17.4+1$$
$$\bar{x}_R = 18.4$$

and

$$s_R = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}} = \sqrt{\frac{2(23)(14)(2(23)(14)-23-14)}{(23+14)^2(23+14-1)}}$$
$$= \sqrt{\frac{(644)(644-23-14)}{(37)^2(36)}} = \sqrt{\frac{(644)(607)}{(1369)(36)}}$$
$$= \sqrt{\frac{390,908}{49,284}} = \sqrt{7.932}$$
$$s_R = 2.816$$

Next, we calculate a z-score. We use the correction for continuity, mean, standard deviation, and number of runs ($R = 17$) to calculate a z-score. The only value that we still need is the correction for continuity h. Recall that $h = +0.5$ if $R < (2n_1n_2/(n_1 + n_2) + 1)$, and $h = -0.5$ if $R > (2n_1n_2/(n_1 + n_2) + 1)$. In our example, $2n_1n_2/(n_1 + n_2) + 1 = 2(23)(14)/(23 + 14) + 1 = 18.4$. Since $17 < 18.4$, we choose $h = +0.5$.

Now, we use our z-score formula with correction for continuity:

$$z^* = \frac{R + h - \overline{x}_R}{s_R} = \frac{17 + (+0.5) - 18.4}{2.816}$$

$$z^* = -0.3196$$

9.3.3.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic Table B.1 in Appendix B is used to establish the critical region of z-scores. For a two-tailed test with $\alpha = 0.05$, we must not reject the null hypothesis if $-1.96 \leq z^* \leq 1.96$.

9.3.3.6 Compare the Obtained Value with the Critical Value We find that z^* is within the critical region of the distribution, $-1.96 \leq -0.3196 \leq 1.96$. Therefore, we do not reject the null hypothesis. This suggests that the order in which the science teacher calls on males and females is random.

9.3.3.7 Interpret the Results We did not reject the null hypothesis, suggesting that the sequence of events is random. Therefore, our data indicate that the order in which the science teacher calls on males and females is random.

9.3.3.8 Reporting the Results Based on our analysis, the runs test indicated that the sequence was random ($R = 17$, $n_1 = 23$, $n_2 = 14$, $p > 0.05$). Therefore, the study provides evidence that the science teacher was demonstrating no gender bias.

9.3.4 Sample Runs Test Referencing a Custom Value

The science teacher in the earlier example wishes to examine the pattern of an "at-risk" student's weekly quiz performance. A passing quiz score is 70. Sometimes the student failed and other times he passed. The teacher wished to determine if the student's performance is random or not. Table 9.1 shows the student's weekly quiz scores for a 12-week period.

TABLE 9.1

Week	Student quiz scores
1	65
2	55
3	95
4	15
5	75
6	65
7	80
8	75
9	60
10	55
11	75
12	80

9.3.4.1 State the Null and Research Hypotheses The null hypothesis states that the sequence of events is random. The research hypothesis states that the sequence of events is not random.

The null hypothesis is

H_O: The sequence in which the student passes and fails a weekly science quiz is random.

The research hypothesis is

H_A: The sequence in which the student passes and fails a weekly science quiz is not random.

9.3.4.2 Set the Level of Risk (or the Level of Significance) Associated with the Null Hypothesis The level of risk, also called an alpha (α), is frequently set at 0.05. We will use $\alpha = 0.05$ in our example. In other words, there is a 95% chance that any observed statistical difference will be real and not due to chance.

9.3.4.3 Choose the Appropriate Test Statistic The observer is examining the data for randomness. Therefore, he is using a runs test for randomness.

9.3.4.4 Compute the Test Statistic The custom value is 69.9 since a passing quiz score is 70. We must identify which quiz scores fall above the custom score and which quiz scores fall below it. As shown in Table 9.2, we mark the quiz scores that fall above the custom score with + and the quiz scores that fall below with −.

Then, we count the number of runs, R. The number of runs in the sequence earlier is $R = 8$.

TABLE 9.2

Week	Student quiz scores	Relation to custom score
1	65	−
2	55	−
3	95	+
4	15	−
5	75	+
6	65	−
7	80	+
8	75	+
9	60	−
10	55	−
11	70	+
12	80	+

9.3.4.5 Determine the Value Needed for Rejection of the Null Hypothesis Using the Appropriate Table of Critical Values for the Particular Statistic
Since the sample sizes are small, we refer to Table B.10 in Appendix B, which lists the critical values for the runs test. The critical values are found on the table at the point for $n_1 = 6$ and $n_2 = 6$. We set $\alpha = 0.05$. The critical region for the runs test is $3 < R < 11$. If the number of runs, R, is 3 or less, or 11 or greater, we reject our null hypothesis.

9.3.4.6 Compare the Obtained Value with the Critical Value We found that $R = 8$. This value is within our critical region ($3 < R < 11$). Therefore, we must not reject the null hypothesis.

9.3.4.7 Interpret the Results We did not reject the null hypothesis, suggesting that the sequence of events is random. Therefore, we can state that based on a passing score of 70, the student's weekly science quiz performance is random.

9.3.4.8 Reporting the Results For this example, the runs test indicated that the sequence was random ($R = 8$, $n_1 = 6$, $n_2 = 6$, $p > 0.05$). Therefore, the evidence suggests that the pattern of the student's weekly science quiz performance is random in terms of achieving a passing score of 70.

9.3.5 Performing the Runs Test for a Custom Value Using SPSS

We will analyze the data from the earlier example using SPSS.

9.3.5.1 Define Your Variables First, click the "Variable View" tab at the bottom of your screen. As shown in Figure 9.12, type the names of your variables in the "Name" column. We call our variable "Quiz."

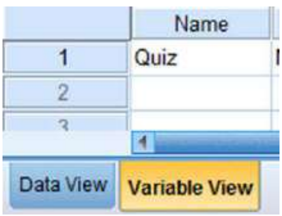


FIGURE 9.12

9.3.5.2 Type in Your Values Click the "Data View" tab at the bottom of your screen, as shown in Figure 9.13. Type the values into the column in the same order they occurred.

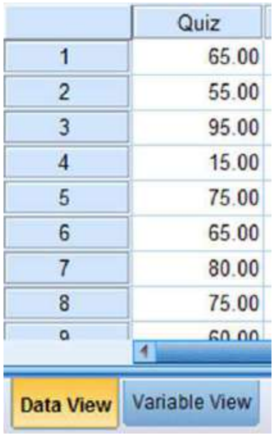


FIGURE 9.13

9.3.5.3 Analyze Your Data As shown in Figure 9.14, use the pull-down menus to choose “Analyze,” “Nonparametric Tests,” “Legacy Dialogs,” and “Runs. . . .”

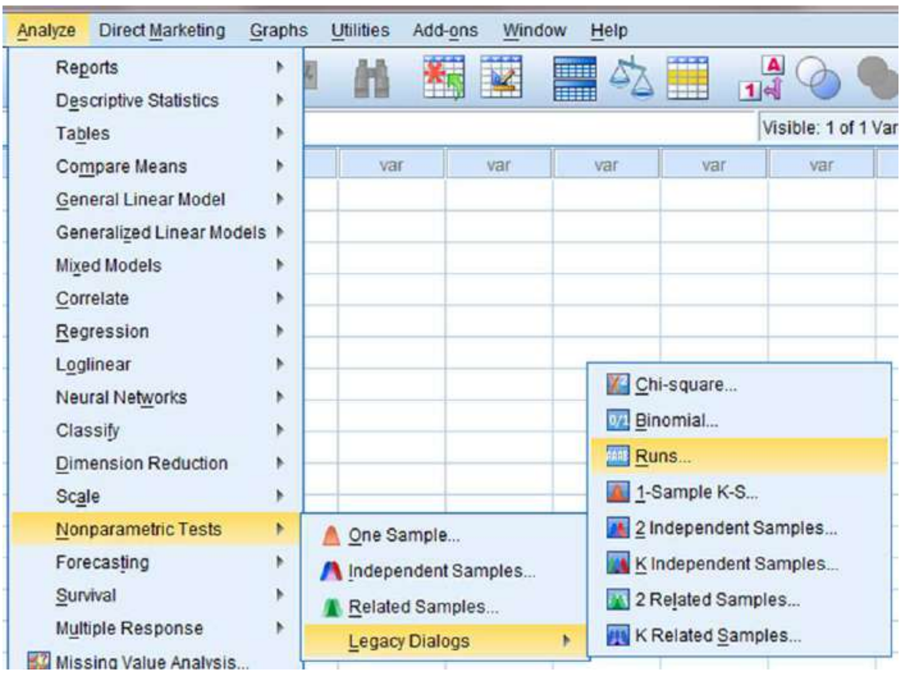


FIGURE 9.14

As shown in Figure 9.15 under “Cut Point,” uncheck “Median” and check the box next to “Custom:.” Type the custom value in the box. For our example, we use a custom value of 69.9. Next, use the arrow button to place the variable with your data values in the box labeled “Test Variable List.” In our example, we choose the variable “Quiz.” Finally, click the “OK” button to perform the analysis.

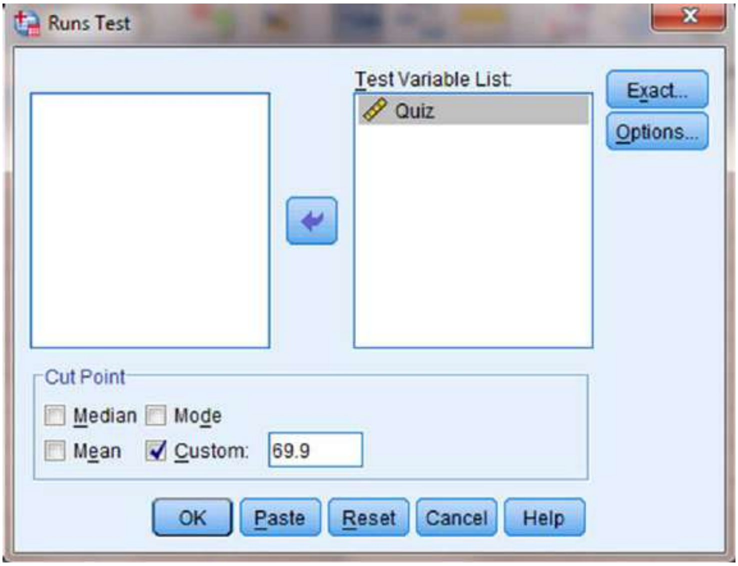


FIGURE 9.15

9.3.5.4 Interpret the Results from the SPSS Output Window The runs test output table (see SPSS Output 9.3) returns the custom test value (69.9), the total number of observations ($N = 12$), and the number of runs ($R = 8$). SPSS also calculates the z-score ($z^* = 0.303$) and the two-tailed significance ($p = 0.762$).

Runs Test

	Quiz
Test Value[a]	69.9000
Total Cases	12
Number of Runs	8
Z	.303
Asymp. Sig. (2-tailed)	.762

a. User-specified.

SPSS OUTPUT 9.3

9.3.5.5 Determine the Observation Frequencies for Each Event In order to determine the number of observations for each event, an additional set of steps is required.

As shown in Figure 9.16, use the pull-down menus to choose “Analyze,” “Descriptive Statistics,” and “Frequencies. . . .”

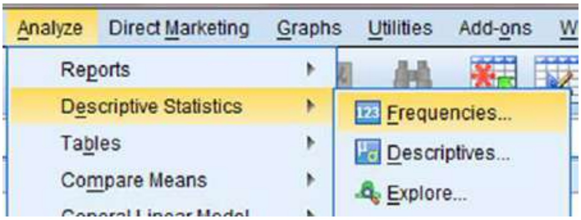


FIGURE 9.16

Next, use the arrow button to place the variable with your data values in the box labeled "Variable(s):," as shown in Figure 9.17. The variable test is called "Quiz." Finally, click "OK" to perform the analysis.

The second output table (see SPSS Output 9.4) displays the frequencies for each value. You must count the number of values above the custom value and the number values below it to determine the frequency for each event.

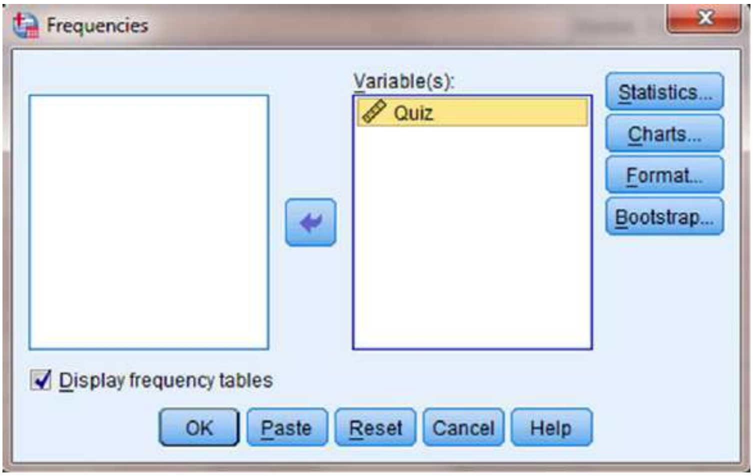


FIGURE 9.17

Quiz

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	15.00	1	8.3	8.3	8.3
	55.00	2	16.7	16.7	25.0
	60.00	1	8.3	8.3	33.3
	65.00	2	16.7	16.7	50.0
	70.00	1	8.3	8.3	58.3
	75.00	2	16.7	16.7	75.0
	80.00	2	16.7	16.7	91.7
	95.00	1	8.3	8.3	100.0
	Total	12	100.0	100.0	

SPSS OUTPUT 9.4

Based on the results from SPSS, the runs test indicated that the sequence was random ($R = 8$, $n_1 = 6$, $n_2 = 6$, $p > 0.05$). Therefore, the pattern of the student's weekly science quiz performance is random in terms of achieving a passing score of 70.

9.4 EXAMPLES FROM THE LITERATURE

Listed are varied examples of the nonparametric procedures described in this chapter. We have summarized each study's research problem and researchers' rationale(s) for choosing a nonparametric approach. We encourage you to obtain these studies if you are interested in their results.

Dorsey-Palmateer and Smith (2004) called into question a classical statistics experiment that debunked a commonly held belief that basketball players' shooting accuracy is based on the performance immediately preceding a given shot. The authors explored this notion of hot hands among professional bowlers. They examined a series of rolls and differentiate between strikes and nonstrikes. They used a runs test to analyze the sequence of bowlers' performance for randomness.

Vergin (2000) explored the presence of momentum among Major League Baseball (MLB) teams and National Basketball Association (NBA) teams. He described momentum as the tendency for a winning team to continue to win and a losing team to continue to lose. Therefore, he used a Wald–Wolfowitz runs test to examine the winning and losing streaks of the 28 MLB teams in 1996 and of the 29 NBA teams during the 1996–1997 and 1997–1998 seasons.

Pollay et al. (1992) investigated the possibility that cigarette companies segregated and segmented advertising efforts toward black consumers. They used a runs test to compare the change in annual frequency of cigarette ads that appeared in Life Magazine versus Ebony.

9.5 SUMMARY

The runs test is a statistical procedure for examining a series of events for randomness. This nonparametric test has no parametric equivalent. In this chapter, we described how to perform and interpret a runs test for both small samples and large samples. We also explained how to perform the procedure using SPSS. Finally, we offered varied examples of these nonparametric statistics from the literature.

9.6 PRACTICE QUESTIONS

1. Represented in the data is the daily performance of a popular stock. Letter A represents a gain and letter B represents a loss. Use a runs test to analyze the stock's performance for randomness. Set $\alpha = 0.05$. Report the results.
   BAABBAABBBBBAABAAAAB

2. A machine on an automated assembly line produces a unique type of bolt. If the machine fails more than three times in an hour, the total production on the line is slowed down. The machine has often exceeded the number of acceptable failures for the last week. The machine is expensive and more cost-effective to repair, but the maintenance crew cannot find the problem. The plant manager asks you to determine if the failure rates are random or if a pattern exists. Table 9.3 shows the number of failures per hour for a 24-h period.

TABLE 9.3

Hour	Number of failures
1	6
2	4
3	2
4	2
5	7
6	5
7	7
8	9
9	2
10	0
11	0
12	0
13	7
14	6
15	5
16	9
17	1
18	0
19	1
20	8
21	5
22	9
23	4
24	5

Use a runs test with a custom value of 3.1 to analyze the acceptable/unacceptable failure rate for randomness. Set $\alpha = 0.05$. Report the results.

9.7 SOLUTIONS TO PRACTICE QUESTIONS

1. The results from the analysis are displayed in SPSS Output 9.5a and SPSS Output 9.5b.

Runs Test

	Performance
Test Value[a]	1.0000
Total Cases	20
Number of Runs	9
Z	-.689
Asymp. Sig. (2-tailed)	.491

a. User-specified.

SPSS OUTPUT 9.5A

Performance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	gain	10	47.6	50.0	50.0
	loss	10	47.6	50.0	100.0
	Total	20	95.2	100.0	
Missing	System	1	4.8		
Total		21	100.0		

SPSS OUTPUT 9.5B

The sequence of the stock's gains and losses was random ($R = 9$, $n_1 = 10$, $n_2 = 10$, $p > 0.05$).

2. The results from the analysis are displayed in SPSS Output 9.6a and SPSS Output 9.6b.

Runs Test

	Failure_Rate
Test Value[a]	3.1000
Total Cases	24
Number of Runs	7
Z	-2.121
Asymp. Sig. (2-tailed)	.034

a. User-specified.

SPSS OUTPUT 9.6A

The sequence of the machine's acceptable/unacceptable failure rate was not random ($R = 7$, $n_1 = 9$, $n_2 = 15$, $p < 0.05$).

Failure_Rate

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	4	16.7	16.7	16.7
	1.00	2	8.3	8.3	25.0
	2.00	3	12.5	12.5	37.5
	4.00	2	8.3	8.3	45.8
	5.00	4	16.7	16.7	62.5
	6.00	2	8.3	8.3	70.8
	7.00	3	12.5	12.5	83.3
	8.00	1	4.2	4.2	87.5
	9.00	3	12.5	12.5	100.0
	Total	24	100.0	100.0	

SPSS OUTPUT 9.6B

APPENDIX A

SPSS AT A GLANCE

A.1 INTRODUCTION

Statistical Package for Social Sciences, or SPSS®, is a powerful tool for performing statistical analyses. Once you learn some basics, you will be able to save hours of algebraic computations while producing meaningful results. This section of the appendices includes a very basic overview of SPSS. We recommend you also run the tutorial when the program initially starts. The tutorial offers you a more detailed description of how to use the program.

A.2 OPENING SPSS

Begin by launching SPSS like any normal application. After SPSS begins, a window will appear, as seen in Figure A.1. Choose "Type in data" and click "OK." From this screen, you can also view the SPSS tutorial or open a file with existing data.

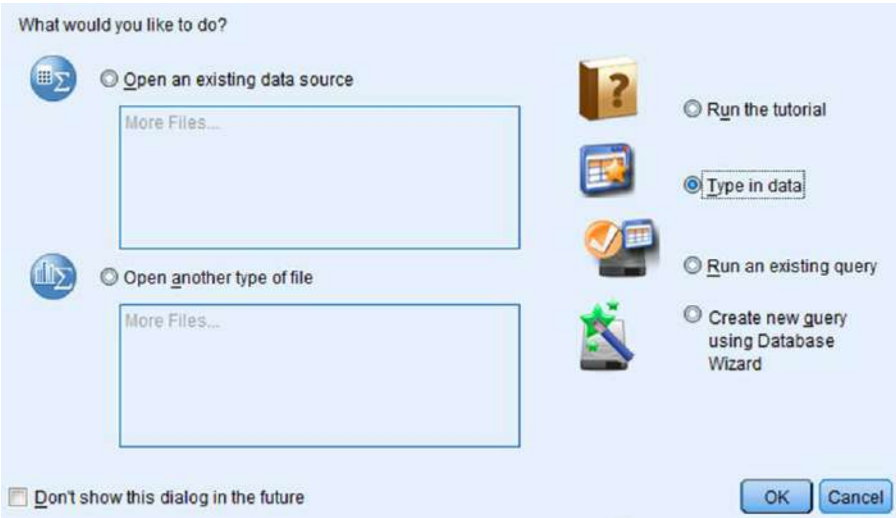


FIGURE A.1

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

A.3 INPUTTING DATA

The SPSS Data Editor window is shown in Figure A.2a. This window will allow you to type in your data. Notice the “Data View” and the “Variable View” tabs at the bottom of the window. Before inputting values, we must setup our variables.

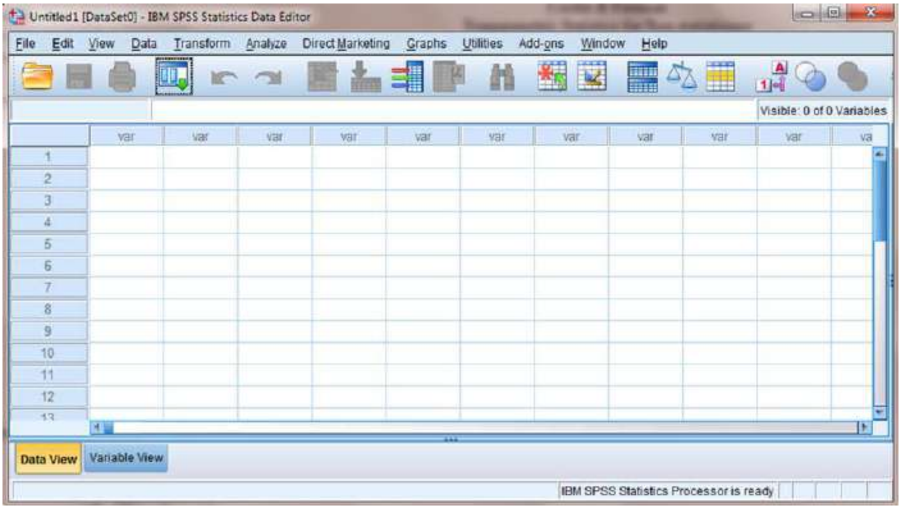


FIGURE A.2a

1. Select the “Variable View” tab located at the bottom of the SPSS Data Editor window to define the characteristics of your variables.
2. Once you change the window to Variable View, as seen in Figure A.2b, type in the names for each variable in the “Names” field. SPSS will not accept spaces at this step, so use underscores. For example, use “Test_A” instead of “Test A.”
3. In the “Width” field, choose the maximum number of characters for each value.

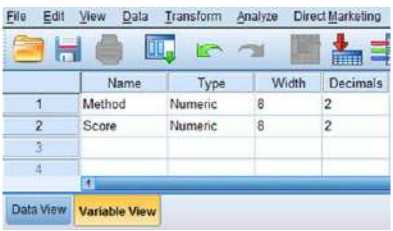


FIGURE A.2b

4. In the “Decimals” field, choose the number of decimals for each value. For the “Score” variable, we have changed the width to three characters and the decimals to 0, as seen in Figure A.2c.

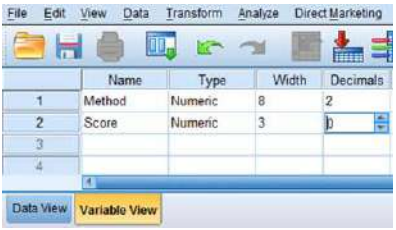


FIGURE A.2c

5. Use the “Label” field to assign names to the variables. Those names will appear in the output report that SPSS returns after an analysis. In addition, the “Label” field will allow you to use spaces, unlike the “Name” field. Figure A.2d illustrates that “Teaching Method” will identify the “Method” variable in the SPSS output report.

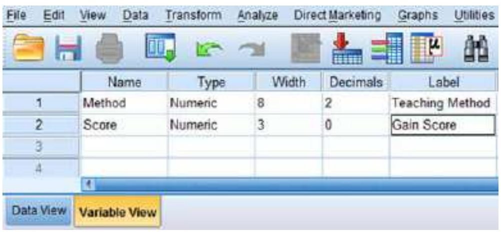


FIGURE A.2d

6. Use the “Values” field to assign a value to categorical data. As shown in Figure A.2e, clicking on the gray box in the right side of the column will cause a new window to appear, which allows you to input your settings. As seen in Figure A.2e, the “One-on-One” teaching method is assigned a value of 1 and the “Small Group” teaching method is assigned a value of 2.
7. The “Align” field allows you to change the alignment of the values as they appear in the “Data view” tab.

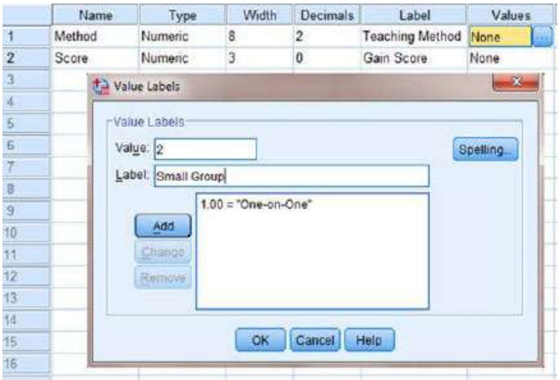


FIGURE A.2e

8. The "Measure" field allows you to select the type of scale. In Figure A.2f, we are preparing to change the "Method" variable from a scale to a nominal measure.

FIGURE A.2f

9. Finally, click on the "Data view" tab at the bottom of the screen and manually type in your data values or paste them from a spreadsheet. Figure A.2g shows the values entered into the Data Editor.

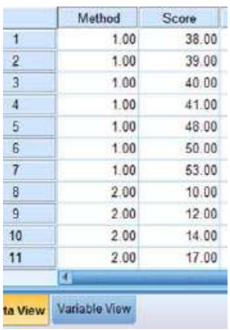


FIGURE A.2g

A.4 ANALYZING DATA

Choose “Analyze” from the pull-down menus at the top of the Data Editor window and select the appropriate test. In Figure A.3, notice that “Nonparametric Tests” has been selected. This menu presents all of the tests discussed in this book.

If your menu choices look slightly different than Figure A.3, it may be that you are using the student version of SPSS. The student version is far more powerful than most people will ever need, and even though the figures in this book are from the full version of SPSS, we doubt that you will notice any difference.

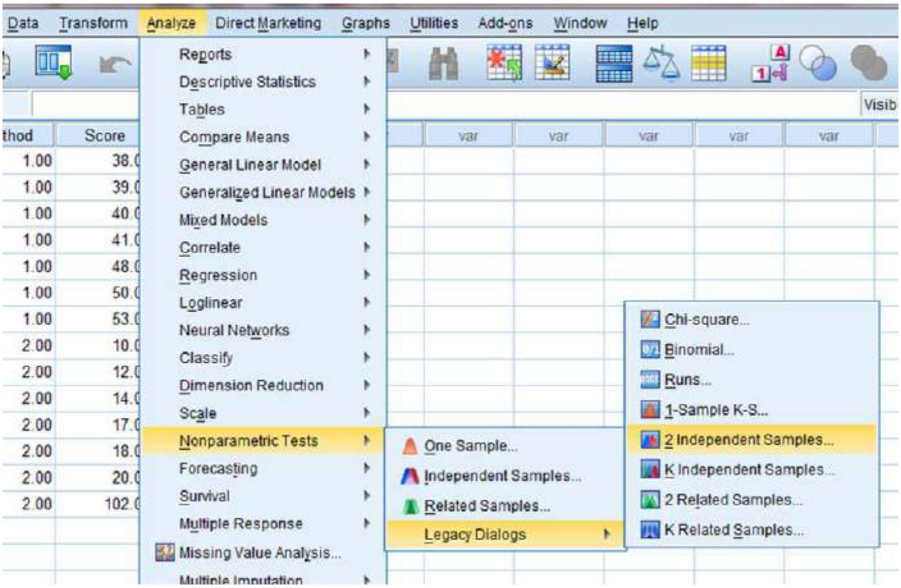


FIGURE A.3

A.5 THE SPSS OUTPUT

Once SPSS has performed an analysis, a new window will appear called the “SPSS Viewer,” as seen in Figure A.4. The window is separated into two panes. The left pane is called the outline pane and shows all of the information stored in the viewer. The pane to the right is called the contents pane and shows the actual output of the analysis.

You may wish to include any tables or graphs from the output pane in a report. You can select the objects in the contents pane and copy them onto a word processing document. The small red arrow identifies the selected object.

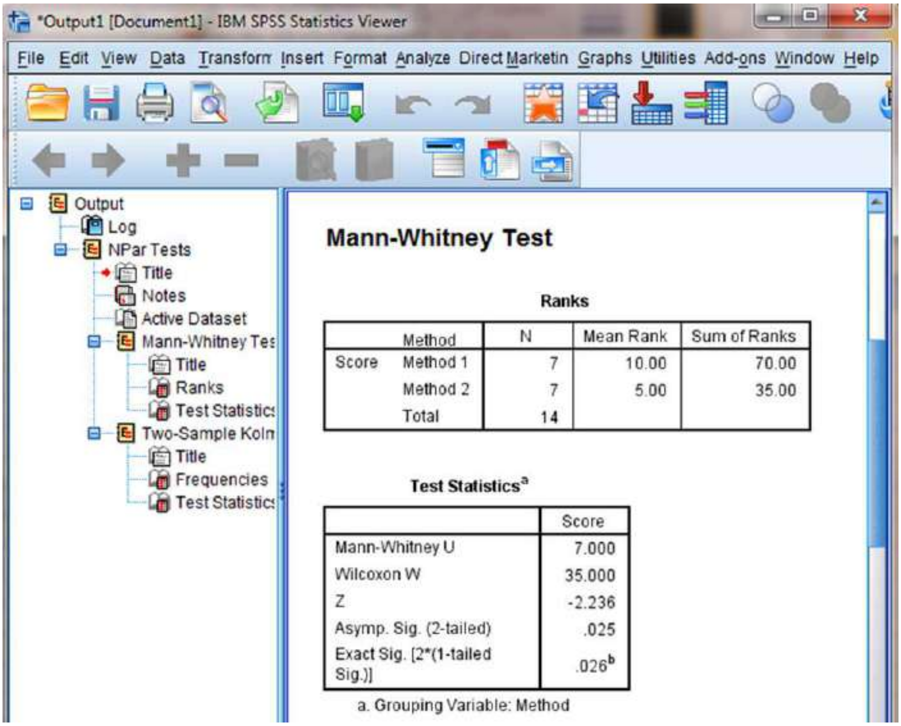

*Output1 [Document1] - IBM SPSS Statistics Viewer
File Edit View Data Transform Insert Format Analyze Direct Marketin Graphs Utilities Add-ons Window Help
Output
Log
NPar Tests
Title
Notes
Active Dataset
Mann-Whitney Tes
Title
Ranks
Test Statistics
Two-Sample Kolm
Title
Frequencies
Test Statistics
Mann-Whitney Test
Ranks
Method
N
Mean Rank
Sum of Ranks
Score
Method 1
7
10.00
70.00
Method 2
7
5.00
35.00
Total
14
Test Statistics[a]
Score
Mann-Whitney U
7.000
Wilcoxon W
35.000
Z
-2.236
Asymp. Sig. (2-tailed)
.025
Exact Sig. [2*(1-tailed Sig.)]
.026[b]
a. Grouping Variable: Method

FIGURE A.4

APPENDIX B

CRITICAL VALUE TABLES

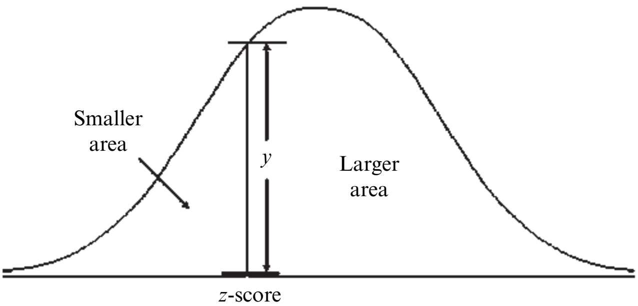


FIGURE B.1

TABLE B.1 The Normal Distribution.

z-score	Smaller area	Larger area	*y*
0.00	0.5000	0.5000	0.3989
0.01	0.4960	0.5040	0.3989
0.02	0.4920	0.5080	0.3989
0.03	0.4880	0.5120	0.3988
0.04	0.4840	0.5160	0.3986
0.05	0.4801	0.5199	0.3984
0.06	0.4761	0.5239	0.3982
0.07	0.4721	0.5279	0.3980
0.08	0.4681	0.5319	0.3977
0.09	0.4641	0.5359	0.3973
0.10	0.4602	0.5398	0.3970
0.11	0.4562	0.5438	0.3965
0.12	0.4522	0.5478	0.3961
0.13	0.4483	0.5517	0.3956
0.14	0.4443	0.5557	0.3951
0.15	0.4404	0.5596	0.3945

(*Continued*)

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	y
0.16	0.4364	0.5636	0.3939
0.17	0.4325	0.5675	0.3932
0.18	0.4286	0.5714	0.3925
0.19	0.4247	0.5753	0.3918
0.20	0.4207	0.5793	0.3910
0.21	0.4168	0.5832	0.3902
0.22	0.4129	0.5871	0.3894
0.23	0.4090	0.5910	0.3885
0.24	0.4052	0.5948	0.3876
0.25	0.4013	0.5987	0.3867
0.26	0.3974	0.6026	0.3857
0.27	0.3936	0.6064	0.3847
0.28	0.3897	0.6103	0.3836
0.29	0.3859	0.6141	0.3825
0.30	0.3821	0.6179	0.3814
0.31	0.3783	0.6217	0.3802
0.32	0.3745	0.6255	0.3790
0.33	0.3707	0.6293	0.3778
0.34	0.3669	0.6331	0.3765
0.35	0.3632	0.6368	0.3752
0.36	0.3594	0.6406	0.3739
0.37	0.3557	0.6443	0.3725
0.38	0.3520	0.6480	0.3712
0.39	0.3483	0.6517	0.3697
0.40	0.3446	0.6554	0.3683
0.41	0.3409	0.6591	0.3668
0.42	0.3372	0.6628	0.3653
0.43	0.3336	0.6664	0.3637
0.44	0.3300	0.6700	0.3621
0.45	0.3264	0.6736	0.3605
0.46	0.3228	0.6772	0.3589
0.47	0.3192	0.6808	0.3572
0.48	0.3156	0.6844	0.3555
0.49	0.3121	0.6879	0.3538
0.50	0.3085	0.6915	0.3521
0.51	0.3050	0.6950	0.3503
0.52	0.3015	0.6985	0.3485
0.53	0.2981	0.7019	0.3467
0.54	0.2946	0.7054	0.3448
0.55	0.2912	0.7088	0.3429
0.56	0.2877	0.7123	0.3410
0.57	0.2843	0.7157	0.3391
0.58	0.2810	0.7190	0.3372

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	y
0.59	0.2776	0.7224	0.3352
0.60	0.2743	0.7257	0.3332
0.61	0.2709	0.7291	0.3312
0.62	0.2676	0.7324	0.3292
0.63	0.2643	0.7357	0.3271
0.64	0.2611	0.7389	0.3251
0.65	0.2578	0.7422	0.3230
0.66	0.2546	0.7454	0.3209
0.67	0.2514	0.7486	0.3187
0.68	0.2483	0.7517	0.3166
0.69	0.2451	0.7549	0.3144
0.70	0.2420	0.7580	0.3123
0.71	0.2389	0.7611	0.3101
0.72	0.2358	0.7642	0.3079
0.73	0.2327	0.7673	0.3056
0.74	0.2296	0.7704	0.3034
0.75	0.2266	0.7734	0.3011
0.76	0.2236	0.7764	0.2989
0.77	0.2206	0.7794	0.2966
0.78	0.2177	0.7823	0.2943
0.79	0.2148	0.7852	0.2920
0.80	0.2119	0.7881	0.2897
0.81	0.2090	0.7910	0.2874
0.82	0.2061	0.7939	0.2850
0.83	0.2033	0.7967	0.2827
0.84	0.2005	0.7995	0.2803
0.85	0.1977	0.8023	0.2780
0.86	0.1949	0.8051	0.2756
0.87	0.1922	0.8078	0.2732
0.88	0.1894	0.8106	0.2709
0.89	0.1867	0.8133	0.2685
0.90	0.1841	0.8159	0.2661
0.91	0.1814	0.8186	0.2637
0.92	0.1788	0.8212	0.2613
0.93	0.1762	0.8238	0.2589
0.94	0.1736	0.8264	0.2565
0.95	0.1711	0.8289	0.2541
0.96	0.1685	0.8315	0.2516
0.97	0.1660	0.8340	0.2492
0.98	0.1635	0.8365	0.2468
0.99	0.1611	0.8389	0.2444
1.00	0.1587	0.8413	0.2420

(*Continued*)

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	y
1.01	0.1562	0.8438	0.2396
1.02	0.1539	0.8461	0.2371
1.03	0.1515	0.8485	0.2347
1.04	0.1492	0.8508	0.2323
1.05	0.1469	0.8531	0.2299
1.06	0.1446	0.8554	0.2275
1.07	0.1423	0.8577	0.2251
1.08	0.1401	0.8599	0.2227
1.09	0.1379	0.8621	0.2203
1.10	0.1357	0.8643	0.2179
1.11	0.1335	0.8665	0.2155
1.12	0.1314	0.8686	0.2131
1.13	0.1292	0.8708	0.2107
1.14	0.1271	0.8729	0.2083
1.15	0.1251	0.8749	0.2059
1.16	0.1230	0.8770	0.2036
1.17	0.1210	0.8790	0.2012
1.18	0.1190	0.8810	0.1989
1.19	0.1170	0.8830	0.1965
1.20	0.1151	0.8849	0.1942
1.21	0.1131	0.8869	0.1919
1.22	0.1112	0.8888	0.1895
1.23	0.1093	0.8907	0.1872
1.24	0.1075	0.8925	0.1849
1.25	0.1056	0.8944	0.1826
1.26	0.1038	0.8962	0.1804
1.27	0.1020	0.8980	0.1781
1.28	0.1003	0.8997	0.1758
1.29	0.0985	0.9015	0.1736
1.30	0.0968	0.9032	0.1714
1.31	0.0951	0.9049	0.1691
1.32	0.0934	0.9066	0.1669
1.33	0.0918	0.9082	0.1647
1.34	0.0901	0.9099	0.1626
1.35	0.0885	0.9115	0.1604
1.36	0.0869	0.9131	0.1582
1.37	0.0853	0.9147	0.1561
1.38	0.0838	0.9162	0.1539
1.39	0.0823	0.9177	0.1518
1.40	0.0808	0.9192	0.1497
1.41	0.0793	0.9207	0.1476
1.42	0.0778	0.9222	0.1456
1.43	0.0764	0.9236	0.1435

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	*y*
1.44	0.0749	0.9251	0.1415
1.45	0.0735	0.9265	0.1394
1.46	0.0721	0.9279	0.1374
1.47	0.0708	0.9292	0.1354
1.48	0.0694	0.9306	0.1334
1.49	0.0681	0.9319	0.1315
1.50	0.0668	0.9332	0.1295
1.51	0.0655	0.9345	0.1276
1.52	0.0643	0.9357	0.1257
1.53	0.0630	0.9370	0.1238
1.54	0.0618	0.9382	0.1219
1.55	0.0606	0.9394	0.1200
1.56	0.0594	0.9406	0.1182
1.57	0.0582	0.9418	0.1163
1.58	0.0571	0.9429	0.1145
1.59	0.0559	0.9441	0.1127
1.60	0.0548	0.9452	0.1109
1.61	0.0537	0.9463	0.1092
1.62	0.0526	0.9474	0.1074
1.63	0.0516	0.9484	0.1057
1.64	0.0505	0.9495	0.1040
1.65	0.0495	0.9505	0.1023
1.66	0.0485	0.9515	0.1006
1.67	0.0475	0.9525	0.0989
1.68	0.0465	0.9535	0.0973
1.69	0.0455	0.9545	0.0957
1.70	0.0446	0.9554	0.0940
1.71	0.0436	0.9564	0.0925
1.72	0.0427	0.9573	0.0909
1.73	0.0418	0.9582	0.0893
1.74	0.0409	0.9591	0.0878
1.75	0.0401	0.9599	0.0863
1.76	0.0392	0.9608	0.0848
1.77	0.0384	0.9616	0.0833
1.78	0.0375	0.9625	0.0818
1.79	0.0367	0.9633	0.0804
1.80	0.0359	0.9641	0.0790
1.81	0.0351	0.9649	0.0775
1.82	0.0344	0.9656	0.0761
1.83	0.0336	0.9664	0.0748
1.84	0.0329	0.9671	0.0734
1.85	0.0322	0.9678	0.0721

(*Continued*)

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	y
1.86	0.0314	0.9686	0.0707
1.87	0.0307	0.9693	0.0694
1.88	0.0301	0.9699	0.0681
1.89	0.0294	0.9706	0.0669
1.90	0.0287	0.9713	0.0656
1.91	0.0281	0.9719	0.0644
1.92	0.0274	0.9726	0.0632
1.93	0.0268	0.9732	0.0620
1.94	0.0262	0.9738	0.0608
1.95	0.0256	0.9744	0.0596
1.96	0.0250	0.9750	0.0584
1.97	0.0244	0.9756	0.0573
1.98	0.0239	0.9761	0.0562
1.99	0.0233	0.9767	0.0551
2.00	0.0228	0.9772	0.0540
2.01	0.0222	0.9778	0.0529
2.02	0.0217	0.9783	0.0519
2.03	0.0212	0.9788	0.0508
2.04	0.0207	0.9793	0.0498
2.05	0.0202	0.9798	0.0488
2.06	0.0197	0.9803	0.0478
2.07	0.0192	0.9808	0.0468
2.08	0.0188	0.9812	0.0459
2.09	0.0183	0.9817	0.0449
2.10	0.0179	0.9821	0.0440
2.11	0.0174	0.9826	0.0431
2.12	0.0170	0.9830	0.0422
2.13	0.0166	0.9834	0.0413
2.14	0.0162	0.9838	0.0404
2.15	0.0158	0.9842	0.0396
2.16	0.0154	0.9846	0.0387
2.17	0.0150	0.9850	0.0379
2.18	0.0146	0.9854	0.0371
2.19	0.0143	0.9857	0.0363
2.20	0.0139	0.9861	0.0355
2.21	0.0136	0.9864	0.0347
2.22	0.0132	0.9868	0.0339
2.23	0.0129	0.9871	0.0332
2.24	0.0125	0.9875	0.0325
2.25	0.0122	0.9878	0.0317
2.26	0.0119	0.9881	0.0310
2.27	0.0116	0.9884	0.0303
2.28	0.0113	0.9887	0.0297

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	y
2.29	0.0110	0.9890	0.0290
2.30	0.0107	0.9893	0.0283
2.31	0.0104	0.9896	0.0277
2.32	0.0102	0.9898	0.0270
2.33	0.0099	0.9901	0.0264
2.34	0.0096	0.9904	0.0258
2.35	0.0094	0.9906	0.0252
2.36	0.0091	0.9909	0.0246
2.37	0.0089	0.9911	0.0241
2.38	0.0087	0.9913	0.0235
2.39	0.0084	0.9916	0.0229
2.40	0.0082	0.9918	0.0224
2.41	0.0080	0.9920	0.0219
2.42	0.0078	0.9922	0.0213
2.43	0.0075	0.9925	0.0208
2.44	0.0073	0.9927	0.0203
2.45	0.0071	0.9929	0.0198
2.46	0.0069	0.9931	0.0194
2.47	0.0068	0.9932	0.0189
2.48	0.0066	0.9934	0.0184
2.49	0.0064	0.9936	0.0180
2.50	0.0062	0.9938	0.0175
2.51	0.0060	0.9940	0.0171
2.52	0.0059	0.9941	0.0167
2.53	0.0057	0.9943	0.0163
2.54	0.0055	0.9945	0.0158
2.55	0.0054	0.9946	0.0154
2.56	0.0052	0.9948	0.0151
2.57	0.0051	0.9949	0.0147
2.58	0.0049	0.9951	0.0143
2.59	0.0048	0.9952	0.0139
2.60	0.0047	0.9953	0.0136
2.61	0.0045	0.9955	0.0132
2.62	0.0044	0.9956	0.0129
2.63	0.0043	0.9957	0.0126
2.64	0.0041	0.9959	0.0122
2.65	0.0040	0.9960	0.0119
2.66	0.0039	0.9961	0.0116
2.67	0.0038	0.9962	0.0113
2.68	0.0037	0.9963	0.0110
2.69	0.0036	0.9964	0.0107
2.70	0.0035	0.9965	0.0104

(*Continued*)

TABLE B.1 (*Continued*)

z-score	Smaller area	Larger area	*y*
2.71	0.0034	0.9966	0.0101
2.72	0.0033	0.9967	0.0099
2.73	0.0032	0.9968	0.0096
2.74	0.0031	0.9969	0.0093
2.75	0.0030	0.9970	0.0091
2.76	0.0029	0.9971	0.0088
2.77	0.0028	0.9972	0.0086
2.78	0.0027	0.9973	0.0084
2.79	0.0026	0.9974	0.0081
2.80	0.0026	0.9974	0.0079
2.81	0.0025	0.9975	0.0077
2.82	0.0024	0.9976	0.0075
2.83	0.0023	0.9977	0.0073
2.84	0.0023	0.9977	0.0071
2.85	0.0022	0.9978	0.0069
2.86	0.0021	0.9979	0.0067
2.87	0.0021	0.9979	0.0065
2.88	0.0020	0.9980	0.0063
2.89	0.0019	0.9981	0.0061
2.90	0.0019	0.9981	0.0060
2.91	0.0018	0.9982	0.0058
2.92	0.0018	0.9982	0.0056
2.93	0.0017	0.9983	0.0055
2.94	0.0016	0.9984	0.0053
2.95	0.0016	0.9984	0.0051
2.96	0.0015	0.9985	0.0050
2.97	0.0015	0.9985	0.0048
2.98	0.0014	0.9986	0.0047
2.99	0.0014	0.9986	0.0046
3.00	0.0013	0.9987	0.0044
3.10	0.0010	0.9990	0.0033
3.20	0.0007	0.9993	0.0024
3.30	0.0005	0.9995	0.0017
3.50	0.0002	0.9998	0.0009
3.75	0.0001	0.9999	0.0004
4.00	0.0000	1.0000	0.0001

Source: Adapted by the authors from Hastings (1955).

TABLE B.2 The χ^2 Distribution.

df	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01
1	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63
2	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21
3	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34
4	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28
5	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09
6	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81
7	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48
8	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09
9	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67
10	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21
11	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73
12	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22
13	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69
14	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14
15	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58
16	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00
17	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41
18	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81
19	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19
20	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57
21	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93
22	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
26	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89

TABLE B.3 Critical Values for the Wilcoxon Signed Rank Test Statistics *T*.

n	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.02$ $\alpha_{\text{one-tailed}} \leq 0.01$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
5	0			
6	2	0		
7	3	2	0	
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15
16	35	29	23	19
17	41	34	27	23
18	47	40	32	27
19	53	46	37	32
20	60	52	43	37
21	67	58	49	42
22	75	65	55	48
23	83	73	62	54
24	91	81	69	61
25	100	89	76	68
26	110	98	84	75
27	119	107	92	83
28	130	116	101	91
29	140	126	110	100
30	151	137	120	109

Source: Adapted from McCornack, R. L. (1965). Extended tables of the Wilcoxon matched pair signed rank statistic. *Journal of the American Statistical Association*, *60*, 864–871. Reprinted with permission from *The Journal of the American Statistical Association*.

TABLE B.4 Critical Values for the Mann–Whitney *U*-Test Statistic.

		n																			
α	*m*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.10	1																				
	2																				
	3		0	1																	
	4		0	1	3																
	5		1	2	4	5															
	6		1	3	5	7	9														
	7		1	4	6	8	11	13													
	8		2	5	7	10	13	16	19												
	9	0	2	5	9	12	15	18	22	25											
	10	0	3	6	10	13	17	21	24	28	32										
	11	0	3	7	11	15	19	23	27	31	36	40									
	12	0	4	8	12	17	21	26	30	35	39	44	49								
	13	0	4	9	13	18	23	28	33	38	43	48	53	58							
	14	0	5	10	15	20	25	31	36	41	47	52	58	63	69						
	15	0	5	10	16	22	27	33	39	45	51	57	63	68	74	80					
	16	0	5	11	17	23	29	36	42	48	54	61	67	74	80	86	93				
	17	0	6	12	18	25	31	38	45	52	58	65	72	79	85	92	99	106			
	18	0	6	13	20	27	34	41	48	55	62	69	77	84	91	98	106	113	120		
	19	1	7	14	21	28	36	43	51	58	66	73	81	89	97	104	112	120	128	135	
	20	1	7	15	22	30	38	46	54	62	70	78	86	94	102	110	119	127	135	143	151
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.05	1																				
	2																				
	3			0																	
	4			0	1																
	5		0	1	2	4															
	6		0	2	3	5	7														
	7		0	2	4	6	8	11													
	8		1	3	5	8	10	13	15												
	9		1	4	6	9	12	15	18	21											
	10		1	4	7	11	14	17	20	24	27										
	11		1	5	8	12	16	19	23	27	31	34									
	12		2	5	9	13	17	21	26	30	34	38	42								
	13		2	6	10	15	19	24	28	33	37	42	47	51							
	14		3	7	11	16	21	26	31	36	41	46	51	56	61						
	15		3	7	12	18	23	28	33	39	44	50	55	61	66	72					
	16		3	8	14	19	25	30	36	42	48	54	60	65	71	77	83				
	17		3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96			
	18		4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109		
	19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	
	20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

(*Continued*)

TABLE B.4 (*Continued*)

		n																			
α	*m*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.025	1																				
	2																				
	3																				
	4				0																
	5			0	1	2															
	6			1	2	3	5														
	7			1	3	5	6	8													
	8		0	2	4	6	8	10	13												
	9		0	2	4	7	10	12	15	17											
	10		0	3	5	8	11	14	17	20	23										
	11		0	3	6	9	13	16	19	23	26	30									
	12		1	4	7	11	14	18	22	26	29	33	37								
	13		1	4	8	12	16	20	24	28	33	37	41	45							
	14		1	5	9	13	17	22	26	31	36	40	45	50	55						
	15		1	5	10	14	19	24	29	34	39	44	49	54	59	64					
	16		1	6	11	15	21	26	31	37	42	47	53	59	64	70	75				
	17		2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87			
	18		2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99		
	19		2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	
	20		2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0.01	1																				
	2																				
	3																				
	4																				
	5				0	1															
	6				1	2	3														
	7			0	1	3	4	6													
	8			0	2	4	6	7	9												
	9			1	3	5	7	9	11	14											
	10			1	3	6	8	11	13	16	19										
	11			1	4	7	9	12	15	18	22	25									
	12			2	5	8	11	14	17	21	24	28	31								
	13		0	2	5	9	12	16	20	23	27	31	35	39							
	14		0	2	6	10	13	17	22	26	30	34	38	43	47						
	15		0	3	7	11	15	19	24	28	33	37	42	47	51	56					
	16		0	3	7	12	16	21	26	31	36	41	46	51	56	61	66				
	17		0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77			
	18		0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88		
	19		1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	
	20		1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Source: Adapted from Milton, R. C. (1964). An extended table of critical values for the Mann-Whitney (Wilcoxon) two-sample statistic. *Journal of the American Statistical Association*, *59*, 925–934. Reprinted with permission from *The Journal of the American Statistical Association*.

TABLE B.5 Critical Values for the Friedman Test Statistic F_r.

k	N	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.025$	$\alpha \leq 0.01$
3	3	6.000	6.000		
	4	6.000	6.500	8.000	8.000
	5	5.200	6.400	7.600	8.400
	6	5.333	7.000	8.333	9.000
	7	5.429	7.143	7.714	8.857
	8	5.250	6.250	7.750	9.000
	9	5.556	6.222	8.000	8.667
	10	5.000	6.200	7.800	9.600
	11	4.909	6.545	7.818	9.455
	12	5.167	6.500	8.000	9.500
	13	4.769	6.000	7.538	9.385
	14	5.143	6.143	7.429	9.000
	15	4.933	6.400	7.600	8.933
4	2	6.000	6.000		
	3	6.600	7.400	8.200	9.000
	4	6.300	7.800	8.400	9.600
	5	6.360	7.800	8.760	9.960
	6	6.400	7.600	8.800	10.200
	7	6.429	7.800	9.000	10.371
	8	6.300	7.650	9.000	10.500
	9	6.467	7.800	9.133	10.867
	10	6.360	7.800	9.120	10.800
	11	6.382	7.909	9.327	11.073
	12	6.400	7.900	9.200	11.100
	13	6.415	7.985	7.369	11.123
	14	6.343	7.886	9.343	11.143
	15	6.440	8.040	9.400	11.240
5	2	7.200	7.600	8.000	8.000
	3	7.467	8.533	9.600	10.133
	4	7.600	8.800	9.800	11.200
	5	7.680	8.960	10.240	11.680
	6	7.733	9.067	10.400	11.867
	7	7.771	9.143	10.514	12.114
	8	7.800	9.300	10.600	12.300
	9	7.733	9.244	10.667	12.444
	10	7.760	9.280	10.720	12.480
6	2	8.286	9.143	9.429	9.714
	3	8.714	9.857	10.810	11.762
	4	9.000	10.286	11.429	12.714
	5	9.000	10.486	11.743	13.229
	6	9.048	10.571	12.000	13.619
	7	9.122	10.674	12.061	13.857

(*Continued*)

TABLE B.5 (*Continued*)

k	N	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.025$	$\alpha \leq 0.01$
	8	9.143	10.714	12.214	14.000
	9	9.127	10.778	12.302	14.143
	10	9.143	10.800	12.343	14.229

Source: Adapted from Martin, L., Leblanc, R., & Toan, N. K. (1993). Tables for the Friedman rank test. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, *21*(1), 39–43. Reprinted with permission from *The Canadian Journal of Statistics*.

TABLE B.6 The Critical Values for the Kruskal–Wallis *H*-Test Statistic.

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
2	2	2	4.571429	–	–
3	1	1	–	–	–
3	2	1	4.285714	–	–
3	2	2	4.464286	4.714286	–
3	3	1	4.571429	5.142857	–
3	3	2	4.555556	5.361111	–
3	3	3	4.622222	5.600000	6.488889
4	2	1	4.500000	–	–
4	2	2	4.458333	5.333333	–
4	3	1	4.055556	5.208333	–
4	3	2	4.511111	5.444444	6.444444
4	3	3	4.700000	5.790909	6.745455
4	4	1	4.166667	4.966667	6.666667
4	4	2	4.554545	5.454545	7.036364
4	4	3	4.545455	5.598485	7.143939
4	4	4	4.653846	5.692308	7.653846
5	2	1	4.200000	5.000000	–
5	2	2	4.373333	5.160000	6.533333
5	3	1	4.017778	4.871111	–
5	3	2	4.650909	5.250909	6.821818
5	3	3	4.533333	5.648485	7.078788
5	4	1	3.987273	4.985455	6.954545
5	4	2	4.540909	5.272727	7.204545
5	4	3	4.548718	5.656410	7.444872
5	4	4	4.668132	5.657143	7.760440
5	5	1	4.109091	5.127273	7.309091
5	5	2	4.623077	5.338462	7.338462
5	5	3	4.545055	5.626374	7.578022
5	5	4	4.522857	5.665714	7.791429
5	5	5	4.560000	5.780000	8.000000

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
6	2	1	4.200000	4.822222	–
6	2	2	4.436364	5.345455	6.654545
6	3	1	3.909091	4.854545	6.581818
6	3	2	4.681818	5.348485	6.969697
6	3	3	4.538462	5.615385	7.192308
6	4	1	4.037879	4.946970	7.083333
6	4	2	4.493590	5.262821	7.339744
6	4	3	4.604396	5.604396	7.467033
6	4	4	4.523810	5.666667	7.795238
6	5	1	4.128205	4.989744	7.182051
6	5	2	4.595604	5.318681	7.375824
6	5	3	4.535238	5.601905	7.590476
6	5	4	4.522500	5.660833	7.935833
6	5	5	4.547059	5.698529	8.027941
6	6	1	4.000000	4.857143	7.065934
6	6	2	4.438095	5.409524	7.466667
6	6	3	4.558333	5.625000	7.725000
6	6	4	4.547794	5.724265	8.000000
6	6	5	4.542484	5.764706	8.118954
6	6	6	4.538012	5.719298	8.222222
7	1	1	4.266667	–	–
7	2	1	4.200000	4.706494	–
7	2	2	4.525974	5.142857	7.000000
7	3	1	4.173160	4.952381	6.649351
7	3	2	4.582418	5.357143	6.838828
7	3	3	4.602826	5.620094	7.227630
7	4	1	4.120879	4.986264	6.986264
7	4	2	4.549451	5.375981	7.304553
7	4	3	4.527211	5.623129	7.498639
7	4	4	4.562500	5.650000	7.814286
7	5	1	4.035165	5.063736	7.060597
7	5	2	4.484898	5.392653	7.449796
7	5	3	4.535238	5.588571	7.697143
7	5	4	4.541597	5.732773	7.931092
7	5	5	4.540056	5.707563	8.100840
7	6	1	4.032653	5.066667	7.254422
7	6	2	4.500000	5.357143	7.490476
7	6	3	4.550420	5.672269	7.756303
7	6	4	4.561625	5.705882	8.016340
7	6	5	4.559733	5.769925	8.156725

(*Continued*)

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
7	6	6	4.530075	5.730075	8.257143
7	7	1	3.985714	4.985714	7.157143
7	7	2	4.490546	5.398109	7.490546
7	7	3	4.590103	5.676937	7.809524
7	7	4	4.558897	5.765664	8.141604
7	7	5	4.545564	5.745564	8.244812
7	7	6	4.568027	5.792517	8.341497
7	7	7	4.593692	5.818182	8.378479
8	1	1	4.418182	–	–
8	2	1	4.011364	4.909091	–
8	2	2	4.586538	5.355769	6.663462
8	3	1	4.009615	4.881410	6.804487
8	3	2	4.450549	5.315934	6.986264
8	3	3	4.504762	5.616667	7.254762
8	4	1	4.038462	5.043956	6.972527
8	4	2	4.500000	5.392857	7.350000
8	4	3	4.529167	5.622917	7.585417
8	4	4	4.560662	5.779412	7.852941
8	5	1	3.967143	4.868571	7.110000
8	5	2	4.466250	5.415000	7.440000
8	5	3	4.514338	5.614338	7.705515
8	5	4	4.549020	5.717647	7.992157
8	5	5	4.555263	5.769298	8.115789
8	6	1	4.014583	5.014583	7.256250
8	6	2	4.441176	5.404412	7.522059
8	6	3	4.573529	5.678105	7.795752
8	6	4	4.562865	5.742690	8.045322
8	6	5	4.550263	5.750263	8.210263
8	6	6	4.598810	5.770238	8.294048
8	7	1	4.045431	5.041229	7.307773
8	7	2	4.450980	5.403361	7.571429
8	7	3	4.555556	5.698413	7.827068
8	7	4	4.548496	5.759211	8.118045
8	7	5	4.550612	5.777449	8.241939
8	7	6	4.552876	5.781231	8.332715
8	7	7	4.573687	5.795031	8.356296
8	8	1	4.044118	5.039216	7.313725
8	8	2	4.508772	5.407895	7.653509
8	8	3	4.555263	5.734211	7.889474
8	8	4	4.578571	5.742857	8.167857
8	8	5	4.572727	5.761039	8.297403

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
8	8	6	4.572134	5.778656	8.366601
8	8	7	4.570652	5.791149	8.418866
8	8	8	4.595000	5.805000	8.465000
9	1	1	4.545455	–	–
9	2	1	3.905983	4.841880	6.346154
9	2	2	4.483516	5.260073	6.897436
9	3	1	4.073260	4.952381	6.886447
9	3	2	4.492063	5.339683	6.990476
9	3	3	4.633333	5.588889	7.355556
9	4	1	3.971429	5.071429	7.171429
9	4	2	4.488889	5.400000	7.363889
9	4	3	4.514706	5.651961	7.613971
9	4	4	4.576253	5.703704	7.909586
9	5	1	4.055556	5.040000	7.148889
9	5	2	4.464706	5.395588	7.447059
9	5	3	4.587364	5.669717	7.733333
9	5	4	4.531384	5.712671	8.024561
9	5	5	4.557193	5.769825	8.169825
9	6	1	3.933824	5.049020	7.247549
9	6	2	4.481481	5.392157	7.494553
9	6	3	4.541910	5.664717	7.822612
9	6	4	4.545614	5.744737	8.108772
9	6	5	4.573651	5.761905	8.230794
9	6	6	4.554113	5.808081	8.307359
9	7	1	4.011204	5.042017	7.270464
9	7	2	4.480089	5.429128	7.636591
9	7	3	4.535338	5.656140	7.860652
9	7	4	4.547732	5.731406	8.131406
9	7	5	4.565492	5.757988	8.287941
9	7	6	4.570864	5.782985	8.353284
9	7	7	4.583851	5.802622	8.403037
9	8	1	3.986355	4.984893	7.394250
9	8	2	4.491667	5.419737	7.642105
9	8	3	4.568651	5.717460	7.927381
9	8	4	4.559163	5.744228	8.203102
9	8	5	4.551252	5.783465	8.318050
9	8	6	4.560688	5.775362	8.408514
9	8	7	4.563770	5.807579	8.450000
9	8	8	4.582821	5.809744	8.494359
9	9	1	4.007018	4.961404	7.333333

(*Continued*)

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
9	9	2	4.460317	5.411111	7.692063
9	9	3	4.565657	5.708514	7.959596
9	9	4	4.550066	5.751647	8.202240
9	9	5	4.587440	5.770048	8.370048
9	9	6	4.555556	5.814444	8.427778
9	9	7	4.567326	5.802198	8.468864
9	9	8	4.570750	5.815052	8.514720
9	9	9	4.582011	5.844797	8.564374
10	1	1	4.653846	4.653846	–
10	2	1	4.114286	4.839560	6.428571
10	2	2	4.434286	5.120000	6.537143
10	3	1	3.996190	5.076190	6.851429
10	3	2	4.470000	5.361667	7.041667
10	3	3	4.529412	5.588235	7.360294
10	4	1	4.042500	5.017500	7.105000
10	4	2	4.462500	5.344853	7.356618
10	4	3	4.587582	5.654248	7.616993
10	4	4	4.564912	5.715789	7.907018
10	5	1	3.988235	4.905882	7.107353
10	5	2	4.454902	5.388235	7.513725
10	5	3	4.552047	5.618713	7.752047
10	5	4	4.556842	5.744211	8.047895
10	5	5	4.574286	5.777143	8.162857
10	6	1	3.967320	5.041830	7.316340
10	6	2	4.479532	5.405848	7.588304
10	6	3	4.542105	5.655789	7.882105
10	6	4	4.550476	5.726190	8.142857
10	6	5	4.554978	5.754978	8.267532
10	6	6	4.575494	5.780237	8.338340
10	7	1	3.981454	4.985965	7.252130
10	7	2	4.491880	5.377444	7.641203
10	7	3	4.545034	5.698095	7.901224
10	7	4	4.550278	5.751206	8.172356
10	7	5	4.567250	5.763862	8.295652
10	7	6	4.563043	5.798758	8.376915
10	7	7	4.562286	5.796571	8.419429
10	8	1	3.963947	5.037632	7.358684
10	8	2	4.482857	5.429286	7.720714
10	8	3	4.533983	5.711688	7.977273
10	8	4	4.550988	5.744466	8.206126
10	9	5	4.556522	5.789130	8.344022

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 3$).

n_1	n_2	n_3	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
10	9	6	4.573333	5.793833	8.397833
10	9	7	4.564484	5.810637	8.480967
10	9	8	4.561538	5.829060	8.494017
10	9	1	4.025714	4.988571	7.436508
10	9	2	4.476479	5.446176	7.693795
10	9	3	4.570751	5.700659	7.997628
10	9	4	4.556401	5.757609	8.223430
10	9	5	4.547556	5.792000	8.380222
10	9	6	4.561231	5.813128	8.449436
10	9	7	4.559707	5.817610	8.507475
10	9	8	4.567063	5.833730	8.544489
10	9	9	4.578982	5.830706	8.575698
10	10	1	3.987013	5.054545	7.501299
10	10	2	4.477470	5.449802	7.726482
10	10	3	4.559420	5.687681	8.026087
10	10	4	4.567000	5.776000	8.263000
10	10	5	4.554462	5.793231	8.403692
10	10	6	4.561823	5.796011	8.472934
10	10	7	4.558277	5.820408	8.536508
10	10	8	4.565025	5.837438	8.565887
10	10	9	4.567050	5.837241	8.606130
10	10	10	4.583226	5.855484	8.640000

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 4$.)

n_1	n_2	n_3	n_4	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
2	2	2	1	5.357143	5.678571	–
2	2	2	2	5.666667	6.166667	6.666667
3	2	1	1	4.892857	–	–
3	2	2	1	5.555556	5.833333	–
3	2	2	2	5.644444	6.333333	7.133333
3	3	1	1	5.333333	6.333333	–
3	3	2	1	5.622222	6.244444	7.044444
3	3	2	2	5.745455	6.527273	7.636364
3	3	3	1	5.654545	6.600000	7.400000
3	3	3	2	5.878788	6.727273	8.015152
3	3	3	3	5.974359	6.897436	8.435897
4	2	1	1	5.250000	5.833333	–
4	2	2	1	5.533333	6.133333	7.000000
4	2	2	2	5.754545	6.545455	7.390909
4	3	1	1	5.066667	6.177778	7.066667

(*Continued*)

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis H-Test Statistic, $k = 4$.)

n_1	n_2	n_3	n_4	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
4	3	2	1	5.572727	6.309091	7.454545
4	3	2	2	5.750000	6.621212	7.871212
4	3	3	1	5.666667	6.545455	7.757576
4	3	3	2	5.858974	6.782051	8.320513
4	3	3	3	6.000000	6.967033	8.653846
4	4	1	1	5.181818	5.945455	7.909091
4	4	2	1	5.568182	6.386364	7.909091
4	4	2	2	5.807692	6.730769	8.346154
4	4	3	1	5.660256	6.634615	8.217949
4	4	3	2	5.901099	6.873626	8.620879
4	4	3	3	6.004762	7.038095	8.866667
4	4	4	1	5.653846	6.725275	8.587912
4	4	4	2	5.914286	6.957143	8.871429
4	4	4	3	6.029167	7.129167	9.075000
4	4	4	4	6.088235	7.235294	9.286765
5	1	1	1	5.333333	–	–
5	2	1	1	5.266667	5.960000	–
5	2	2	1	5.541818	6.109091	7.276364
5	2	2	2	5.636364	6.563636	7.772727
5	3	1	1	5.130909	6.003636	7.400000
5	3	2	1	5.518182	6.363636	7.757576
5	3	2	2	5.771795	6.664103	8.202564
5	3	3	1	5.656410	6.641026	8.117949
5	3	3	2	5.865934	6.821978	8.606593
5	3	3	3	6.020952	7.011429	8.840000
5	4	1	1	5.254545	6.040909	7.909091
5	4	2	1	5.580769	6.419231	8.173077
5	4	2	2	5.782418	6.725275	8.472527
5	4	3	1	5.639560	6.681319	8.408791
5	4	3	2	5.901905	6.925714	8.801905
5	4	3	3	6.029167	7.093333	9.029167
5	4	4	1	5.674286	6.760000	8.725714
5	4	4	2	5.947500	6.990000	9.002500
5	4	4	3	6.035294	7.172794	9.220588
5	4	4	4	6.066667	7.262745	9.392157
5	5	1	1	5.153846	6.076923	8.107692
5	5	2	1	5.564835	6.540659	8.327473
5	5	2	2	5.794286	6.777143	8.634286
5	5	3	1	5.662857	6.737143	8.611429
5	5	3	2	5.921667	6.946667	8.946667
5	5	3	3	6.023529	7.117647	9.188235

TABLE B.6 (*Continued*)

(The Critical Values for the Kruskal–Wallis *H*-Test Statistic, $k = 4$.)

n_1	n_2	n_3	n_4	$\alpha \leq 0.10$	$\alpha \leq 0.05$	$\alpha \leq 0.01$
5	5	4	1	5.670000	6.782500	8.870000
5	5	4	2	5.944853	7.032353	9.156618
5	5	4	3	6.052288	7.217647	9.356863
5	5	4	4	6.070175	7.291228	9.536842
5	5	5	1	5.682353	6.829412	9.052941
5	5	5	2	5.945098	7.074510	9.286275
5	5	5	3	6.043275	7.250292	9.495906
5	5	5	4	6.082105	7.327895	9.669474
5	5	5	5	6.097143	7.377143	9.800000

Source: Adapted from Meyer, J. P., & Seaman, M. A. (2008, March). A comparison of the exact Kruskal-Wallis distribution to asymptotic approximations for N $\leq$ 105. Paper presented at the annual meeting of the American Educational Research Association, New York. Reprinted with permission of the authors.

TABLE B.7 Critical Values for the Spearman Rank-Order Correlation Coefficient r_s.

n	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.02$ $\alpha_{\text{one-tailed}} \leq 0.01$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
4	1.000			
5	0.900	1.000	1.000	
6	0.829	0.886	0.943	1.000
7	0.714	0.786	0.893	0.929
8	0.643	0.738	0.833	0.881
9	0.600	0.700	0.783	0.833
10	0.564	0.648	0.745	0.794
11	0.536	0.618	0.709	0.755
12	0.503	0.587	0.671	0.727
13	0.484	0.560	0.648	0.703
14	0.464	0.538	0.622	0.675
15	0.443	0.521	0.604	0.654
16	0.429	0.503	0.582	0.635
17	0.414	0.485	0.566	0.615
18	0.401	0.472	0.550	0.600
19	0.391	0.460	0.535	0.584
20	0.380	0.447	0.520	0.570
21	0.370	0.435	0.508	0.556
22	0.361	0.425	0.496	0.544
23	0.353	0.415	0.486	0.532

(*Continued*)

TABLE B.7 (*Continued*)

n	$\alpha_{\text{two-tailed}} \le 0.10$ $\alpha_{\text{one-tailed}} \le 0.05$	$\alpha_{\text{two-tailed}} \le 0.05$ $\alpha_{\text{one-tailed}} \le 0.025$	$\alpha_{\text{two-tailed}} \le 0.02$ $\alpha_{\text{one-tailed}} \le 0.01$	$\alpha_{\text{two-tailed}} \le 0.01$ $\alpha_{\text{one-tailed}} \le 0.005$
24	0.344	0.406	0.476	0.321
25	0.337	0.398	0.466	0.511
26	0.331	0.390	0.457	0.501
27	0.324	0.382	0.448	0.491
28	0.317	0.375	0.440	0.483
29	0.312	0.368	0.433	0.475
30	0.306	0.362	0.425	0.467
31	0.301	0.356	0.418	0.459
32	0.296	0.350	0.412	0.452
33	0.291	0.345	0.405	0.446
34	0.287	0.340	0.399	0.439
35	0.283	0.335	0.394	0.433
36	0.279	0.330	0.388	0.427
37	0.275	0.325	0.383	0.421
38	0.271	0.321	0.378	0.415
39	0.267	0.317	0.373	0.410
40	0.264	0.313	0.368	0.405
41	0.261	0.309	0.364	0.400
42	0.257	0.305	0.359	0.395
43	0.254	0.301	0.355	0.391
44	0.251	0.298	0.351	0.386
45	0.248	0.294	0.347	0.382
46	0.246	0.291	0.343	0.378
47	0.243	0.288	0.340	0.374
48	0.240	0.285	0.336	0.370
49	0.238	0.282	0.333	0.366
50	0.235	0.279	0.329	0.363

Source: Adapted from Zar, J. H. (1972). Significance testing of the Spearman rank correlation coefficient. *Journal of the American Statistical Association*, *67*, 578–580. Reprinted with permission from The *Journal of the American Statistical Association*.

TABLE B.8 Critical Values for the Pearson Product-Moment Correlation Coefficient *r*.

df	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.025$ $\alpha_{\text{one-tailed}} \leq 0.0125$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
1	0.988	0.997	0.999	0.999
2	0.900	0.950	0.975	0.990
3	0.805	0.878	0.924	0.959
4	0.729	0.811	0.868	0.917
5	0.669	0.754	0.817	0.875
6	0.621	0.707	0.771	0.834
7	0.582	0.666	0.732	0.798
8	0.549	0.632	0.697	0.765
9	0.521	0.602	0.667	0.735
10	0.497	0.576	0.640	0.708
11	0.476	0.553	0.616	0.684O
12	0.458	0.532	0.594	0.661
13	0.441	0.514	0.575	0.641
14	0.426	0.497	0.557	0.623
15	0.412	0.482	0.541	0.606
16	0.400	0.468	0.526	0.590
17	0.389	0.456	0.512	0.575
18	0.378	0.444	0.499	0.561
19	0.369	0.433	0.487	0.549
20	0.360	0.423	0.476	0.537
21	0.352	0.413	0.466	0.526
22	0.344	0.404	0.456	0.515
23	0.337	0.396	0.447	0.505
24	0.330	0.388	0.439	0.496
25	0.323	0.381	0.430	0.487
26	0.317	0.374	0.423	0.479
27	0.311	0.367	0.415	0.471
28	0.306	0.361	0.409	0.463
29	0.301	0.355	0.402	0.456
30	0.296	0.349	0.396	0.449
31	0.291	0.344	0.390	0.442
32	0.287	0.339	0.384	0.436
33	0.283	0.334	0.378	0.430
34	0.279	0.329	0.373	0.424
35	0.275	0.325	0.368	0.418
36	0.271	0.320	0.363	0.413
37	0.267	0.316	0.359	0.408
38	0.264	0.312	0.354	0.403
39	0.260	0.308	0.350	0.398
40	0.257	0.304	0.346	0.393
41	0.254	0.301	0.342	0.389

(*Continued*)

TABLE B.8 (*Continued*)

df	$\alpha_{\text{two-tailed}} \leq 0.10$ $\alpha_{\text{one-tailed}} \leq 0.05$	$\alpha_{\text{two-tailed}} \leq 0.05$ $\alpha_{\text{one-tailed}} \leq 0.025$	$\alpha_{\text{two-tailed}} \leq 0.025$ $\alpha_{\text{one-tailed}} \leq 0.0125$	$\alpha_{\text{two-tailed}} \leq 0.01$ $\alpha_{\text{one-tailed}} \leq 0.005$
42	0.251	0.297	0.338	0.384
43	0.248	0.294	0.334	0.380
44	0.246	0.291	0.330	0.376
45	0.243	0.288	0.327	0.372
46	0.240	0.285	0.323	0.368
47	0.238	0.282	0.320	0.365
48	0.235	0.279	0.317	0.361
49	0.233	0.276	0.314	0.358
50	0.231	0.273	0.311	0.354

TABLE B.9 Factorials.

n	*n*!
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40,320
9	362,880
10	3,628,800
11	39,916,800
12	479,001,600
13	6,227,020,800
14	87,178,291,200
15	1,307,674,368,000
16	20,922,789,888,000
17	355,687,428,096,000
18	6,402,373,705,728,000
19	121,645,100,408,832,000
20	2,432,902,008,176,640,000
21	51,090,942,171,709,440,000
22	1,124,000,727,777,607,680,000
23	25,852,016,738,884,976,640,000
24	620,448,401,733,239,439,360,000
25	15,511,210,043,330,985,984,000,000

TABLE B.10 Critical Values for the Runs Test for Randomness.

One-tailed alternative; $\alpha = 0.05$.

	n_2										
n_1	2	3	4	5	6	7	8	9	10	11	12
2	–	–	–	–	–	–	2	2	2	2	2
	–	–	–	–	–	–	–	–	–	–	–
3	–	–	–	2	2	2	2	2	3	3	3
	–	–	7	–	–	–	–	–	–	–	–
4	–	–	2	2	3	3	3	3	3	3	4
	–	7	8	9	9	9	–	–	–	–	–
5	–	2	2	3	3	3	3	4	4	4	4
	–	–	9	9	10	10	11	11	11	–	–
6	–	2	3	3	3	4	4	4	5	5	5
	–	–	9	10	11	11	12	12	12	13	13
7	–	2	3	3	4	4	4	5	5	5	6
	–	–	9	10	11	12	13	13	13	14	14
8	2	2	3	3	4	4	5	5	6	6	6
	–	–	–	11	12	13	13	14	14	15	15
9	2	2	3	4	4	5	5	6	6	6	7
	–	–	–	11	12	13	14	14	15	15	16
10	2	3	3	4	5	5	6	6	6	7	7
	–	–	–	11	12	13	14	15	16	16	17
11	2	3	3	4	5	5	6	6	7	7	8
	–	–	–	–	13	14	15	15	16	17	17
12	2	3	4	4	5	6	6	7	7	8	8
	–	–	–	–	13	14	15	16	17	17	18

One-tailed alternative; $\alpha = 0.025$.

	n_2										
n_1	2	3	4	5	6	7	8	9	10	11	12
2	–	–	–	–	–	–	–	–	–	–	2
	–	–	–	–	–	–	–	–	–	–	–
3	–	–	–	–	2	2	2	2	2	2	2
	–	–	–	–	–	–	–	–	–	–	–
4	–	–	–	2	2	2	3	3	3	3	3
	–	–	–	9	9	–	–	–	–	–	–
5	–	–	2	2	3	3	3	3	3	4	4
	–	–	9	10	10	11	11	–	–	–	–
6	–	2	2	3	3	3	3	4	4	4	4
	–	–	9	10	11	12	12	13	13	13	13
7	–	2	2	3	3	3	4	4	5	5	5
	–	–	–	11	12	13	13	14	14	14	14

(*Continued*)

TABLE B.10 (*Continued*)

One-tailed alternative; $\alpha = 0.025$.

	n_2										
n_1	2	3	4	5	6	7	8	9	10	11	12
8	–	2	3	3	3	4	4	5	5	5	6
	–	–	–	11	12	13	14	14	15	15	16
9	–	2	3	3	4	4	5	5	5	6	6
	–	–	–	–	13	14	14	15	16	16	16
10	–	2	3	3	4	5	5	5	6	6	7
	–	–	–	–	13	14	15	16	16	17	17
11	–	2	3	4	4	5	5	6	6	7	7
	–	–	–	–	13	14	15	16	17	17	18
12	2	2	3	4	4	5	6	6	7	7	7
	–	–	–	–	13	14	16	16	17	18	19

Source: Adapted from tables D.5 and D.6 of Janke, S. J., & Tinsley, F. C. (2005). *Introduction to Linear Models and Statistical Inference*. Hoboken, NJ: John Wiley & Sons, Inc. Reprinted with permission of John Wiley & Sons, Inc.

REFERENCES

American Psychological Association. (2001). *The Publication Manual of the American Psychological Association*, 5th ed. Washington, DC: American Psychological Association.

Anderson, N. H. (1961). Scales and statistics: Parametric and nonparametric. *Psychological Bulletin*, *58*(4), 305–316.

Belanger, N. D., & Desrochers, S. (2001). Can 6-month-old infants process causality in different types of causal events? *British Journal of Developmental Psychology*, *19*(1), 11–21.

Blumberg, F. C., & Sokol, L. M. (2004). Boys' and girls' use of cognitive strategy when learning to play video games. *The Journal of General Psychology*, *131*(2), 151–158.

Boser, J., & Poppen, W. A. (1978). Identification of teacher verbal response roles for improving student-teacher relationships. *Journal of Educational Research*, *72*(2), 91–94.

Bryant, B. K., & Trockel, J. F. (1976). Personal history of psychological stress related to locus of control orientation among college women. *Journal of Consulting and Clinical Psychology*, *44*(2), 266–271.

Bryce, T. G. K., & Blown, E. J. (2012). The novice-expert continuum in astronomy knowledge. *International Journal of Science Education*, *34*(4), 545–587.

Cady, J., Meier, S. L., & Lubinski, C. A. (2006). Developing mathematics teachers: The transition from preservice to experienced teacher. *The Journal of Educational Research*, *99*(5), 295–305.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York: Academic Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Daniel, W. W. (1990). *Applied Nonparametric Statistics*, 2nd ed. Boston: PWS-KENT Publishing Company.

Dorsey-Palmateer, R., & Smith, G. (2004). Bowlers' hot hands. *The American Statistician*, *58*(1), 38–45.

Duffy, R. D., & Sedlacek, W. E. (2007). The work values of first-year college students: Exploring group differences. *The Career Development Quarterly*, *55*(4), 359–364.

Enander, R. T., Gagnon, R. N., Hanumara, R. C., Park, E., Armstrong, T., & Gute, D. M. (2007). Environmental health practice: Statistically based performance measurement. *American Journal of Public Health*, *97*(5), 819–824.

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.

Ferrari, J. R., Athey, R. B., Moriarty, M. O., & Appleby, D. C. (2006). Education and employment among alumni academic honor society leaders. *Education*, *127*(2), 244–259.

Field, A. (2005). *Discovering Statistics Using SPSS*, 2nd ed. London: Sage Publications.

Finson, K. D., Pedersen, J., & Thomas, J. (2006). Comparing science teaching styles to students' perceptions of scientists. *School Science and Mathematics*, *106*(1), 8–15.

Fitzgerald, L. M., Delitto, A., & Irrgang, J. J. (2007). Validation of the clinical internship evaluation tool. *Physical Therapy*, *87*(7), 844–860.

Flannelly, K. J., Strock, A. L., & Weaver, A. J. (2005). A review of research on the effects of religion on adolescent tobacco use published between 1990 and 2003. *Adolescence*, *40*, 761–776.

Gaito, J. (1980). Measurement scales and statistics: Resurgence of an old misconception. *Psychological Bulletin*, *87*(3), 564–567.

Gibbons, J. D., & Chakraborti, S. (2010). *Nonparametric Statistical Inference*, 5th ed. Boca Raton, FL: Chapman and Hall/CRC.

Gömleksız, M. N., & Bulut, İ. (2007). An evaluation of the effectiveness of the new primary school mathematics curriculum in practice. *Educational Sciences: Theory & Practice*, *7*(1), 81–94.

Greiner, C. S., & Smith, B. (2006). Determining the effect of selected variables on teacher retention. *Education*, *126*(4), 653–659.

Hardré, P. L., Crowson, H. M., Xie, K., & Ly, C. (2007). Testing differential effects of computer-based, web-based and paper-based administration of questionnaire research instruments. *British Journal of Educational Technology*, *38*(1), 5–22.

Hastings, C. (1955). *Approximations for Digital Computers*. Princeton, NJ: Princeton University Press.

Hegedus, K. S. (1999). Providers' and consumers' perspective of nurses' caring behaviours. *Journal of Advanced Nursing*, *30*(5), 1090–1096.

Helsen, W., Gilis, B., & Weston, M. (2006). Errors in judging "offside" in association football: Test of the optical error versus the perceptual flash-lag hypothesis. *Journal of Sports Sciences*, *24*(5), 521–528.

Limb, G. E., & Organista, K. C. (2003). Comparisons between Caucasian students, students of color, and American Indian students on their views on social work's traditional mission, career motivations, and practice preferences. *Journal of Social Work Education*, *39*(1), 91–109.

Malthouse, E. (2001). Methodological and statistical concerns of the experimental behavioral researcher. *Journal of Consumer Psychology*, *10*(1/2), 111–112.

Mansell, S., Sobsey, D., & Moskal, R. (1998). Clinical findings among sexually abused children with and without developmental disabilities. *Mental Retardation*, *36*(1), 12–22.

Marston, D. (1996). A comparison of inclusion only, pull-out only, and combined service models for students with mild disabilities. *The Journal of Special Education*, *30*(2), 121–132.

McMillian, J., Morgan, S. A., & Ament, P. (2006). Acceptance of male registered nurses by female registered nurses. *Journal of Nursing Scholarship*, *38*(1), 100–106.

Nanna, M. J., & Sawilowsky, S. S. (1998). Analysis of Likert scale data in disability and medical rehabilitation research. *Psychological Methods*, *3*(1), 55–67.

Odaci, H. (2007). Depression, submissive behaviors and negative automatic thoughts in obese Turkish adolescents. *Social Behavior & Personality: An International Journal*, *35*(3), 409–416.

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*, *9*(6), 1–12. Retrieved: January 31, 2008 from http://PAREonline.net/getvn.asp?v=9&n=6

Plata, M., & Trusty, J. (2005). Effect of socioeconomic status on general and at-risk high school boys' willingness to accept same-sex peers with LD. *Adolescence*, *40*(157), 47–66.

Pett, M. A. (1997). *Nonparametric Statistics for Health Care Research: Statistics for Small Samples and Unusual Distributions*. Thousand Oaks, CA: Sage Publications.

Pollay, R. W., Lee, J. S., & Carter-Whitney, D. (1992). Separate, but not equal: Racial segmentation in cigarette advertising. *Journal of Advertising*, *21*(1), 45–57.

Re, A. M., Pedron, M., & Cornoldi, C. (2007). Expressive writing difficulties in children described as exhibiting ADHD symptoms. *Journal of Learning Disabilities*, *40*(3), 244–255.

Rimm-Kaufman, S. E., & Zhang, Y. (2005). Father-school communication in preschool and kindergarten. *The School Psychology Review*, *34*(3), 287–308.

Rinderknecht, K., & Smith, C. (2004). Social Cognitive Theory in an after-school nutrition intervention for urban Native American youth. *Journal of Nutrition Education & Behavior*, *36*(6), 298–304.

Salkind, N. J. (2004). *Statistics for People Who (Think They) Hate Statistics*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Savignon, S. J., & Sysoyev, P. V. (2002). Sociocultural strategies for a dialogue of cultures. *The Modern Language Journal*, *86*(4), 508–524.

Schulze, E., & Tomal, A. (2006). The chilly classroom: Beyond gender. *College Teaching*, *54*(3), 263–239.

Seiver, O. H., & Hatfield, T. H. (2002). Grading systems for retail food facilities: preference reversals of environmental health professionals. *Journal of Environmental Health*, *64*(10), 8–13.

Shorten, J., Seikel, M., & Ahrberg, J. H. (2005). Why do you still use Dewey? Academic libraries that continue with Dewey decimal classification. *Library Resources & Technical Services*, *49*(2), 123–136.

Siegel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. New York: McGraw-Hill.

Smirnov, N. V. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of the Mathematical Statistics*, *19*, 279–281.

Stake, R. E. (1960). Review of elementary statistics by P. G. Hoel. *Educational and Psychological Measurement*, *20*, 871–873.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*, 677–680.

Townsend, J. T., & Ashby, F. G. (1984). Measurement scales and statistics: The misconception misconceived. *Psychological Bulletin*, *96*(2), 394–401.

Vaughn, S., Reiss, M., Rothlein, L., & Hughes, M. T. (1999). Kindergarten teachers' perceptions of instructing students with disabilities. *Remedial and Special Education*, *20*(3), 184–191.

Velleman, P. F., & Wilkinson, L. (1993). Nominal, ordinal, interval, and ratio typologies are misleading. *The American Statistician*, *47*(1), 65–72.

Vergin, R. C. (2000). Winning streaks in sports and the misperception of momentum. *Journal of Sport Behavior*, *23*(2), 181–197.

Warner, R. M. (2008). *Applied Statistics: From Bivariate through Multivariate Techniques*. Thousand Oaks, CA: Sage Publications.

INDEX

Nonparametric Statistics: A Step-by-Step Approach, Second Edition. Gregory W. Corder and Dale I. Foreman.