CHAPTER

15

Correlation

Tools You Will Need

The following items are considered essential background material for this chapter. If you doubt your knowledge of any of these items, you should review the appropriate chapter or section before proceeding.

- Sum of squares (SS) (Chapter 4)
  - Computational formula
  - Definitional formula
- *z*-scores (Chapter 5)
- Hypothesis testing (Chapter 8)

Preview

Having been a student and taken exams for much of your life, you probably have noticed a curious phenomenon. In every class, there are some students who zip through exams and turn in their papers while everyone else is still on page 1. Other students cling to their exams and are still working frantically when the instructor announces that time is up and demands that all papers be turned in. Have you wondered what grades these students receive? Are the students who finish first the best in the class, or are they completely unprepared and simply accepting their failure? Are the *A* students the last to finish because they are compulsively checking and rechecking their answers? To help answer these questions, we carefully observed a recent exam and recorded the amount of time each student spent on the exam and the grade each student received. These data are shown in Figure 15.1. Note that we have listed time along the *X*-axis and grade on the *Y*-axis. Each student is identified by a point on the graph that is located directly above the student's time and directly across from the student's grade. Also note that we have drawn a line through the middle of the data points in Figure 15.1. The line helps make the relationship between time and grade more obvious. The graph shows that the highest grades tend to go to the students who finished their exams early. Students who held their papers to the bitter end tended to have low grades.

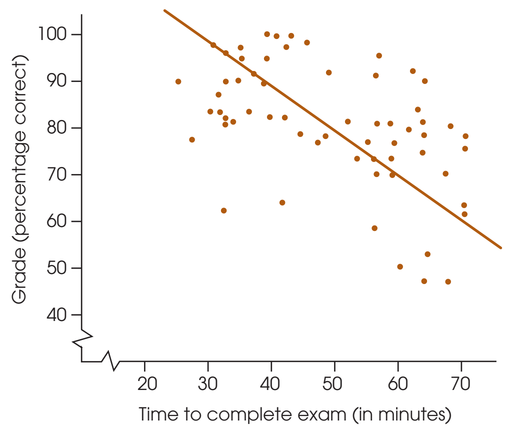


FIGURE 15.1

The relationship between exam grade and time needed to complete the exam. Notice the general trend in these data: Students who finish the exam early tend to have better grades.

The Problem: Although the data in Figure 15.1 appear to show a clear relationship, we need some procedure to measure the relationship and a hypothesis test to determine whether it is significant. In the preceding five chapters, we described relationships between variables in terms of mean differences between two or more groups of scores, and we used hypothesis tests that evaluate the significance of mean differences. For the data in Figure 15.1, there is only one group of scores, and calculating a mean is not going to help describe the relationship. To evaluate these data, a completely different approach is needed for both descriptive and inferential statistics.

The Solution: The data in Figure 15.1 are an example of the results from a correlational research study. In Chapter 1, the correlational design was introduced as a method for examining the relationship between two variables by measuring two different variables for each individual in one group of participants. The relationship obtained in a correlational study is typically described and evaluated with a statistical measure known as a *correlation*. Just as a sample mean provides a concise description of an entire sample, a correlation provides a concise description of a relationship. We look at how correlations are used and interpreted. For example, now that you have seen the relationship between time and grades, do you think it might be a good idea to start turning in your exam papers a little sooner? Wait and see.

15.1 INTRODUCTION

Correlation is a statistical technique that is used to measure and describe the relationship between two variables. Usually the two variables are simply observed as they exist naturally in the environment—there is no attempt to control or manipulate the variables. For example, a researcher could check high school records (with permission) to obtain a measure of each student's academic performance, and then survey each family to obtain a measure of income. The resulting data could be used to determine whether there is relationship between high school grades and family income. Notice that the researcher is not manipulating any student's grade or any family's income, but is simply observing what occurs naturally. You also should notice that a correlation requires two scores for each individual (one score from each of the two variables). These scores normally are identified as X and Y. The pairs of scores can be listed in a table, or they can be presented graphically in a scatter plot (Figure 15.2). In the scatter plot, the values for the X variable are listed on the horizontal axis and the Y values are listed on the vertical axis. Each individual is then represented by a single point in the graph so that the horizontal position corresponds to the individual's X value and the vertical position corresponds to the Y value. The value of a scatter plot is that it allows you to see any patterns or trends that exist in the data. The scores in Figure 15.2, for example, show a clear relationship between family income and student grades; as income increases, grades also increase.

Person	Family Income (in $1000)	Student's Average Grade
A	31	72
B	38	86
C	42	81
D	44	78
E	49	85
F	56	80
G	58	91
H	65	89
I	70	94
J	90	83
K	92	90
L	106	97
M	135	89
N	174	95

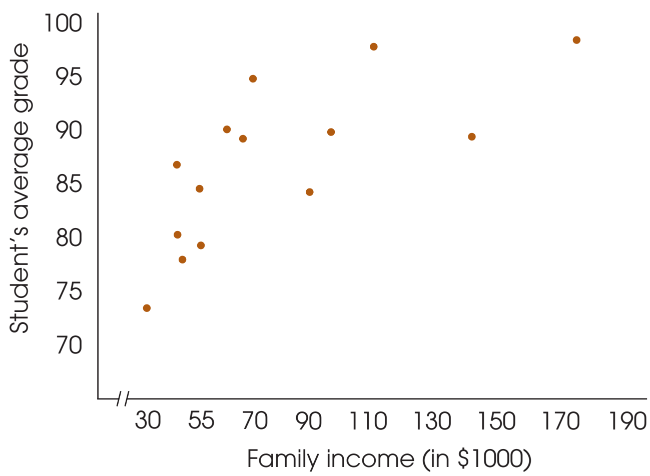


FIGURE 15.2

Correlational data showing the relationship between family income (X) and student grades (Y) for a sample of $n = 14$ high school students. The scores are listed in order from lowest to highest family income and are shown in a scatter plot.

THE CHARACTERISTICS OF A RELATIONSHIP

A correlation is a numerical value that describes and measures three characteristics of the relationship between X and Y. These three characteristics are as follows:

1. **The Direction of the Relationship.** The sign of the correlation, positive or negative, describes the direction of the relationship.

DEFINITIONS

In a **positive correlation,** the two variables tend to change in the same direction: As the value of the *X* variable increases from one individual to another, the *Y* variable also tends to increase; when the *X* variable decreases, the *Y* variable also decreases.

In a **negative correlation,** the two variables tend to go in opposite directions. As the *X* variable increases, the *Y* variable decreases. That is, it is an inverse relationship.

The following examples illustrate positive and negative relationships.

EXAMPLE 15.1

Suppose you run the drink concession at the football stadium. After several seasons, you begin to notice a relationship between the temperature at game time and the beverages you sell. Specifically, you have noted that when the temperature is low, you sell relatively little beer. However, as the temperature goes up, beer sales also go up (Figure 15.3). This is an example of a positive correlation. You also have noted a relationship between temperature and coffee sales: On cold days you sell a lot of coffee, but coffee sales go down as the temperature goes up. This is an example of a negative relationship.

2. **The Form of the Relationship.** In the preceding coffee and beer examples, the relationships tend to have a linear form; that is, the points in the scatter plot tend to cluster around a straight line. We have drawn a line through the middle

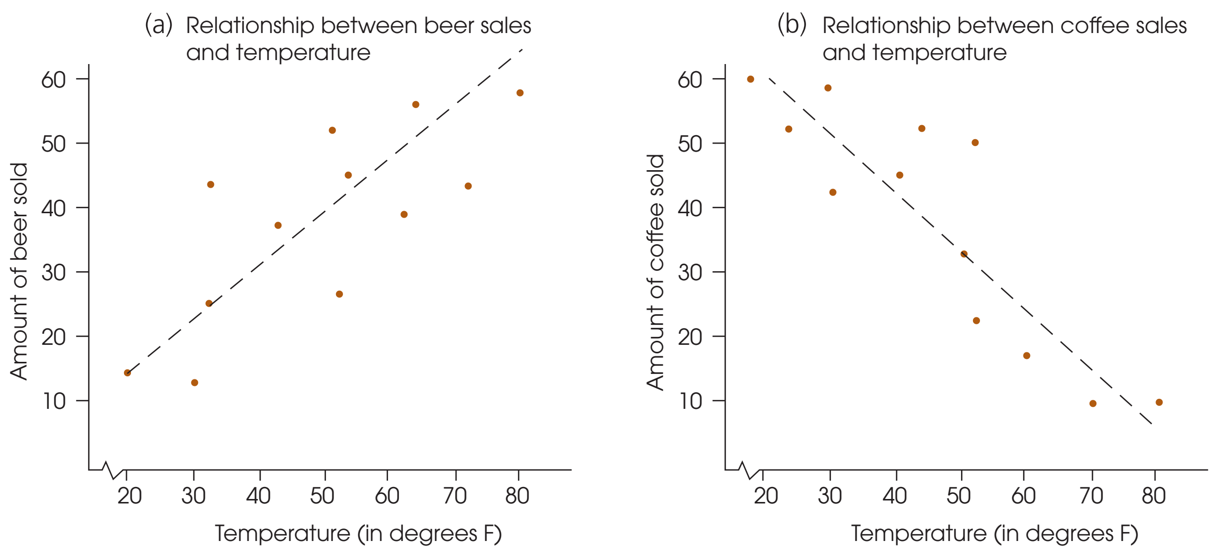


FIGURE 15.3

Examples of positive and negative relationships. (a) Beer sales are positively related to temperature. (b) Coffee sales are negatively related to temperature.

of the data points in each figure to help show the relationship. The most common use of correlation is to measure straight-line relationships. However, other forms of relationships do exist and there are special correlations used to measure them. (We examine alternatives in Section 15.5.)

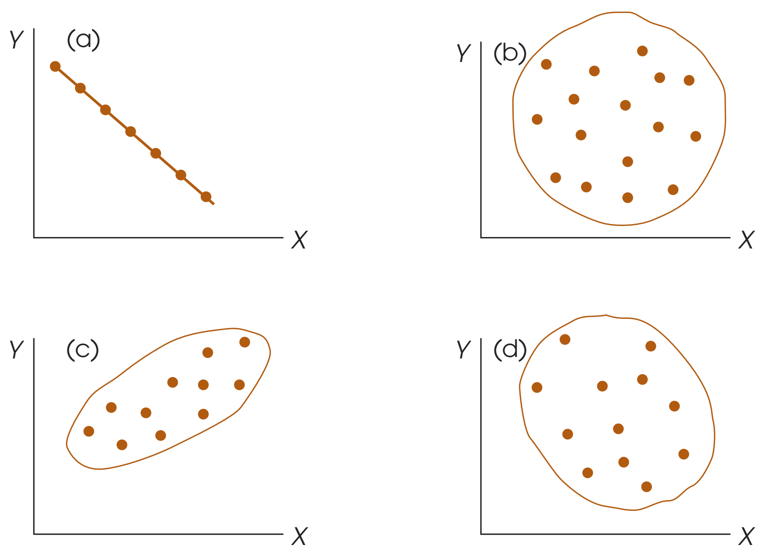
3. **The Strength or Consistency of the Relationship**. Finally, the correlation measures the consistency of the relationship. For a linear elationship, for example, the data points could fit perfectly on a straight line. Every time X increases by one point, the value of Y also changes by a consistent and predictable amount. Figure 15.4(a) shows an example of a perfect linear relationship. However, relationships are usually not perfect. Although there may be a tendency for the value of Y to increase whenever X increases, the amount that Y changes is not always the same, and occasionally, Y decreases when X increases. In this situation, the data points do not fall perfectly on a straight line. The consistency of the relationship is measured by the numerical value of the correlation. A *perfect correlation* always is identified by a correlation of 1.00 and indicates a perfectly consistent relationship. For a correlation of 1.00 (or –1.00), each change in X is accompanied by a perfectly predictable change in Y. At the other extreme, a correlation of 0 indicates no consistency at all. For a correlation of 0, the data points are scattered randomly with no clear trend [see Figure 15.4(b)]. Intermediate values between 0 and 1 indicate the degree of consistency.

   Examples of different values for linear correlations are shown in Figure 15.4. In each example we have sketched a line around the data points. This line, called an *envelope* because it encloses the data, often helps you to see the overall trend in the data. As a rule of thumb, when the envelope is shaped roughly like a football, the correlation is around 0.7. Envelopes that are fatter than a football indicate correlations closer to 0, and narrower shapes indicate correlations closer to 1.00.

You should also note that the sign (+ or −) and the strength of a correlation are independent. For example, a correlation of 1.00 indicates a perfectly consistent relationship whether it is positive (+1.00) or negative (−1.00). Similarly, correlations of +0.80 and −0.80 are equally consistent relationships.

FIGURE 15.4

Examples of different values for linear correlations: (a) a perfect negative correlation, −1.00; (b) no linear trend, 0.00; (c) a strong positive relationship, approximately +0.90; (d) a relatively weak negative correlation, approximately −0.40.



LEARNING CHECK

1. For each of the following, indicate whether you would expect a positive or a negative correlation.
   a. Model year and price for a used Honda
   b. IQ and grade point average for high school students
   c. Daily high temperature and daily energy consumption for 30 winter days in New York City
2. The data points would be clustered more closely around a straight line for a correlation of –0.80 than for a correlation of +0.05. (True or false?)
3. If the data points are clustered close to a line that slopes up from left to right, then a good estimate of the correlation would be +0.90. (True or false?)
4. If a scatter plot shows a set of data points that form a circular pattern, the correlation should be near zero. (True or false?)

ANSWERS

1. **a.** Positive: Higher model years tend to have higher prices.
   b. Positive: More intelligent students tend to get higher grades.
   c. Negative: Higher temperature tends to decrease the need for heating.
2. True. The numerical value indicates the strength of the relationship. The sign only indicates direction.
3. True.
4. True.

15.2 THE PEARSON CORRELATION

By far the most common correlation is the *Pearson correlation* (or the Pearson product–moment correlation) which measures the degree of straight-line relationship.

DEFINITION

The **Pearson correlation** measures the degree and the direction of the linear relationship between two variables.

The Pearson correlation is identified by the letter r. Conceptually, this correlation is computed by

$$r = \frac{\text{degree to which } X \text{ and } Y \text{ vary together}}{\text{degree to which } X \text{ and } Y \text{ vary separately}}$$

$$= \frac{\text{covariability of } X \text{ and } Y}{\text{variability of } X \text{ and } Y \text{ separately}}$$

When there is a perfect linear relationship, every change in the X variable is accompanied by a corresponding change in the Y variable. In Figure 15.4(a), for example, every time the value of X increases, there is a perfectly predictable decrease in the value of Y. The result is a perfect linear relationship, with X and Y always varying together. In this case, the covariability (X and Y together) is identical to the variability of X and Y separately, and the formula produces a correlation with a magnitude of 1.00 or –1.00. At the other extreme, when there is no linear relationship, a change in the

X variable does not correspond to any predictable change in the Y variable. In this case, there is no covariability, and the resulting correlation is zero.

THE SUM OF PRODUCTS OF DEVIATIONS

To calculate the Pearson correlation, it is necessary to introduce one new concept: the *sum of products* of deviations, or SP. This new value is similar to SS (the sum of squared deviations), which is used to measure variability for a single variable. Now, we use SP to measure the amount of covariability between two variables. The value for SP can be calculated with either a definitional formula or a computational formula.

The *definitional formula* for the sum of products is

$$SP = \Sigma(X - M_X)(Y - M_Y) \tag{15.1}$$

where M_X is the mean for the X scores and M_Y is the mean for the Ys.

The definitional formula instructs you to perform the following sequence of operations:

1. Find the X deviation and the Y deviation for each individual.
2. Find the product of the deviations for each individual.
3. Add the products.

Notice that this process "defines" the value being calculated: the sum of the products of the deviations.

The *computational formula* for the sum of products of deviations is

$$SP = \Sigma XY - \frac{\Sigma X \Sigma Y}{n} \tag{15.2}$$

Caution: The n in this formula refers to the number of pairs of scores.

Because the computational formula uses the original scores (X and Y values), it usually results in easier calculations than those required with the definitional formula, especially if M_X or M_Y is not a whole number. However, both formulas always produce the same value for SP.

You may have noted that the formulas for SP are similar to the formulas you have learned for SS (sum of squares). The relationship between the two sets of formulas is described in Box 15.1. The following example demonstrates the calculation of SP with both formulas.

EXAMPLE 15.2

The same set of $n = 4$ pairs of scores is used to calculate SP, first using the definitional formula and then using the computational formula.

For the definitional formula, you need deviation scores for each of the X values and each of the Y values. Note that the mean for the Xs is $M_X = 3$ and the mean for the Ys is $M_Y = 5$. The deviations and the products of deviations are shown in the following table:

Caution: The signs (+ and –) are critical in determining the sum of products, SP.

Scores		Deviations		Products
X	Y	$X - M_X$	$Y - M_Y$	$(X - M_X)(Y - M_Y)$
1	3	−2	−2	+4
2	6	−1	+1	−1
4	4	+1	−1	−1
5	7	+2	+2	+4
				$+6 = SP$

For these scores, the sum of the products of the deviations is $SP = +6$.

For the computational formula, you need the X value, the Y value, and the XY product for each individual. Then you find the sum of the Xs, the sum of the Ys, and the sum of the XY products. These values are as follows:

X	Y	XY	
1	3	3	
2	6	12	
4	4	16	
5	7	35	
12	20	66	Totals

Substituting the totals in the formula gives

$$SP = \Sigma XY - \frac{\Sigma X \Sigma Y}{n}$$

$$= 66 - \frac{12(20)}{4}$$

$$= 66–60$$

$$= 6$$

Both formulas produce the same result, $SP = 6$.

BOX 15.1 COMPARING THE *SP* AND *SS* FORMULAS

It will help you to learn the formulas for *SP* if you note the similarity between the two *SP* formulas and the corresponding formulas for *SS* that were presented in Chapter 4. The definitional formula for *SS* is

$$SS = \Sigma(X - M)^2$$

In this formula, you must square each deviation, which is equivalent to multiplying it by itself. With this in mind, the formula can be rewritten as

$$SS = \Sigma(X - M)(X - M)$$

The similarity between the *SS* formula and the *SP* formula should be obvious—the *SS* formula uses squares and the *SP* formula uses products. This same relationship exists for the computational formulas. For *SS*, the computational formula is

$$SS = \Sigma X^2 - \frac{(\Sigma X)^2}{n}$$

As before, each squared value can be rewritten so that the formula becomes

$$SS = \Sigma XX - \frac{\Sigma X \Sigma X}{n}$$

Again, note the similarity in structure between the *SS* formula and the *SP* formula. If you remember that *SS* uses squares and *SP* uses products, the two new formulas for the sum of products should be easy to learn.

CALCULATION OF THE PEARSON-CORRELATION

As noted earlier, the Pearson correlation consists of a ratio comparing the covariability of X and Y (the numerator) with the variability of X and Y separately (the denominator). In the formula for the Pearson r, we use SP to measure the covariability of X and Y. The variability of X is measured by computing SS for the X scores and the variability of Y is measured by SS for the Y scores. With these definitions, the formula for the Pearson correlation becomes

Note that you *multiply* SS for X by SS for Y in the denominator of the Pearson formula.

$$r = \frac{SP}{\sqrt{SS_X SS_Y}} \tag{15.3}$$

The following example demonstrates the use of this formula with a simple set of scores.

EXAMPLE 15.3

X	Y
0	2
10	6
4	2
8	4
8	6

The Pearson correlation is computed for the set of $n = 5$ pairs of scores shown in the margin.

Before starting any calculations, it is useful to put the data in a scatter plot and make a preliminary estimate of the correlation. These data have been graphed in Figure 15.5. Looking at the scatter plot, it appears that there is a very good (but not perfect) positive correlation. You should expect an approximate value of $r = +0.8$ or $+0.9$. To find the Pearson correlation, we need SP, SS for X, and SS for Y. The calculations for each of these values, using the definitional formulas, are presented in Table 15.1. (Note that the mean for the X values is $M_X = 6$ and the mean for the Y scores is $M_Y = 4$.)

Using the values from Table 15.1, the Pearson correlation is

$$r = \frac{SP}{\sqrt{(SS_X)(SS_Y)}} = \frac{28}{\sqrt{(64)(16)}} = \frac{28}{32} = +0.875$$

Note that the value we obtained for the correlation is perfectly consistent with the pattern shown in Figure 15.5. First, the positive value of the correlation indicates

FIGURE 15.5
Scatter plot of the data from Example 15.3.

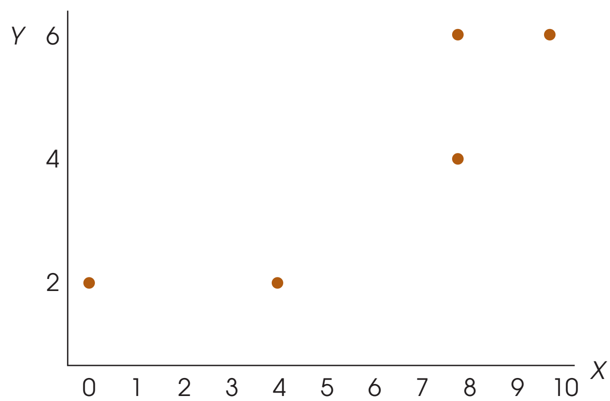


TABLE 15.1

Calculation of SS_X, SS_Y, and SP for a sample of $n = 5$ pairs of scores.

Scores		Deviations		Squared Deviations		Products
X	Y	$X - M_X$	$Y - M_Y$	$(X - M_X)^2$	$(Y - M_Y)^2$	$(X - M_X)(Y - M_Y)$
0	2	−6	−2	36	4	+12
10	6	+4	+2	16	4	+8
4	2	−2	−2	4	4	+4
8	4	+2	0	4	0	0
8	6	+2	+2	4	4	+4
				$SS_X = 64$	$SS_Y = 16$	$SP = +28$

that the points are clustered around a line that slopes up to the right. Second, the high value for the correlation (near 1.00) indicates that the points are very tightly clustered close to the line. Thus, the value of the correlation describes the relationship that exists in the data.

THE PEARSON CORRELATION AND *z*-SCORES

The Pearson correlation measures the relationship between an individual's location in the X distribution and his or her location in the Y distribution. For example, a positive correlation means that individuals who score high on X also tend to score high on Y. Similarly, a negative correlation indicates that individuals with high X scores tend to have low Y scores.

Recall from Chapter 5 that z-scores identify the exact location of each individual score within a distribution. With this in mind, each X value can be transformed into a z-score, z_X, using the mean and standard deviation for the set of Xs. Similarly, each Y score can be transformed into z_Y. If the X and Y values are viewed as a sample, the transformation is completed using the sample formula for z (Equation 5.3). If the X and Y values form a complete population, the z-scores are computed using Equation 5.1. After the transformation, the formula for the Pearson correlation can be expressed entirely in terms of z-scores.

$$\text{For a sample, } r = \frac{\Sigma z_X z_Y}{(n - 1)} \tag{15.4}$$

$$\text{For a population, } \rho = \frac{\Sigma z_X z_Y}{N} \tag{15.5}$$

Note that the population value is identified with a Greek letter, in this case the letter rho (ρ), which is the Greek equivalent of the letter r.

LEARNING CHECK

1. Describe what is measured by a Pearson correlation.
2. Can SP ever have a value less than zero?
3. Calculate the sum of products of deviations (SP) for the following set of scores. Use the definitional formula and then the computational formula. Verify that you get the same answer with both formulas.

X	Y
0	1
4	3
5	3
2	2
4	1

4. For the following data:

a. Sketch a scatter plot and make an estimate of the Pearson correlation.

b. Compute the Pearson correlation.

X	Y
2	6
1	5
3	3
0	7
4	4

ANSWERS

1. The Pearson correlation measures the degree and direction of the linear relationship between two variables.

2. Yes. *SP* can be positive, negative, or zero depending on the relationship between X and Y.

3. $SP = 5$

4. $r = -8/10 = -0.80$

15.3 USING AND INTERPRETING THE PEARSON CORRELATION

WHERE AND WHY CORRELATIONS ARE USED

Although correlations have a number of different applications, a few specific examples are presented next to give an indication of the value of this statistical measure.

1. **Prediction.** If two variables are known to be related in a systematic way, then it is possible to use one of the variables to make accurate predictions about the other. For example, when you applied for admission to college, you were required to submit a great deal of personal information, including your scores on the Scholastic Achievement Test (SAT). College officials want this information so that they can predict your chances of success in college. It has been demonstrated over several years that SAT scores and college grade point averages are correlated. Students who do well on the SAT tend to do well in college; students who have difficulty with the SAT tend to have difficulty in college. Based on this relationship, college admissions officers can make a prediction about the potential success of each applicant. You should note that this prediction is not perfectly accurate. Not everyone who does poorly on the

SAT has trouble in college. That is why you also submit letters of recommendation, high school grades, and other information with your application.

2. **Validity.** Suppose that a psychologist develops a new test for measuring intelligence. How could you show that this test truly measures what it claims; that is, how could you demonstrate the validity of the test? One common technique for demonstrating validity is to use a correlation. If the test actually measures intelligence, then the scores on the test should be related to other measures of intelligence—for example, standardized IQ tests, performance on learning tasks, problem-solving ability, and so on. The psychologist could measure the correlation between the new test and each of these other measures of intelligence to demonstrate that the new test is valid.
3. **Reliability.** In addition to evaluating the validity of a measurement procedure, correlations are used to determine reliability. A measurement procedure is considered reliable to the extent that it produces stable, consistent measurements. That is, a reliable measurement procedure produces the same (or nearly the same) scores when the same individuals are measured twice under the same conditions. For example, if your IQ was measured as 113 last week, you would expect to obtain nearly the same score if your IQ was measured again this week. One way to evaluate reliability is to use correlations to determine the relationship between two sets of measurements. When reliability is high, the correlation between two measurements should be strong and positive. Further discussion of the concept of reliability is presented in Box 15.2.
4. **Theory Verification.** Many psychological theories make specific predictions about the relationship between two variables. For example, a theory may predict a relationship between brain size and learning ability; a developmental theory may predict a relationship between the parents' IQs and the child's IQ; a social psychologist may have a theory predicting a relationship between personality type and behavior in a social situation. In each case, the prediction of the theory could be tested by determining the correlation between the two variables.

INTERPRETING CORRELATIONS

When you encounter correlations, there are four additional considerations that you should bear in mind:

1. Correlation simply describes a relationship between two variables. It does not explain why the two variables are related. Specifically, a correlation should not and cannot be interpreted as proof of a cause-and-effect relationship between the two variables.
2. The value of a correlation can be affected greatly by the range of scores represented in the data.
3. One or two extreme data points, often called *outliers,* can have a dramatic effect on the value of a correlation.
4. When judging how "good" a relationship is, it is tempting to focus on the numerical value of the correlation. For example, a correlation of +0.5 is halfway between 0 and 1.00 and, therefore, appears to represent a moderate degree of relationship. However, a correlation should not be interpreted as a proportion. Although a correlation of 1.00 does mean that there is a 100% perfectly predictable relationship between X and Y, a correlation of 0.5 does not mean that you can make predictions with 50% accuracy. To describe how accurately one variable predicts the other, you must square the correlation. Thus, a

BOX 15.2 RELIABILITY AND ERROR IN MEASUREMENT

The idea of reliability of measurement is tied directly to the notion that each individual measurement includes an element of error. Expressed as an equation,

$$\text{measured score} = \text{true score} + \text{error}$$

For example, if I try to measure your intelligence with an IQ test, the score that I get is determined partially by your actual level of intelligence (your true score) but it also is influenced by a variety of other factors such as your current mood, your level of fatigue, your general health, and so on. These other factors are lumped together as *error,* and are typically a part of any measurement.

It is generally assumed that the error component changes randomly from one measurement to the next and that this causes your score to change. For example, your IQ score is likely to be higher when you are well rested and feeling good compared to a measurement that is taken when you are tired and depressed. Although your actual intelligence hasn't changed, the error component causes your score to change from one measurement to another.

As long as the error component is relatively small, then your scores will be relatively consistent from one measurement to the next, and the measurements are said to be reliable. If you are feeling especially happy and well rested, it may affect your IQ score by a few points, but it is not going to boost your IQ from 110 to 170.

On the other hand, if the error component is relatively large, then you will find huge differences from one measurement to the next and the measurements are not reliable. Measurements of reaction time, for example, tend to be very unreliable. Suppose, for example, that you are seated at a desk with your finger on a button and a light bulb in front of you. Your job is to push the button as quickly as possible when the light goes on. On some trials, you are focused on the light with your finger tensed and ready to push. On other trials, you are distracted, or day dreaming, or blink when the light goes on so that time passes before you finally push the button. As a result, there is a huge error component to the measurement and your reaction time can change dramatically from one trial to the next. When measurements are unreliable you cannot trust any single measurement to provide an accurate indication of the individual's true score. To deal with this problem, researchers typically measure reaction time repeatedly and then average it over a large number of measurements.

Correlations can be used to help researchers measure and describe reliability. By taking two measurements for each individual, it is possible to compute the correlation between the first score and the second score. A strong, positive correlation indicates a good level of reliability: people who scored high on the first measurement also scored high on the second. A weak correlation indicates that there is not a consistent relationship between the first score and the second score; that is, a weak correlation indicates poor reliability.

correlation of $r = .5$ means that one variable *partially* predicts the other, but the predictable portion is only $r^2 = 0.5^2 = 0.25$ (or 25%) of the total variability.

We now discuss each of these four points in detail.

CORRELATION AND CAUSATION

One of the most common errors in interpreting correlations is to assume that a correlation necessarily implies a cause-and-effect relationship between the two variables. (Even Pearson blundered by asserting causation from correlational data [Blum, 1978].) We constantly are bombarded with reports of relationships: Cigarette smoking is related to heart disease; alcohol consumption is related to birth defects; carrot consumption is related to good eyesight. Do these relationships mean that cigarettes cause heart disease or carrots cause good eyesight? The answer is *no*. Although there may be a causal relationship, the simple existence of a correlation does not prove it. Earlier, for example, we discussed a study showing a relationship between high school grades and family income. However, this result does not mean that having a higher family income *causes* students to get better grades. For example, if mom gets an unexpected bonus at work, it is unlikely that her child's grades will also show a sudden increase. To establish a cause-and-effect relationship, it is necessary to conduct a true experiment (see p. 14) in which one variable is manipulated by a researcher and other variables are rigorously controlled. The fact that a correlation does not establish causation is demonstrated in the following example.

EXAMPLE 15.4

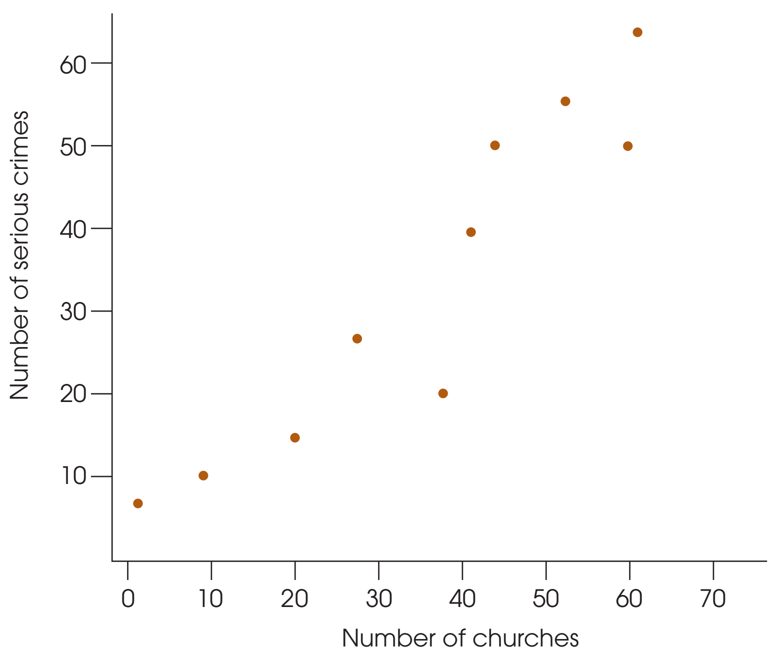
Suppose we select a variety of different cities and towns throughout the United States and measure the number of churches (*X* variable) and the number of serious crimes (*Y* variable) for each. A scatter plot showing hypothetical data for this study is presented in Figure 15.6. Notice that this scatter plot shows a strong, positive correlation between churches and crime. You also should note that these are realistic data. It is reasonable that small towns would have less crime and fewer churches and that large cities would have large values for both variables. Does this relationship mean that churches cause crime? Does it mean that crime causes churches? It should be clear that both answers are no. Although a strong correlation exists between number of churches and crime, the real cause of the relationship is the size of the population.

CORRELATION AND RESTRICTED RANGE

Whenever a correlation is computed from scores that do not represent the full range of possible values, you should be cautious in interpreting the correlation. Suppose, for example, that you are interested in the relationship between IQ and creativity. If you select a sample of your fellow college students, your data probably will represent only a limited range of IQ scores (most likely from 110 to 130). The correlation

FIGURE 15.6

Hypothetical data showing the logical relationship between the number of churches and the number of serious crimes for a sample of U.S. cities.



within this *restricted range* could be completely different from the correlation that would be obtained from a full range of IQ scores. For example, Figure 15.7 shows a strong positive relationship between X and Y when the entire range of scores is considered. However, this relationship is obscured when the data are limited to a restricted range.

To be safe, you should not generalize any correlation beyond the range of data represented in the sample. For a correlation to provide an accurate description for the general population, there should be a wide range of X and Y values in the data.

OUTLIERS

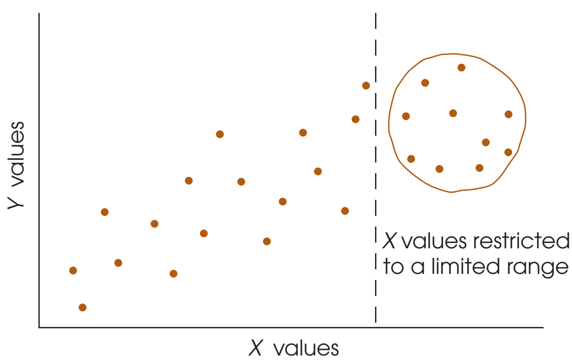
An outlier is an individual with X and/or Y values that are substantially different (larger or smaller) from the values obtained for the other individuals in the data set. The data point of a single outlier can have a dramatic influence on the value obtained for the correlation. This effect is illustrated in Figure 15.8. Figure 15.8(a) shows a set of $n = 5$ data points for which the correlation between the X and Y variables is nearly zero (actually $r = -0.08$). In Figure 15.8(b), one extreme data point (14, 12) has been added to the original data set. When this outlier is included in the analysis, a strong, positive correlation emerges (now $r = +0.85$). Note that the single outlier drastically alters the value for the correlation and, thereby, can affect one's interpretation of the relationship between variables X and Y. Without the outlier, one would conclude there is no relationship between the two variables. With the extreme data point, $r = +0.85$, which implies a strong relationship with Y increasing consistently as X increases. The problem of outliers is a good reason for looking at a scatter plot, instead of simply basing your interpretation on the numerical value of the correlation. If you only "go by the numbers," you might overlook the fact that one extreme data point inflated the size of the correlation.

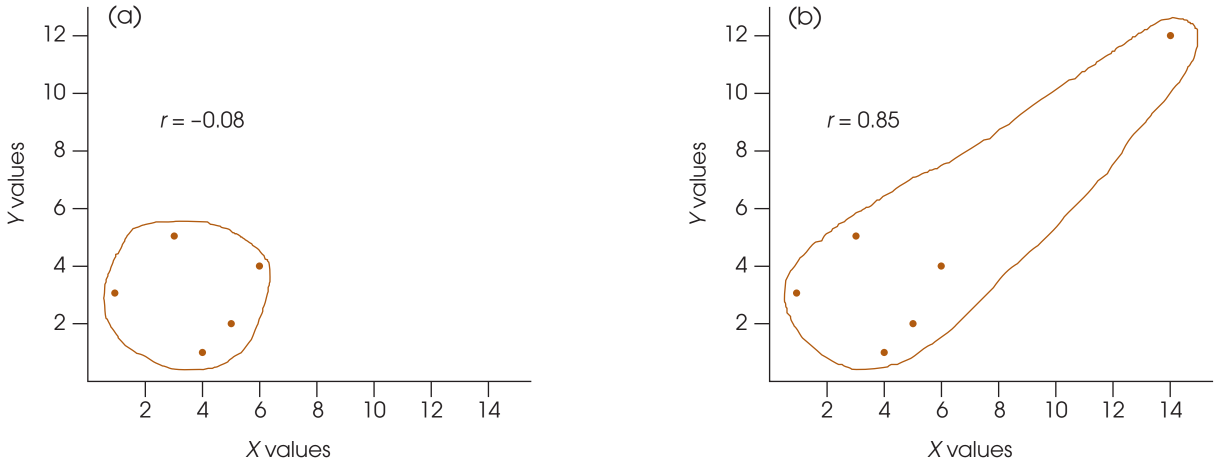
CORRELATION AND THE STRENGTH OF THE RELATIONSHIP

A correlation measures the degree of relationship between two variables on a scale from 0 to 1.00. Although this number provides a measure of the degree of relationship, many researchers prefer to square the correlation and use the resulting value to measure the strength of the relationship.

One of the common uses of correlation is for prediction. If two variables are correlated, you can use the value of one variable to predict the other. For example, college admissions officers do not just guess which applicants are likely to do well; they use other variables (SAT scores, high school grades, and so on) to predict which students are most likely to be successful. These predictions are based on correlations. By using

FIGURE 15.7
In this example, the full range of X and Y values shows a strong, positive correlation, but the restricted range of scores produces a correlation near zero.





Original Data		
Subject	*X*	*Y*
A	1	3
B	3	5
C	6	4
D	4	1
E	5	2

Data with Outlier Included		
Subject	*X*	*Y*
A	1	3
B	3	5
C	6	4
D	4	1
E	5	2
F	14	12

FIGURE 15.8

A demonstration of how one extreme data point (an outlier) can influence the value of a correlation.

correlations, the admissions officers expect to make more accurate predictions than would be obtained by chance. In general, the squared correlation (r^2) measures the gain in accuracy that is obtained from using the correlation for prediction. The squared correlation measures the proportion of variability in the data that is explained by the relationship between *X* and *Y*. It is sometimes called the *coefficient of determination.*

D E F I N I T I O N

The value r^2 is called the **coefficient of determination** because it measures the proportion of variability in one variable that can be determined from the relationship with the other variable. A correlation of $r = 0.80$ (or –0.80), for example, means that $r^2 = 0.64$ (or 64%) of the variability in the *Y* scores can be predicted from the relationship with *X*.

In earlier chapters (see pp. 299, 328, and 361) we introduced r^2 as a method for measuring effect size for research studies where mean differences were used to compare treatments. Specifically, we measured how much of the variance in the scores was accounted for by the differences between treatments. In experimental terminology, r^2 measures how much of the variance in the dependent variable is accounted for by the independent variable. Now we are doing the same thing, except that there is no independent or dependent variable. Instead, we simply have two variables, *X* and *Y*, and we use r^2 to measure how much of the variance in one variable can be determined from its relationship with the other variable. The following example demonstrates this concept.

EXAMPLE 15.5

Figure 15.9 shows three sets of data representing different degrees of linear relationship. The first set of data [Figure 15.9(a)] shows the relationship between IQ and shoe size. In this case, the correlation is $r = 0$ (and $r^2 = 0$), and you have no ability to predict a person's IQ based on his or her shoe size. Knowing a person's shoe size provides no information (0%) about the person's IQ. In this case, shoe size provides no help explaining why different people have different IQs.

Now consider the data in Figure 15.9(b). These data show a moderate, positive correlation, $r = +0.60$, between IQ scores and college grade point averages (GPA). Students with high IQs tend to have higher grades than students with low IQs. From this relationship, it is possible to predict a student's GPA based on his or her IQ. However, you should realize that the prediction is not perfect. Although students with high IQs *tend* to have high GPAs, this is not always true. Thus, knowing a student's IQ provides some information about the student's grades, or knowing a student's grades provides some information about the student's IQ. In this case, IQ scores help explain the fact that different students have different GPAs. Specifically, you can say that *part* of the differences in GPA are accounted for by IQ. With a correlation of $r = +0.60$, we obtain $r^2 = 0.36$, which means that 36% of the variance in GPA can be explained by IQ.

Finally, consider the data in Figure 15.9(c). This time we show a perfect linear relationship ($r = +1.00$) between monthly salary and yearly salary for a group of college employees. With $r = 1.00$ and $r^2 = 1.00$, there is 100% predictability. If you know a person's monthly salary, you can predict perfectly the person's annual salary. If two people have different annual salaries, the difference can be completely explained (100%) by the difference in their monthly salaries.

Just as r^2 was used to evaluate effect size for mean differences in Chapters 9, 10, and 11, r^2 can now be used to evaluate the size or strength of the correlation. The same standards that were introduced in Table 9.3 (p. 299), apply to both uses of the r^2 measure. Specifically, an r^2 value of 0.01 indicates a small effect or a small correlation, an r^2 value of 0.09 indicates a medium correlation, and r^2 of 0.25 or larger indicates a large correlation.

More information about the coefficient of determination (r^2) is presented in Section 15.5 and in Chapter 16. For now, you should realize that whenever two variables are consistently related, it is possible to use one variable to predict values for the

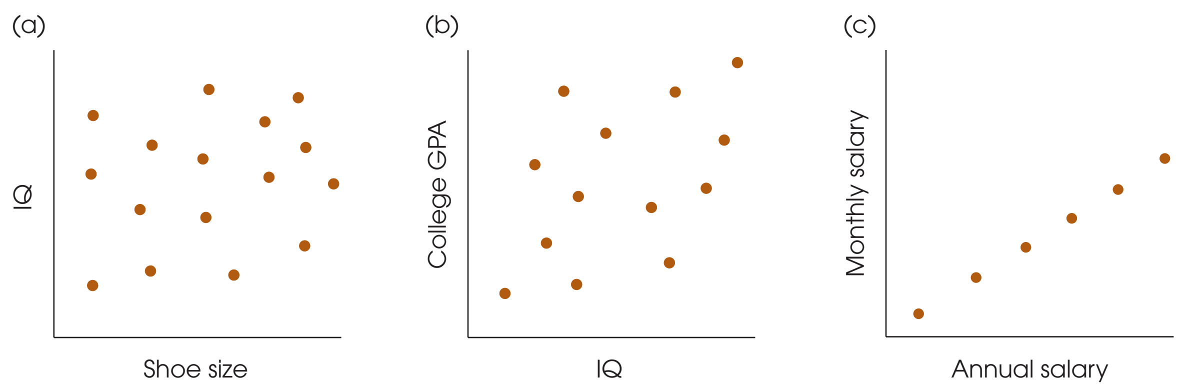


FIGURE 15.9

Three sets of data showing three different degrees of linear relationship.

second variable. One final comment concerning the interpretation of correlations is presented in Box 15.3.

BOX 15.3

REGRESSION TOWARD THE MEAN

Consider the following problem.

> Explain why the rookie of the year in major-league baseball usually does not perform as well in his second season.

Notice that this question does not appear to be statistical or mathematical in nature. However, the answer to the question is directly related to the statistical concepts of correlation and regression (Chapter 16). Specifically, there is a simple observation about correlations known as *regression toward the mean*.

DEFINITION When there is a less-than-perfect correlation between two variables, extreme scores (high or low) for one variable tend to be paired with the less extreme scores (more toward the mean) on the second variable. This fact is called **regression toward the mean.**

Figure 15.10 shows a scatter plot with a less-than-perfect correlation between two variables. The data points in this figure might represent batting averages for baseball rookies in 2010 (variable 1) and batting averages for the same players in 2011 (variable 2). Because the correlation is less than perfect, the highest scores on variable 1 are generally *not* the highest scores on variable 2. In baseball terms, the rookies who had the highest averages in 2010 do not have the highest averages in 2011.

Remember that a correlation does not explain *why* one variable is related to the other; it simply says that there is a relationship. The correlation cannot explain why the best rookie does not perform as well in his second year. But, because the correlation is not perfect, it is a statistical fact that extremely high scores in one year generally will *not* be paired with extremely high scores in the next year.

Regression toward the mean often poses a problem for interpreting experimental results. Suppose, for example, that you want to evaluate the effects of a special preschool program for disadvantaged children. You select a sample of children who score extremely low on an academic performance test. After participating in your preschool program, these children score significantly higher on the test. Why did their scores improve? One answer is that the special program helped. But an alternative answer is regression toward the mean. If there is a less-than-perfect correlation between scores on the first test and scores on the second test (which is usually the case), individuals with extremely low scores on test 1 will tend to have higher scores on test 2. It is a statistical fact of life, not necessarily the result of any special program.

Now try using the concept of regression toward the mean to explain the following phenomena:

1. You have a truly outstanding meal at a restaurant. However, when you go back with a group of friends, you find that the food is disappointing.
2. You have the highest score on exam I in your statistics class, but score only a little above average on exam II.

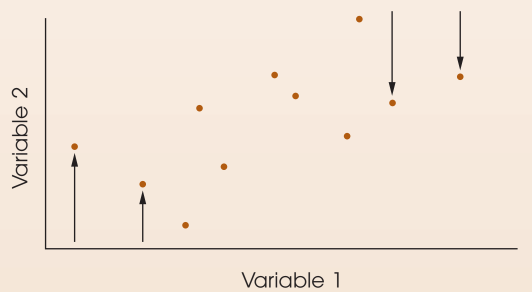


FIGURE 15.10

A demonstration of regression toward the mean. The figure shows a scatterplot for a set of data with a less-than-perfect correlation. Notice that the highest scores on variable 1 (extreme right-hand points) are not the highest scores on variable 2, but are displaced downward toward the mean. Also, the lowest scores on variable 1 (extreme left-hand points) are not the lowest scores on variable 2, but are displaced upward toward the mean.

LEARNING CHECK

1. A researcher finds a correlation of $r =$ (0.71 between the time spent playing video games each week and grade point average for a group of high school boys. This means that playing video games causes students to get lower grades. (True or false?)
2. A researcher finds a correlation of $r = 0.60$ between salary and the number of years of education for a group of 40-year-old men. How much of the variance in salary is explained by the years of education?

ANSWERS

1. False. You cannot conclude that there is a cause-and-effect relationship based on a correlation.
2. $r^2 = 0.36$, or 36%

15.4 HYPOTHESIS TESTS WITH THE PEARSON CORRELATION

The Pearson correlation is generally computed for sample data. As with most sample statistics, however, a sample correlation is often used to answer questions about the corresponding population correlation. For example, a psychologist would like to know whether there is a relationship between IQ and creativity. This is a general question concerning a population. To answer the question, a sample would be selected, and the sample data would be used to compute the correlation value. You should recognize this process as an example of inferential statistics: using samples to draw inferences about populations. In the past, we have been concerned primarily with using sample means as the basis for answering questions about population means. In this section, we examine the procedures for using a sample correlation as the basis for testing hypotheses about the corresponding population correlation.

THE HYPOTHESES

The basic question for this hypothesis test is whether a correlation exists in the population. The null hypothesis is "No. There is no correlation in the population," or "The population correlation is zero." The alternative hypothesis is "Yes. There is a real, nonzero correlation in the population." Because the population correlation is traditionally represented by ρ (the Greek letter rho), these hypotheses would be stated in symbols as

H_0: $\rho = 0$ (There is no population correlation.)

H_1: $\rho \neq 0$ (There is a real correlation.)

When there is a specific prediction about the direction of the correlation, it is possible to do a directional, or one-tailed, test. For example, if a researcher is predicting a positive relationship, the hypotheses would be

H_0: $\rho \leq 0$ (The population correlation is not positive.)

H_1: $\rho > 0$ (The population correlation is positive.)

The correlation from the sample data is used to evaluate the hypotheses. For the regular, nondirectional test, a sample correlation near zero provides support for H_0 and a sample value far from zero tends to refute H_0. For a directional test, a positive value for the sample correlation would tend to refute a null hypothesis stating that the population correlation is not positive.

Although sample correlations are used to test hypotheses about population correlations, you should keep in mind that samples are not expected to be identical to the populations from which they come; there is some discrepancy (sampling error) between a sample statistic and the corresponding population parameter. Specifically, you should always expect some error between a sample correlation and the population correlation it represents. One implication of this fact is that even when there is no correlation in the population ($\rho = 0$), you are still likely to obtain a nonzero value for the sample correlation. This is particularly true for small samples. Figure 15.11 illustrates how a small sample from a population with a near-zero correlation could result in a correlation that deviates from zero. The colored dots in the figure represent the entire population and the three circled dots represent a random sample. Note that the three sample points show a relatively good, positive correlation even through there is no linear trend ($\rho = 0$) for the population.

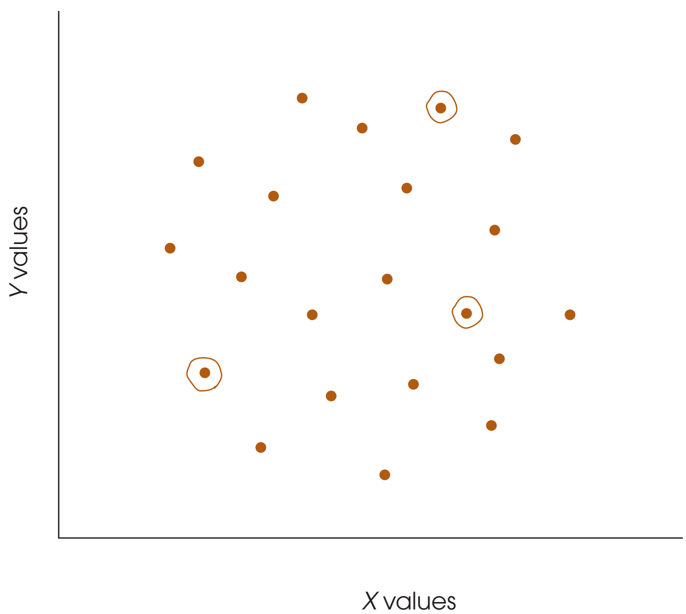
When you obtain a nonzero correlation for a sample, the purpose of the hypothesis test is to decide between the following two interpretations:

1. There is no correlation in the population ($\rho = 0$), and the sample value is the result of sampling error. Remember, a sample is not expected to be identical to the population. There always is some error between a sample statistic and the corresponding population parameter. This is the situation specified by H_0.
2. The nonzero sample correlation accurately represents a real, nonzero correlation in the population. This is the alternative stated in H_1.

The correlation from the sample helps to determine which of these two interpretations is more likely. A sample correlation near zero supports the conclusion that the population correlation is also zero. A sample correlation that is substantially different from zero supports the conclusion that there is a real, nonzero correlation in the population.

FIGURE 15.11

Scatterplot of a population of X and Y values with a near-zero correlation. However, a small sample of $n = 3$ data points from this population shows a relatively strong, positive correlation. Data points in the sample are circled.



DEGREES OF FREEDOM FOR THE CORRELATION TEST

The hypothesis test for the Pearson correlation has degrees of freedom defined by $df = n - 2$. An intuitive explanation for this value is that a sample with only $n = 2$ data points has no degrees of freedom. Specifically, if there are only two points, they will fit perfectly on a straight line, and the sample produces a perfect correlation of $r = +1.00$ or $r = -1.00$. Because the first two points always produce a perfect correlation, the sample correlation is free to vary only when the data set contains more than two points. Thus, $df = n - 2$.

THE HYPOTHESIS TEST

Although it is possible to conduct the hypothesis test by computing either a t statistic or an F-ratio, the computations for evaluating r have already been completed and are summarized in Table B.6 in Appendix B. The table is based on the concept that a sample is expected to be representative of the population from which it was obtained. In particular, a sample correlation should be similar to the population correlation. If the population correlation is zero, as specified in the null hypothesis, then the sample correlation should be near zero. Thus, a sample correlation that is close to zero provides support for H_0 and a sample correlation that is far from zero contradicts the null hypothesis.

The table lists critical values in terms of degrees of freedom: $df = n - 2$. Remember to subtract 2 when using this table.

Table B.6 identifies exactly which sample correlations are likely to be obtained from a population with $\rho = 0$ and which values are very unlikely. To use the table, you need to know the sample size (n) and the alpha level. With a sample size of $n = 20$ and an alpha level of .05, for example, you locate $df = n - 2 = 18$ in the left-hand column and the value .05 for either one tail or two tails across the top of the table. For $df = 18$ and $\alpha = .05$ for a two-tailed test, the table shows a value of 0.444. Thus, if the null hypothesis is true and there is no correlation in the population, then the sample correlation should be near to zero. According to the table, the sample correlation should have a value between +0.444 and –0.444. If H_0 is true, it is very unlikely ($\alpha = .05$) to obtain a sample correlation outside this range. Therefore, a sample correlation beyond ±0.444 leads to rejecting the null hypothesis. The following examples demonstrate the use of the table.

EXAMPLE 15.6

A researcher is using a regular, two-tailed test with $\alpha = .05$ to determine whether a nonzero correlation exists in the population. A sample of $n = 30$ individuals is obtained. With $\alpha = .05$ and $n = 30$, the table lists a value of 0.361. Thus, the sample correlation (independent of sign) must have a value greater than or equal to 0.361 to reject H_0 and conclude that there is a significant correlation in the population. Any sample correlation between 0.361 and –0.361 is considered within the realm of sampling error and, therefore, is not significant.

EXAMPLE 15.7

This time the researcher is using a directional, one-tailed test to determine whether there is a positive correlation in the population.

H_0: $\rho \leq 0$ (There is not a positive correlation.)

H_1: $\rho > 0$ (There is a positive correlation.)

With $\alpha = .05$ and a sample of $n = 30$, the table lists a value of 0.306 for a one-tailed test. To reject H_0 and conclude that there is a significant positive correlation in the population, the sample correlation must be positive (as predicted) and have a value greater than or equal to 0.306.

IN THE LITERATURE

REPORTING CORRELATIONS

When correlations are computed, the results are reported using APA format. The statement should include the sample size, the calculated value for the correlation, whether it is a statistically significant relationship, the probability level, and the type of test used (one- or two-tailed). For example, a correlation might be reported as follows:

> A correlation for the data revealed a significant relationship between amount of education and annual income, $r = +0.65$, $n = 30$, $p < .01$, two tails.

Sometimes a study might look at several variables, and correlations between all possible variable pairings are computed. Suppose, for example, that a study measured people's annual income, amount of education, age, and intelligence. With four variables, there are six possible pairings leading to six different correlations. The results from multiple correlations are most easily reported in a table called a *correlation matrix,* using footnotes to indicate which correlations are significant. For example, the report might state:

The analysis examined the relationships among income, amount of education, age, and intelligence for $n = 30$ participants. The correlations between pairs of variables are reported in Table 1. Significant correlations are noted in the table.

TABLE 1

Correlation matrix for income, amount of education, age, and intelligence

	Education	Age	IQ
Income	+.65*	+.41**	+.27
Education		+.11	+.38**
Age			−.02

$n = 30$
*$p < .01$, two tails
**$p < .05$, two tails

LEARNING CHECK

1. A researcher obtains a correlation of $r = -0.39$ for a sample of $n = 25$ individuals. Does this sample provide sufficient evidence to conclude that there is a significant, nonzero correlation in the population? Assume a two-tailed test with $\alpha = .05$.
2. For a sample of $n = 15$, how large a correlation is needed to conclude at the .05 level of significance that there is a nonzero correlation in the population? Assume a two-tailed test.
3. As sample size gets smaller, what happens to the magnitude of the correlation necessary for significance? Explain why this occurs.

ANSWERS

1. No. For $n = 25$, the critical value is $r = 0.396$. The sample value is not in the critical region.
2. For $n = 15$, $df = 13$ and the critical value is $r = 0.514$.
3. As the sample size gets smaller, the magnitude of the correlation needed for significance gets larger. With a small sample, it is easy to get a relatively large correlation just by chance. Therefore, a small sample requires a very large correlation before you can be confident there is a real (nonzero) relationship in the population.

PARTIAL CORRELATIONS

Occasionally a researcher may suspect that the relationship between two variables is being distorted by the influence of a third variable. Earlier in the chapter, for example, we found a strong positive relationship between the number of churches and the number of serious crimes for a sample of different towns and cities (see Example 15.4, p 522). However, it is unlikely that there is a direct relationship between churches and crime. Instead, both variables are influenced by population: Large cities have a lot of churches and high crime rates compared to smaller towns, which have fewer churches and less crime. If population were controlled, there probably would be no real correlation between churches and crime.

Fortunately, there is a statistical technique, known as *partial correlation,* that allows a researcher to measure the relationship between two variables while eliminating or holding constant the influence of a third variable. Thus, a researcher could use a partial correlation to examine the relationship between churches and crime without the risk that the relationship is distorted by the size of the population.

DEFINITION

A **partial correlation** measures the relationship between two variables while controlling the influence of a third variable by holding it constant.

In a situation with three variables, X, Y, and Z, it is possible to compute three individual Pearson correlations:

1. r_{XY} measuring the correlation between X and Y
2. r_{XZ} measuring the correlation between X and Z
3. r_{YZ} measuring the correlation between Y and Z

These three individual correlations can then be used to compute a partial correlation. For example, the partial correlation between X and Y, holding Z constant, is determined by the formula

$$r_{XY-Z} = \frac{r_{XY} - (r_{XZ} r_{YZ})}{\sqrt{(1 - r_{XZ}^2)(1 - r_{YZ}^2)}} \qquad (15.6)$$

The following example demonstrates the calculation and interpretation of a partial correlation.

EXAMPLE 15.8

We begin with the hypothetical data shown in Table 15.2. These scores have been constructed to simulate the church/crime/population situation for a sample of $n = 15$ cities. The X variable represents the number of churches, Y represents the number of

TABLE 15.2

Hypothetical data showing the relationship between the number of churches, the number of crimes, and the population of a set of $n = 15$ cities.

Number of Churches (X)	Number of Crimes (Y)	Population (Z)
1	4	1
2	3	1
3	1	1
4	2	1
5	5	1
7	8	2
8	11	2
9	9	2
10	7	2
11	10	2
13	15	3
14	14	3
15	16	3
16	17	3
17	13	3

crimes, and Z represents the population for each city. For these scores, the individual Pearson correlations are all large and positive:

a. The correlation between churches and crime is $r_{XY} = 0.923$.

b. The correlation between churches and population is $r_{XZ} = 0.961$.

c. The correlation between crime and population is $r_{YZ} = 0.961$.

The data points for the 15 cities are shown in the scatter plot in Figure 15.12. Note that there are three categories for the size of the population (three values for Z) corresponding to small, medium, and large cities. Also note that the population variable, Z, separates the scores into three distinct groups: When $Z = 1$, the population is low and churches and crime (X and Y) are also low; when $Z = 2$, the population is moderate and churches and crime (X and Y) are also moderate; and when $Z = 3$, the population is large and churches and crime are both high. Thus, as the population increases from one city to another, the number of churches and crimes also increase, and the result is a strong positive correlation between churches and crime.

Within each of the three population categories, however, there is no linear relationship between churches and crime. Specifically, within each group, the population variable is constant and the five data points for X and Y form a circular pattern, indicating no consistent linear relationship. The partial correlation allows us to hold population constant across the entire sample and measure the underlying relationship between churches and crime without any influence from population. For these data, the partial correlation is

$$r_{XY-Z} = \frac{0.923 - 0.961(0.961)}{\sqrt{(1-0.961^2)(1-0.961^2)}}$$

$$= \frac{0}{0.076}$$

$$= 0$$

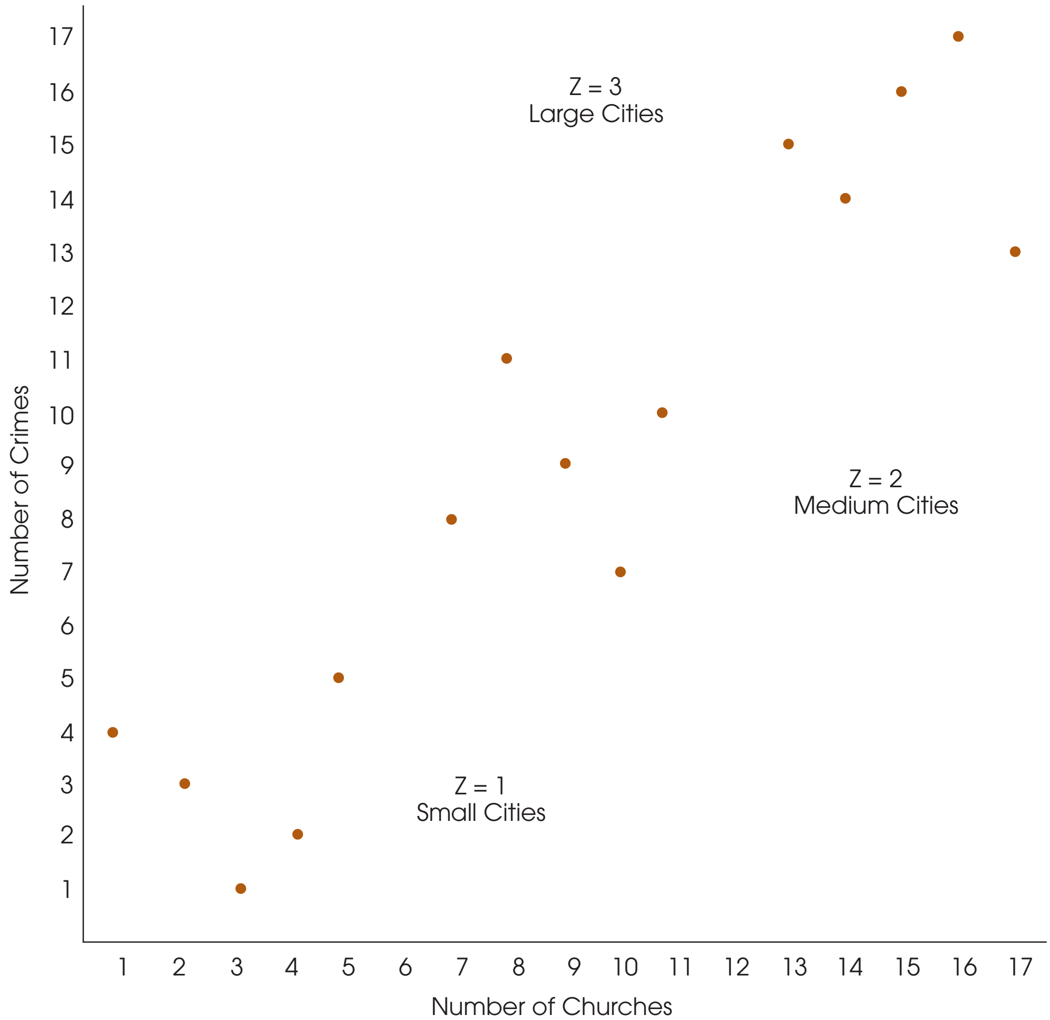


FIGURE 15.12

Hypothetical data showing the relationship between the number of churches and the number of crimes for three groups of cities. Those with small populations ($Z = 1$), those with medium populations ($Z = 2$), and those with large populations ($Z = 3$).

Thus, when the population differences are eliminated, there is no correlation remaining between churches and crime ($r = 0$).

In Example 15.8, the population differences, which correspond to the different values of the Z variable, were eliminated mathematically in the calculation of the partial correlation. However, it is possible to visualize how these differences are eliminated in the actual data. Looking at Figure 15.12, focus on the five points in the bottom left corner. These are the five cities with small populations, few churches, and little crime. The five points in the upper right corner represent the five cities with large populations, many churches, and a lot of crime. The partial correlation controls population size by mathematically equalizing the populations for all 15 cities. Population is increased for the five small cities. However, increasing the population also increases churches and crime. Similarly, population is decreased for the five large cities, which also decreases churches and crime. In Figure 15.12, imagine the five points in the bottom left moving up and to the right so that they overlap with the points in the center. At the same time, the five points in the upper right move down and to the left so that they also overlap the points in the center. When population is equalized, the resulting set of 15 cities is shown in Figure 15.13. Note that controlling the population appears to have eliminated the relationship between churches and crime. This appearance is verified by the correlation for the 15 data points in Figure 15.13, which is $r = 0$, exactly the same as the partial correlation.

In Example 15.8 we used a partial correlation to demonstrate that an apparent relationship between churches and crime was actually caused by the influence of a third variable, population. It also is possible to use partial correlations to demonstrate that a relationship is not caused by the influence of a third variable. As an example, consider research examining the relationship between exposure to sexual content on television and sexual behavior among adolescents (Collins et al., 2004). The study consisted of a survey of 1,792 adolescents, 12 to 17 years old, who reported their television viewing

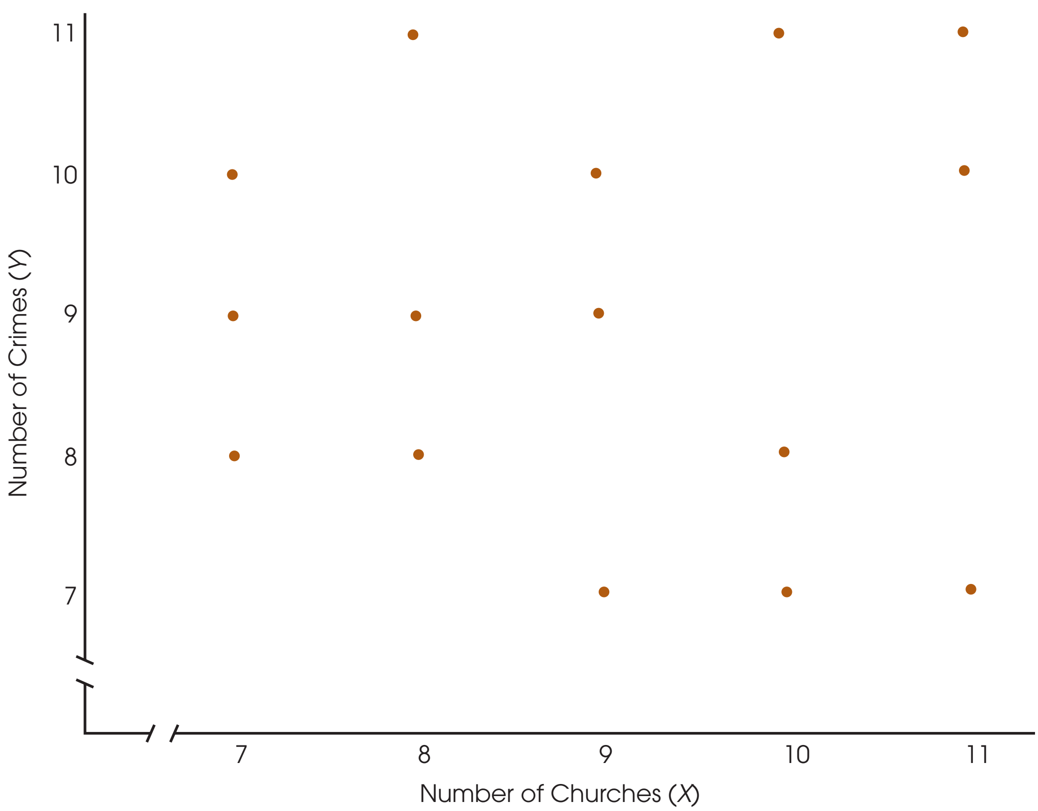


FIGURE 15.13

The relationship between the number of churches and the number of crimes for the same 15 cities shown in Figure 15.12 after the populations have been equalized.

habits and their sexual behaviors. The results showed a clear relationship between television viewing and behaviors. Specifically, the more sexual content the adolescents watched on television, the more likely they were to engage in sexual behaviors. One concern for the researchers was that the observed relationship may be influenced by the age of the participants. For example, as the adolescents mature from age 12 to age 17, they increasingly watch television programs with sexual content and they increase their own sexual behaviors. Although the viewing of sexual content on television and the participants' sexual behaviors are increasing together, the observed relationship may simply be the result of age differences. To address this problem, the researcher used a partial correlation technique to eliminate or hold constant the age variable. The results clearly showed that a relationship still exists between television sexual content and sexual behavior even after the influence of the participants' ages was accounted for.

Testing the significance of a partial correlation The statistical significance of a partial correlation is determined using the same procedure as is used to evaluate a regular Pearson correlation. Specifically, the partial correlation is compared with the critical values listed in Table B6. For a partial correlation, however, you must use $df = n - 3$ instead of the $n - 2$ value that is used for the Pearson correlation. A significant correlation means that it is very unlikely ($p < \alpha$) that the sample correlation would occur without a corresponding relationship in the population.

LEARNING CHECK

1. Sales figures show a positive relationship between temperature and ice cream consumption; as temperature increases, ice cream consumption also increases. Other research shows a positive relationship between temperature and crime rate (Cohn & Rotton, 2000). When the temperature increases, both ice cream consumption and crime rates tend to increase. As a result, there is a positive correlation between ice cream consumption and crime rate. However, what do you think is the true relationship between ice cream consumption and crime rate? Specifically, what value would you predict for the partial correlation between the two variables if temperature were held constant?

ANSWER

1. There should be no systematic relationship between ice cream consumption and crime rate. The partial correlation should be near zero.

15.5 ALTERNATIVES TO THE PEARSON CORRELATION

The Pearson correlation measures the degree of linear relationship between two variables when the data (X and Y values) consist of numerical scores from an interval or ratio scale of measurement. However, other correlations have been developed for nonlinear relationships and for other types of data. In this section we examine three additional correlations: the Spearman correlation, the point-biserial correlation, and the phi-coefficient. As you will see, all three can be viewed as special applications of the Pearson correlation.

THE SPEARMAN CORRELATION

When the Pearson correlation formula is used with data from an ordinal scale (ranks), the result is called the *Spearman correlation*. The Spearman correlation is used in two situations.

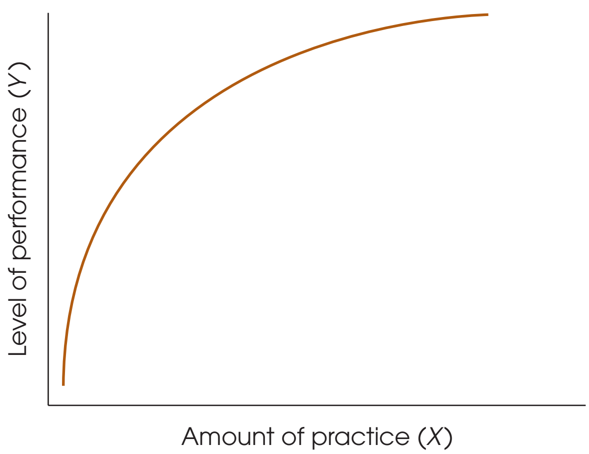
First, the Spearman correlation is used to measure the relationship between X and Y when both variables are measured on ordinal scales. Recall from Chapter 1 that an ordinal scale typically involves ranking individuals rather than obtaining numerical scores. Rank-order data are fairly common because they are often easier to obtain than interval or ratio scale data. For example, a teacher may feel confident about rank-ordering students' leadership abilities but would find it difficult to measure leadership on some other scale.

In addition to measuring relationships for ordinal data, the Spearman correlation can be used as a valuable alternative to the Pearson correlation, even when the original raw scores are on an interval or a ratio scale. As we have noted, the Pearson correlation measures the degree of *linear relationship* between two variables—that is, how well the data points fit on a straight line. However, a researcher often expects the data to show a consistently one-directional relationship but not necessarily a linear relationship. For example, Figure 15.14 shows the typical relationship between practice and performance. For nearly any skill, increasing amounts of practice tend to be associated with improvements in performance (the more you practice, the better you get). However, it is not a straight-line relationship. When you are first learning a new skill, practice produces large improvements in performance. After you have been performing a skill for several years, however, additional practice produces only minor changes in performance. Although there is a consistent relationship between the amount of practice and the quality of performance, it clearly is not linear. If the Pearson correlation were computed for these data, it would not produce a correlation of 1.00 because the data do not fit perfectly on a straight line. In a situation like this, the Spearman correlation can be used to measure the consistency of the relationship, independent of its form.

The reason that the Spearman correlation measures consistency, rather than form, comes from a simple observation: When two variables are consistently related, their ranks are linearly related. For example, a perfectly consistent positive relationship means that every time the X variable increases, the Y variable also increases. Thus, the smallest value of X is paired with the smallest value of Y, the second-smallest value of X is paired with the second smallest value of Y, and so on. Every time the rank for X goes up by 1 point, the rank for Y also goes up by 1 point. As a result, the ranks fit perfectly on a straight line. This phenomenon is demonstrated in the following example.

FIGURE 15.14

Hypothetical data showing the relationship between practice and performance. Although this relationship is not linear, there is a consistent positive relationship. An increase in performance tends to accompany an increase in practice.



EXAMPLE 15.9

Table 15.3 presents X and Y scores for a sample of $n = 4$ people. Note that the data show a perfectly consistent relationship. Each increase in X is accompanied by an increase in Y. However the relationship is not linear, as can be seen in the graph of the data in Figure 15.15(a).

Next, we convert the scores to ranks. The lowest X is assigned a rank of 1, the next lowest a rank of 2, and so on. The Y scores are then ranked in the same way. The ranks are listed in Table 15.3 and shown in Figure 15.15(b). Note that the perfect consistency for the scores produces a perfect linear relationship for the ranks.

The preceding example demonstrates that a consistent relationship among scores produces a linear relationship when the scores are converted to ranks. Thus, if you want to measure the consistency of a relationship for a set of scores, you can simply convert the scores to ranks and then use the Pearson correlation formula to measure the linear relationship for the ranked data. The degree of linear relationship for the ranks provides a measure of the degree of consistency for the original scores.

TABLE 15.3

Scores and ranks for Example 15.9.

Person	X	Y	X−Rank	Y−Rank
A	4	9	3	3
B	2	2	1	1
C	10	10	4	4
D	3	8	2	2

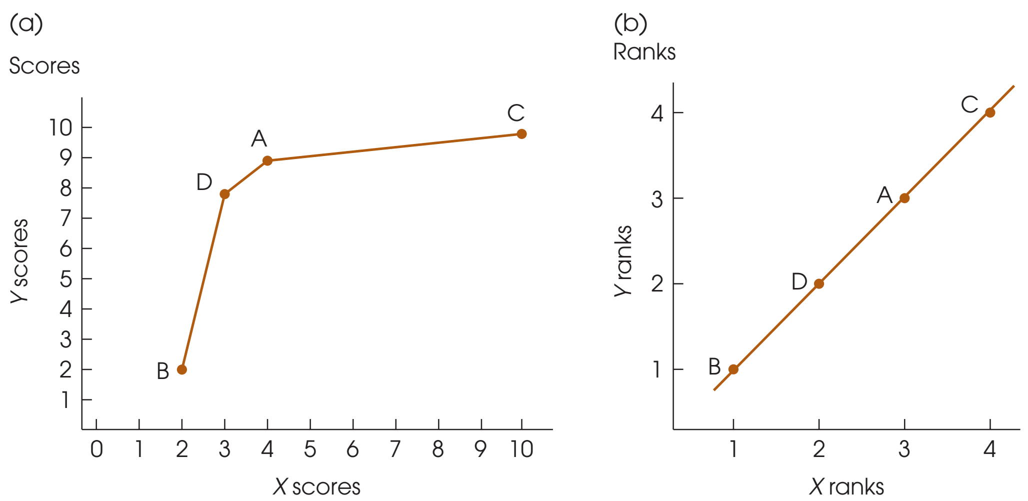


FIGURE 15.15

Scatter plots showing (a) the scores and (b) the ranks for the data in Example 15.9. Notice that there is a consistent, positive relationship between the X and Y scores, although it is not a linear relationship. Also notice that the scatter plot of the ranks shows a perfect linear relationship.

To summarize, the Spearman correlation measures the relationship between two variables when both are measured on ordinal scales (ranks). There are two general situations in which the Spearman correlation is used:

1. Spearman is used when the original data are ordinal; that is, when the X and Y values are ranks. In this case, you simply apply the Pearson correlation formula to the set of ranks.
2. Spearman is used when a researcher wants to measure the consistency of a relationship between X and Y, independent of the specific form of the relationship. In this case, the original scores are first converted to ranks, then the Pearson correlation formula is used with the ranks. Because the Pearson formula measures the degree to which the ranks fit on a straight line, it also measures the degree of consistency in the relationship for the original scores. Incidentally, when there is a consistently one-directional relationship between two variables, the relationship is said to be *monotonic*. Thus, the Spearman correlation measures the degree of monotonic relationship between two variables.

The word *monotonic* describes a sequence that is consistently increasing (or decreasing). Like the word *monotonous,* it means constant and unchanging.

In either case, the Spearman correlation is identified by the symbol r_S to differentiate it from the Pearson correlation. The complete process of computing the Spearman correlation, including ranking scores, is demonstrated in Example 15.10.

EXAMPLE 15.10

The following data show a nearly perfect monotonic relationship between X and Y. When X increases, Y tends to decrease, and there is only one reversal in this general trend. To compute the Spearman correlation, we first rank the X and Y values, and we then compute the Pearson correlation for the ranks.

We have listed the X values in order so that the trend is easier to recognize.

Original Data

X	Y
3	12
4	10
10	11
11	9
12	2

Ranks

X	Y	XY
1	5	5
2	3	6
3	4	12
4	2	8
5	1	5
		$36 = \Sigma XY$

To compute the correlation, we need SS for X, SS for Y, and SP. Remember that all of these values are computed with the ranks, not the original scores. The X ranks are simply the integers 1, 2, 3, 4, and 5. These values have $\Sigma X = 15$ and $\Sigma X^2 = 55$. The SS for the X ranks is

$$SS_X = \Sigma X^2 - \frac{(\Sigma X)^2}{n} = 55 - \frac{(15)^2}{5} = 10$$

Note that the ranks for Y are identical to the ranks for X; that is, they are the integers 1, 2, 3, 4, and 5. Therefore, the SS for Y is identical to the SS for X:

$$SS_Y = 10$$

To compute the *SP* value, we need ΣX, ΣY, and ΣXY for the ranks. The *XY* values are listed in the table with the ranks, and we already have found that both the *X*s and the *Y*s have a sum of 15. Using these values, we obtain

$$SP = \Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{n} = 36 - \frac{(15)(15)}{5} = -9$$

Finally, the Spearman correlation simply uses the Pearson formula for the ranks.

$$r_s = \frac{SP}{\sqrt{(SS_X)(SS_Y)}} = \frac{-9}{\sqrt{10(10)}} = -0.9$$

The Spearman correlation indicates that the data show a consistent (nearly perfect) negative trend.

RANKING TIED SCORES

When you are converting scores into ranks for the Spearman correlation, you may encounter two (or more) identical scores. Whenever two scores have exactly the same value, their ranks should also be the same. This is accomplished by the following procedure:

1. List the scores in order from smallest to largest. Include tied values in the list.
2. Assign a rank (first, second, etc.) to each position in the ordered list.
3. When two (or more) scores are tied, compute the mean of their ranked positions, and assign this mean value as the final rank for each score.

The process of finding ranks for tied scores is demonstrated here. These scores have been listed in order from smallest to largest.

Scores	Rank Position	Final Rank	
3	1	1.5	Mean of 1 and 2
3	2	1.5	
5	3	3	
6	4	5	Mean of 4, 5, and 6
6	5	5	
6	6	5	
12	7	7	

Note that this example has seven scores and uses all seven ranks. For $X = 12$, the largest score, the appropriate rank is 7. It cannot be given a rank of 6 because that rank has been used for the tied scores.

SPECIAL FORMULA FOR THE SPEARMAN CORRELATION

After the original *X* values and *Y* values have been ranked, the calculations necessary for *SS* and *SP* can be greatly simplified. First, you should note that the *X* ranks and the *Y* ranks are really just a set of integers: 1, 2, 3, 4, ... , *n*. To compute the mean for these integers, you can locate the midpoint of the series by $M = (n + 1)/2$. Similarly, the SS for this series of integers can be computed by

$$SS = \frac{n(n^2 - 1)}{12} \quad \text{(Try it out.)}$$

Also, because the X ranks and the Y ranks are the same values, the SS for X is identical to the SS for Y.

Because calculations with ranks can be simplified and because the Spearman correlation uses ranked data, these simplifications can be incorporated into the final calculations for the Spearman correlation. Instead of using the Pearson formula after ranking the data, you can put the ranks directly into a simplified formula:

$$r_S = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)} \tag{15.7}$$

Caution: In this formula, you compute the value of the fraction and then subtract from 1. The 1 is not part of the fraction.

where D is the difference between the X rank and the Y rank for each individual. This special formula produces the same result that would be obtained from the Pearson formula. However, note that this special formula can be used only after the scores have been converted to ranks and only when there are no ties among the ranks. If there are relatively few tied ranks, the formula still may be used, but it loses accuracy as the number of ties increases. The application of this formula is demonstrated in the following example.

EXAMPLE 15.11

To demonstrate the special formula for the Spearman correlation, we use the same data that were presented in Example 15.10. The ranks for these data are shown again here:

Ranks		Difference	
X	Y	D	D^2
1	5	4	16
2	3	1	1
3	4	1	1
4	2	−2	4
5	1	−4	16
			$38 = \Sigma D^2$

Using the special formula for the Spearman correlation, we obtain

$$r_S = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)} = 1 - \frac{6(38)}{5(25 - 1)} = 1 - \frac{228}{120} = 1 - 1.90 = -0.90$$

This is exactly the same answer that we obtained in Example 15.10, using the Pearson formula on the ranks.

TESTING THE SIGNIFICANCE OF THE SPEARMAN CORRELATION

Testing a hypothesis for the Spearman correlation is similar to the procedure used for the Pearson r. The basic question is whether a correlation exists in the population. The sample correlation could be the result of chance, or perhaps it reflects an actual relationship between the variables in the population. For the Pearson correlation, the Greek letter rho (ρ) was used for the population correlation. For the Spearman, ρ_S is used for the population parameter. Note that this symbol is consistent with the sample statistic, r_S. The null hypothesis states that there is no correlation (no monotonic relationship) between the variables for the population, or, in symbols:

$$H_0: \quad \rho_S = 0 \qquad \text{(The population correlation is zero.)}$$

The alternative hypothesis predicts that a nonzero correlation exists in the population, which can be stated in symbols as

$$H_1: \quad \rho_S \neq 0 \qquad \text{(There is a real correlation.)}$$

To determine whether the Spearman correlation is statistically significant (that is, H_0 should be rejected), consult Table B.7. This table is similar to the one used to determine the significance of Pearson's r (Table B.6); however, the first column is sample size (n) rather than degrees of freedom. To use the table, line up the sample size in the first column with one of the alpha levels listed across the top. The values in the body of the table identify the magnitude of the Spearman correlation that is necessary to be significant. The table is built on the concept that a sample correlation should be representative of the corresponding population value. In particular, if the population correlation is $\rho_S = 0$ (as specified in H_0), then the sample correlation should be near zero. For each sample size and alpha level, the table identifies the minimum sample correlation that is significantly different from zero. The following example demonstrates the use of the table.

EXAMPLE 15.12

An industrial psychologist selects a sample of $n = 15$ employees. These employees are ranked in order of work productivity by their manager. They also are ranked by a peer. The Spearman correlation computed for these data revealed a correlation of $r_S = .60$. Using Table B.7 with $n = 15$ and $\alpha = .05$, a correlation of at least ± 0.45 is needed to reject H_0. The observed correlation for the sample easily surpasses this critical value. The correlation between manager and peer ratings is statistically significant.

LEARNING CHECK

1. Describe what is measured by a Spearman correlation, and explain how this correlation is different from the Pearson correlation.
2. If the following scores are converted into ranks, what rank will be assigned to the individuals who have scores of $X = 7$?

   Scores: 1, 1, 1, 3, 6, 7, 7, 8, 10
3. Rank the following scores and compute the Spearman correlation:

X	Y
2	7
12	38
9	6
10	19

ANSWERS

1. The Spearman correlation measures the consistency of the direction of the relationship between two variables. The Spearman correlation does not depend on the form of the relationship, whereas the Pearson correlation measures how well the data fit a linear form.
2. Both scores get a rank of 6.5 (the average of 6 and 7).
3. $r_S = 0.80$

THE POINT-BISERIAL CORRELATION AND MEASURING EFFECT SIZE WITH r^2

In Chapters 9, 10, and 11 we introduced r^2 as a measure of effect size that often accompanies a hypothesis test using the t statistic. The r^2 used to measure effect size and the r used to measure a correlation are directly related, and we now have an opportunity to demonstrate the relationship. Specifically, we compare the independent-measures t test (Chapter 10) and a special version of the Pearson correlation known as the *point-biserial correlation.*

The point-biserial correlation is used to measure the relationship between two variables in situations in which one variable consists of regular, numerical scores, but the second variable has only two values. A variable with only two values is called a *dichotomous variable* or a *binomial variable*. Some examples of dichotomous variables are

1. Male versus female
2. College graduate versus not a college graduate
3. First-born child versus later-born child
4. Success versus failure on a particular task
5. Older than 30 years versus younger than 30 years

It is customary to use the numerical values 0 and 1, but any two different numbers would work equally well and would not affect the value of the correlation.

To compute the point-biserial correlation, the dichotomous variable is first converted to numerical values by assigning a value of zero (0) to one category and a value of one (1) to the other category. Then the regular Pearson correlation formula is used with the converted data.

To demonstrate the point-biserial correlation and its association with the r^2 measure of effect size, we use the data from Example 10.1 (p. 326). The original example compared high school grades for two groups of students: one group who regularly watched Sesame Street as 5-year-old children and one who did not watch the program. The data from the independent-measures study are presented on the left side of Table 15.4. Notice that the data consist of two separate samples and the independent-measures t was used to determine whether there was a significant mean difference between the two populations represented by the samples.

On the right-hand side of Table 15.4, we have reorganized the data into a form that is suitable for a point-biserial correlation. Specifically, we used each student's high school grade as the X value and we have created a new variable, Y, to represent the group, or condition, for each student. In this case, we have used $Y = 1$ for students who watched Sesame Street and $Y = 0$ for students who did not watch the program.

When the data in Table 15.4 were originally presented in Chapter 10, we conducted an independent-measures t hypothesis test and obtained $t = 4.00$ with $df = 18$. We measured the size of the treatment effect by calculating r^2, the percentage of variance accounted for, and obtained $r^2 = 0.47$.

Calculating the point-biserial correlation for these data also produces a value for r. Specifically, the X scores produce $SS = 680$; the Y values produce $SS = 5.00$, and the

TABLE 15.4

The same data are organized in two different formats. On the left-hand side, the data appear as two separate samples appropriate for an independent-measures *t* hypothesis test. On the right-hand side, the same data are shown as a single sample, with two scores for each individual: the original high school grade and a dichotomous score (Y) that identifies the group in which the participant is located (Seasame Street = 1 and No-Sesame Street = 0). The data on the right are appropriate for a point-biserial correlation.

Data for the Independent-Measures *t* test. Two separate samples, each with n = 10 scores.

Average High School Grade			
Watched Seasame Street		Did Not Watch Sesame Street	
86	99	90	79
87	97	89	83
91	94	82	86
97	89	83	81
98	92	85	92
$n = 10$		$n = 10$	
$M = 93$		$M = 85$	
$SS = 200$		$SS = 160$	

Data for the Point-Biserial Correlation. Two scores, X and Y for each of the n = 20 participants.

Participant	Grade X	Group Y
A	86	1
B	87	1
C	91	1
D	97	1
E	98	1
F	99	1
G	97	1
H	94	1
I	89	1
J	92	1
K	90	0
L	89	0
M	82	0
N	83	0
O	85	0
P	79	0
Q	83	0
R	86	0
S	81	0
T	92	0

sum of the products of the X and Y deviations produces $SP = 40$. The point-biserial correlation is

$$r = \frac{SP}{\sqrt{(SS_X)(SS_Y)}} = \frac{40}{\sqrt{(680)(5)}} = \frac{40}{58.31} = 0.686$$

Notice that squaring the value of the point-biserial correlation produces $r^2 = (0.686)^2 = 0.47$, which is exactly the value of r^2 we obtained measuring effect size.

In some respects, the point-biserial correlation and the independent-measures hypothesis test are evaluating the same thing. Specifically, both are examining the relationship between the TV-viewing habits of 5-year-old children and their future academic performance in high school.

1. The correlation is measuring the *strength* of the relationship between the two variables. A large correlation (near 1.00 or –1.00) would indicate that there is a consistent, predictable relationship between high school grades and watching Sesame Street as a 5-year-old child. In particular, the value of r^2 measures how much of the variability in grades can be predicted by knowing whether the participants watched Sesame Street.
2. The *t* test evaluates the *significance* of the relationship. The hypothesis test determines whether the mean difference in grades between the two groups is greater than can be reasonably explained by chance alone.

As we noted in Chapter 10 (pp. 332–333), the outcome of the hypothesis test and the value of r^2 are often reported together. The t value measures statistical significance and r^2 measures the effect size. Also, as we noted in Chapter 10, the values for t and r^2 are directly related. In fact, either can be calculated from the other by the equations

$$r^2 = \frac{t^2}{t^2 + df} \quad \text{and} \quad t^2 = \frac{r^2}{(1 - r^2)/df}$$

where df is the degrees of freedom for the t statistic.

However, you should note that r^2 is determined entirely by the size of the correlation, whereas t is influenced by the size of the correlation and the size of the sample. For example, a correlation of $r = 0.30$ produces $r^2 = 0.09$ (9%) no matter how large the sample may be. On the other hand, a point-biserial correlation of $r = 0.30$ for a total sample of 10 people ($n = 5$ in each group) produces a nonsignificant value of $t = 0.889$. If the sample is increased to 50 people ($n = 25$ in each group), the same correlation produces a significant t value of $t = 2.18$. Although t and r are related, they are measuring different things.

POINT-BISERIAL CORRELATION, PARTIAL CORRELATION, AND EFFECT SIZE FOR THE REPEATED-MEASURES *t* TEST

In the previous section we demonstrated that the point-biserial correlation produces an r value that is directly related to the r^2 value used to measure effect size for the independent-measures t test. With one modification, this same process can be duplicated for the repeated-measures t test. The modification involves using a partial correlation (see pp. 531–535) to control for individual differences.

You should recall from Chapters 11 and 13 that one of the major distinctions between independent-measures and repeated-measures designs is that the repeated-measures designs eliminate the influence of individual differences. When computing a point-biserial correlation for repeated-measures data, we can use a partial correlation to eliminate individual differences once again.

The left-hand side of Table 15.5 shows data from a repeated-measures study comparing two treatments with a sample of $n = 4$ participants. Note that we have added a column of P values, or participant totals, showing the sum of the two scores for each participant. For example, participant A has scores of 3 and 5, which add to $P = 8$. The P values provide an indication of the individual differences. Participant A, for example, has consistently smaller scores and a smaller P value than all of the other participants. These data produce $t = 2.00$ with $df = 3$, which results in $r^2 = 4/(4 + 3) = 0.5714$ as the measure of effect size.

TABLE 15.5

The data on the left represent scores from a repeated-measures study comparing two treatments with a sample of $n = 4$ participants. The data on the right are the same scores in a format compatible with the point-biserial correlation. The P values in each set of data show the sum of the two scores for each participant and provide a measure of individual differences.

	Treatment		
Participant	I	II	P
A	3	5	8
B	4	14	18
C	5	7	12
D	4	6	10

Score (X)	Treatment (Y)	P
3	0	8
4	0	18
5	0	12
4	0	10
5	1	8
14	1	18
7	1	12
6	1	10

On the right-hand side of Table 15.5 we have reorganized the data into a format compatible with the point-biserial correlation. The individual scores, or X values are listed in the first column. The second column, or Y values, are numerical codes corresponding to the two treatment conditions: Treatment I = 0 and Treatment II = 1. The third column contains the P value for each individual, which measures the individual differences between participants. For these data, the partial correlation between X and Y, controlling for the P values, is

$$r_{XY\text{-}P} = 0.756$$

Note that this is a slightly modified point-biserial correlation. The modification is that we used a partial correlation to control the individual differences. However, squaring this correlation produces $r^2 = (0.756)^2 = 0.5715$, which is identical, within rounding error, to the r^2 value that measures effect size for the repeated-measures t test.

THE PHI-COEFFICIENT

When both variables (X and Y) measured for each individual are dichotomous, the correlation between the two variables is called the *phi-coefficient.* To compute phi (ϕ), you follow a two-step procedure:

1. Convert each of the dichotomous variables to numerical values by assigning a 0 to one category and a 1 to the other category for each of the variables.
2. Use the regular Pearson formula with the converted scores.

This process is demonstrated in the following example.

EXAMPLE 15.13

A researcher is interested in examining the relationship between birth-order position and personality. A random sample of $n = 8$ individuals is obtained, and each individual is classified in terms of birth-order position as first-born or only child versus later-born. Then each individual's personality is classified as either introvert or extrovert.

The original measurements are then converted to numerical values by the following assignments:

Birth Order	Personality
1st or only child = 0	Introvert = 0
Later-born child = 1	Extrovert = 1

The original data and the converted scores are as follows:

Original Data		Converted Scores	
Birth Order (X)	Personality (Y)	Birth Order (X)	Personality (Y)
1st	Introvert	0	0
3rd	Extrovert	1	1
Only	Extrovert	0	1
2nd	Extrovert	1	1
4th	Extrovert	1	1
2nd	Introvert	1	0
Only	Introvert	0	0
3rd	Extrovert	1	1

The Pearson correlation formula is then used with the converted data to compute the phi-coefficient.

Because the assignment of numerical values is arbitrary (either category could be designated 0 or 1), the sign of the resulting correlation is meaningless. As with most correlations, the *strength* of the relationship is best described by the value of r^2, the coefficient of determination, which measures how much of the variability in one variable is predicted or determined by the association with the second variable.

We also should note that although the phi-coefficient can be used to assess the relationship between two dichotomous variables, the more common statistical procedure is a chi-square statistic, which is examined in Chapter 17.

LEARNING CHECK

1. Define a *dichotomous* variable.

2. The following data represent job-related stress scores for a sample of $n = 8$ individuals. These people also are classified by salary level.

a. Convert the data into a form suitable for the point-biserial correlation.

b. Compute the point-biserial correlation for these data.

Salary More than \$40,000	Salary Less than \$40,000
8	4
6	2
5	1
3	3

3. A researcher would like to know whether there is a relationship between gender and manual dexterity for 3-year-old children. A sample of $n = 10$ boys and $n = 10$ girls is obtained and each child is given a manual-dexterity test. Five of the girls failed the test and only two of the boys failed. Describe how these data could be coded into a form suitable for computing a phi-coefficient to measure the strength of the relationship.

ANSWERS

1. A dichotomous variable has only two possible values.

2. a. Salary level is a dichotomous variable and can be coded as $Y = 1$ for individuals with salary more than \$40,000 and $Y = 0$ for salary less than \$40,000. The stress scores produce $SS_X = 36$, the salary codes produce $SS_Y = 2$, and $SP = 6$.

b. The point-biserial correlation is 0.71.

3. Gender could be coded with male = 0 and female = 1. Manual dexterity could be coded with failure = 0 and success = 1. Eight boys would have scores of 0 and 1 and two would have scores of 0 and 0. Five girls would have scores of 1 and 1 and five would have scores of 1 and 0.

SUMMARY

1. A correlation measures the relationship between two variables, X and Y. The relationship is described by three characteristics:

   a. *Direction.* A relationship can be either positive or negative. A positive relationship means that X and Y vary in the same direction. A negative relationship means that X and Y vary in opposite directions. The sign of the correlation (+ or –) specifies the direction.

   b. *Form.* The most common form for a relationship is a straight line, which is measured by the Pearson correlation. Other correlations measure the consistency or strength of the relationship, independent of any specific form.

   c. *Strength or consistency.* The numerical value of the correlation measures the strength or consistency of the relationship. A correlation of 1.00 indicates a perfectly consistent relationship and 0.00 indicates no relationship at all. For the Pearson correlation, $r = 1.00$ (or –1.00) means that the data points fit perfectly on a straight line.

2. The most commonly used correlation is the Pearson correlation, which measures the degree of linear relationship. The Pearson correlation is identified by the letter r and is computed by

$$r = \frac{SP}{\sqrt{SS_X SS_Y}}$$

   In this formula, SP is the sum of products of deviations and can be calculated with either a definitional formula or a computational formula:

   definitional formula: $SP = \Sigma(X - M_X)(Y - M_Y)$

   computational formula: $SP = \Sigma XY - \dfrac{\Sigma X \Sigma Y}{n}$

3. A correlation between two variables should not be interpreted as implying a causal relationship. Simply because X and Y are related does not mean that X causes Y or that Y causes X.

4. To evaluate the strength of a relationship, you square the value of the correlation. The resulting value, r^2, is called the *coefficient of determination* because it measures the portion of the variability in one variable that can be predicted using the relationship with the second variable.

5. A partial correlation measures the linear relationship between two variables by eliminating the influence of a third variable by holding it constant.

6. The Spearman correlation (r_S) measures the consistency of direction in the relationship between X and Y—that is, the degree to which the relationship is one-directional, or monotonic. The Spearman correlation is computed by a two-stage process:

   a. Rank the X scores and the Y scores separately.

   b. Compute the Pearson correlation using the ranks.

7. The point-biserial correlation is used to measure the strength of the relationship when one of the two variables is dichotomous. The dichotomous variable is coded using values of 0 and 1, and the regular Pearson formula is applied. Squaring the point-biserial correlation produces the same r^2 value that is obtained to measure effect size for the independent-measures t test. When both variables, X and Y, are dichotomous, the phi-coefficient can be used to measure the strength of the relationship. Both variables are coded 0 and 1, and the Pearson formula is used to compute the correlation.

KEY TERMS

correlation (510)
positive correlation (512)
negative correlation (512)
perfect correlation (513)
Pearson correlation (514)
sum of products (*SP*) (515)
restricted range (522)
coefficient of determination (524)
regression toward the mean (526)
correlation matrix (530)
partial correlation (531)
Spearman correlation (535)
point-biserial correlation (542)
phi-coefficient (545)

RESOURCES

Book Companion Website: www.cengage.com/psychology/gravetter

You can find a tutorial quiz and other learning exercises for Chapter 15 on the book companion website. The website also provides access to two workshops entitled *Correlation* and *Bivariate Scatter Plots,* which include information on regression.

aplia

Improve your understanding of statistics with Aplia's auto-graded problem sets and immediate, detailed explanations for every question. To learn more, visit **www.aplia.com/statistics.**

CENGAGE brain .com

Log in to CengageBrain to access the resources your instructor requires. For this book, you can access:

Psychology CourseMate brings course concepts to life with interactive learning, study, and exam preparation tools that support the printed textbook. A textbook-specific website, Psychology CourseMate includes an integrated interactive eBook and other interactive learning tools including quizzes, flashcards, and more.

Visit www.cengagebrain.com to access your account and purchase materials.

SPSS

General instructions for using SPSS are presented in Appendix D. Following are detailed instructions for using SPSS to perform **The Pearson, Spearman, point-biserial, and partial correlations.** *Note:* We focus on the Pearson correlation and then describe how slight modifications to this procedure can be made to compute the Spearman, point-biserial, and partial correlations. Separate instructions for the **phi-coefficient** are presented at the end of this section.

Data Entry

The data are entered into two columns in the data editor, one for the X values (VAR00001) and one for the Y values (VAR00002), with the two scores for each individual in the same row.

Data Analysis

1. Click **Analyze** on the tool bar, select **Correlate,** and click on **Bivariate.**
2. One by one, move the labels for the two data columns into the **Variables** box. (Hightlight each label and click the arrow to move it into the box.)

3. The **Pearson** box should be checked but, at this point, you can switch to the Spearman correlation by clicking the appropriate box.
4. Click **OK.**

SPSS Output

We used SPSS to compute the correlation for the data in Example 15.3 and the output is shown in Figure 15.16. The program produces a correlation matrix showing all the possible correlations, including the correlation of X with X and the correlation of Y with Y (both are perfect correlations). You want the correlation of X and Y, which is contained in the upper right corner (or the lower left). The output includes the significance level (p value or alpha level) for the correlation.

To compute a partial correlation, click **Analyze** on the tool bar, select **Correlate,** and click on **Partial.** Move the column labels for the two variables to be correlated into the **Variables** box and move the column label for the variable to be held constant into the **Controlling for** box and click **OK.**

To compute the **Spearman** correlation, enter either the X and Y ranks or the X and Y scores into the first two columns. Then follow the same Data Analysis instructions that were presented for the Pearson correlation. At step 3 in the instructions, click on the **Spearman** box before the final OK. (*Note:* If you enter X and Y scores into the data editor, SPSS converts the scores to ranks before computing the Spearman correlation.)

To compute the **point-biserial** correlation, enter the scores (X values) in the first column and enter the numerical values (usually 0 and 1) for the dichotomous variable in the second column. Then, follow the same Data Analysis instructions that were presented for the Pearson correlation.

The **phi-coefficient** can also be computed by entering the complete string of 0s and 1s into two columns of the SPSS data editor, then following the same Data Analysis instructions that were presented for the Pearson correlation. However, this can be tedious, especially with a large set of scores. The following is an alternative procedure for computing the phi-coefficient with large data sets.

FIGURE 15.16

The SPSS output for the correlation in Example 15.3.

Correlations

		VAR00001	VAR00002
VAR00001	Pearson Correlation	1	.875
	Sig. (2-tailed)		.052
	N	5	5
VAR00002	Pearson Correlation	.875	1
	Sig. (2-tailed)	.052	
	N	5	5

Data Entry

1. Enter the values, 0, 0, 1, 1 (in order) into the first column of the SPSS data editor.
2. Enter the values 0, 1, 0, 1 (in order) into the second column.
3. Count the number of individuals in the sample who are classified with $X = 0$ and $Y = 0$. Enter this frequency in the top box in the third column of the data editor. Then, count how many have $X = 0$ and $Y = 1$ and enter the frequency in the second box of the third column. Continue with the number who have $X = 1$ and $Y = 0$, and finally the number who have $X = 1$ and $Y = 1$. You should end up with 4 values in column three.
4. Click **Data** on the Tool Bar at the top of the SPSS Data Editor page and select **Weight Cases** at the bottom of the list.
5. Click the circle labeled **Weight cases by,** and then highlight the label for the column containing your frequencies (VAR00003) on the left and move it into the **Frequency Variable** box by clicking on the arrow.
6. Click **OK**.
7. Click **Analyze** on the tool bar, select **Correlate,** and click on **Bivariate.**
8. One by one, move the labels for the two data columns containing the 0s and 1s (probably VAR00001 and VAR00002) into the **Variables** box. (Highlight each label and click the arrow to move it into the box.)
9. Verify that the **Pearson** box is checked.
10. Click **OK.**

SPSS Output

The program produces the same correlation matrix that was described for the Pearson correlation. Again, you want the correlation between *X* and *Y*, which is in the upper right corner (or lower left). Remember, with the phi-coefficient, the sign of the correlation is meaningless.

FOCUS ON PROBLEM SOLVING

1. A correlation always has a value from +1.00 to –1.00. If you obtain a correlation outside this range, then you have made a computational error.

2. When interpreting a correlation, do not confuse the sign (+ or –) with its numerical value. The sign and the numerical value must be considered separately. Remember that the sign indicates the direction of the relationship between *X* and *Y*. On the other hand, the numerical value reflects the strength of the relationship or how well the points approximate a linear (straight-line) relationship. Therefore, a correlation of –0.90 is as strong as a correlation of +0.90. The signs tell us that the first correlation is an inverse relationship.

3. Before you begin to calculate a correlation, sketch a scatter plot of the data and make an estimate of the correlation. (Is it positive or negative? Is it near 1 or near 0?) After computing the correlation, compare your final answer with your original estimate.

4. The definitional formula for the sum of products (*SP*) should be used only when you have a small set (n) of scores and the means for *X* and *Y* are both whole numbers. Otherwise, the computational formula produces quicker, easier, and more accurate results.

5. For computing a correlation, n is the number of individuals (and therefore the number of *pairs* of X and Y values).

DEMONSTRATION 15.1

CORRELATION

Calculate the Pearson correlation for the following data:

Person	X	Y	
A	0	4	$M_X = 4$ with $SS_X = 40$
B	2	1	$M_Y = 6$ with $SS_Y = 54$
C	8	10	$SP = 40$
D	6	9	
E	4	6	

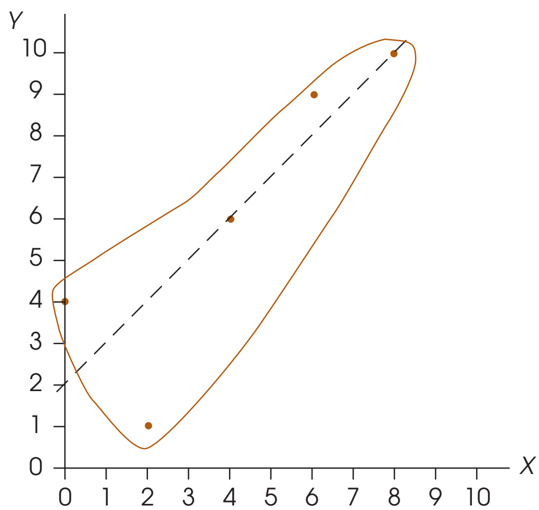
STEP 1 **Sketch a scatter plot.** We have constructed a scatter plot for the data (Figure 15.17) and placed an envelope around the data points to make a preliminary estimate of the correlation. Note that the envelope is narrow and elongated. This indicates that the correlation is large—perhaps 0.80 to 0.90. Also, the correlation is positive because increases in X are generally accompanied by increases in Y.

STEP 2 **Compute the Pearson correlation.** For these data, the Pearson correlation is

$$r = \frac{SP}{\sqrt{SS_X SS_Y}} = \frac{40}{\sqrt{40(54)}} = \frac{40}{\sqrt{2160}} = \frac{40}{46.48} = 0.861$$

FIGURE 15.17

The scatter plot for the data of Demonstration 15.1. An envelope is drawn around the points to estimate the magnitude of the correlation. A line is drawn through the middle of the envelope.



In step 1, our preliminary estimate for the correlation was between +0.80 and +0.90. The calculated correlation is consistent with this estimate.

STEP 3 **Evaluate the significance of the correlation.** The null hypothesis states that, for the population, there is no linear relationship between X and Y, and that the value obtained for the sample correlation is simply the result of sampling error. Specifically, H_0 says that the population correlation is zero ($\rho = 0$). With $n = 5$ pairs of X and Y values the test has $df = 3$. Table B.6 lists a critical value of 0.878 for a two-tailed test with $\alpha = .05$. Because our correlation is smaller than this value, we fail to reject the null hypothesis and conclude that the correlation is not significant.

PROBLEMS

1. What information is provided by the sign (+ or –) of the Pearson correlation?

2. What information is provided by the numerical value of the Pearson correlation?

3. Calculate SP (the sum of products of deviations) for the following scores. *Note:* Both means are whole numbers, so the definitional formula works well

X	Y
0	2
1	4
4	5
3	3
7	6

4. Calculate SP (the sum of products of deviations) for the following scores. *Note:* Both means are decimal values, so the computational formula works well.

X	Y
0	2
0	1
1	0
2	1
1	2
0	3

5. For the following scores,

X	Y
7	6
9	6
6	3
12	5
9	6
5	4

a. Sketch a scatter plot showing the six data points.
b. Just looking at the scatter plot, estimate the value of the Pearson correlation.
c. Compute the Pearson correlation.

6. For the following scores,

X	Y
1	3
3	5
2	1
2	3

a. Sketch a scatter plot and estimate the Pearson correlation.
b. Compute the Pearson correlation.

7. For the following scores,

X	Y
1	7
4	2
1	3
1	6
2	0
0	6
2	3
1	5

a. Sketch a scatter plot and estimate the Pearson correlation.
b. Compute the Pearson correlation.

8. For the following scores,

X	Y
1	6
4	1
1	4
1	3
3	1

a. Sketch a scatter plot and estimate the value of the Pearson correlation.
b. Compute the Pearson correlation.

9. With a small sample, a single point can have a large effect on the magnitude of the correlation. To create the following data, we started with the scores from problem 8 and changed the first X value from $X = 1$ to $X = 6$.

X	Y
6	6
4	1
1	4
1	3
3	1

a. Sketch a scatter plot and estimate the value of the Pearson correlation.
b. Compute the Pearson correlation.

10. For the following set of scores,

X	Y
6	4
3	1
5	0
6	7
4	2
6	4

a. Compute the Pearson correlation.
b. Add 2 points to each X value and compute the correlation for the modified scores. How does adding a constant to every score affect the value of the correlation?
c. Multiply each of the original X values by 2 and compute the correlation for the modified scores. How does multiplying each score by a constant affect the value of the correlation?

11. Correlation studies are often used to help determine whether certain characteristics are controlled more by genetic influences or by environmental influences. These studies often examine adopted children and compare their behaviors with the behaviors of their birth parents and their adoptive parents. One study examined how much time individuals spend watching TV (Plomin, Corley, DeFries, & Fulker, 1990). The following data are similar to the results obtained in the study.

Amount of Time Spent Watching TV

Adopted Children	Birth Parents	Adoptive Parents
2	0	1
3	3	4
6	4	2
1	1	0
3	1	0
0	2	3
5	3	2
2	1	3
5	3	3

a. Compute the correlation between the children and their birth parents.
b. Compute the correlation between the children and their adoptive parents.
c. Based on the two correlations, does TV watching appear to be inherited from the birth parents or is it learned from the adoptive parents?

12. Judge and Cable (2010) report the results of a study demonstrating a negative relationship between weight and income for a group of women professionals. Following are data similar to those obtained in the study. To simplify the weight variable, the women are classified into five categories that measure actual weight relative to height, from 1 = thinnest to 5 = heaviest. Income figures are annual income (in thousands), rounded to the nearest $1,000.
a. Calculate the Pearson correlation for these data.
b. Is the correlation statistically significant? Use a two-tailed test with $\alpha = .05$.

Weight (X)	Income (Y)
1	125
2	78
4	49
3	63
5	35
2	84
5	38
3	51
1	93
4	44

13. The researchers cited in the previous problem also examined the weight/salary relationship for men and found a positive relationship, suggesting that we have very different standards for men than for women

(Judge & Cable, 2010). The following are data similar to those obtained for working men. Again, weight relative to height is coded in five categories from 1 = thinnest to 5 = heaviest. Income is recorded as thousands earned annually.

a. Calculate the Pearson correlation for these data.
b. Is the correlation statistically significant? Use a two-tailed test with $\alpha = .05$.

Weight (X)	Income (Y)
4	156
3	88
5	49
2	73
1	45
3	92
1	53
5	148

14. Identifying individuals with a high risk of Alzheimer's disease usually involves a long series of cognitive tests. However, researchers have developed a 7-Minute Screen, which is a quick and easy way to accomplish the same goal. The question is whether the 7-Minute Screen is as effective as the complete series of tests. To address this question, Ijuin et al. (2008) administered both tests to a group of patients and compared the results. The following data represent results similar to those obtained in the study.

Patient	7-Minute Screen	Cognitive Series
A	3	11
B	8	19
C	10	22
D	8	20
E	4	14
F	7	13
G	4	9
H	5	20
I	14	25

a. Compute the Pearson correlation to measure the degree of relationship between the two test scores.
b. Is the correlation statistically significant? Use a two-tailed test with $\alpha = .01$.
c. What percentage of variance for the cognitive scores is predicted from the 7-Minute Screen scores? (Compute the value of r^2.)

15. Assuming a two-tailed test with $\alpha = .05$, how large a correlation is needed to be statistically significant for each of the following samples?
a. A sample of $n = 8$
b. A sample of $n = 18$
c. A sample of $n = 28$

16. As we have noted in previous chapters, even a very small effect can be significant if the sample is large enough. For each of the following, determine how large a sample is necessary for the correlation to be significant. Assume a two-tailed test with $\alpha = .05$. (*Note:* The table does not list all the possible *df* values. Use the sample size corresponding to the appropriate *df* value that is listed in the table.)
a. A correlation of $r = 0.30$.
b. A correlation of $r = 0.25$.
c. A correlation of $r = 0.20$.

17. A researcher measures three variables, X, Y, and Z, for each individual in a sample of $n = 25$. The Pearson correlations for this sample are $r_{XY} = 0.8$, $r_{XZ} = 0.6$, and $r_{YZ} = 0.7$.
a. Find the partial correlation between X and Y, holding Z constant.
b. Find the partial correlation between X and Z, holding Y constant. (*Hint:* Simply switch the labels for the variables Y and Z to correspond with the labels in the equation.)

18. A researcher records the annual number of serious crimes and the amount spent on crime prevention for several small towns, medium cities, and large cities across the country. The resulting data show a strong positive correlation between the number of serious crimes and the amount spent on crime prevention. However, the researcher suspects that the positive correlation is actually caused by population; as population increases, both the amount spent on crime prevention and the number of crimes also increases. If population is controlled, there probably should be a negative correlation between the amount spent on crime prevention and the number of serious crimes. The following data show the pattern of results obtained. Note that the municipalities are coded in three categories. Use a partial correlation, holding population constant, to measure the true relationship between crime rate and the amount spent on prevention.

Number of Crimes	Amount for Prevention	Population Size
3	6	1
4	7	1
6	3	1
7	4	1
8	11	2
9	12	2
11	8	2
12	9	2
13	16	3
14	17	3
16	13	3
17	14	3

19. A common concern for students (and teachers) is the assignment of grades for essays or term papers. Because there are no absolute right or wrong answers, these grades must be based on a judgment of quality. To demonstrate that these judgments actually are reliable, an English instructor asked a colleague to rank-order a set of term papers. The ranks and the instructor's grades for these papers are as follows:

Rank	Grade
1	A
2	B
3	A
4	B
5	B
6	C
7	D
8	C
9	C
10	D
11	F

a. Compute the Spearman correlation for these data. (*Note:* You must convert the letter grades to ranks, using tied ranks to represent tied grades.)
b. Is the Spearman correlation statistically significant? Use a two-tailed test with $\alpha = .05$.

20. It appears that there is a significant relationship between cognitive ability and social status, at least for birds. Boogert, Reader, and Laland (2006) measured social status and individual learning ability for a group of starlings. The following data represent results similar to those obtained in the study. Because social status is an ordinal variable consisting of five ordered categories, the Spearman correlation is appropriate for these data. Convert the social status categories and the learning scores to ranks, and compute the Spearman correlation.

Subject	Social Status	Learning Score
A	1	3
B	3	10
C	2	7
D	3	11
E	5	19
F	4	17
G	5	17
H	2	4
I	4	12
J	2	3

21. Problem 12 presented data showing a negative relationship between weight and income for a sample of working women. However, weight was coded in five categories, which could be viewed as an ordinal scale rather than an interval or ratio scale. If so, a Spearman correlation is more appropriate than a Pearson correlation.
a. Convert the weights and the incomes into ranks and compute the Spearman correlation for the scores in problem 12.
b. Is the Spearman correlation large enough to be significant?

22. Problem 22 in Chapter 10 presented data showing that mature soccer players, who have a history of hitting soccer balls with their heads, had significantly lower cognitive scores than mature swimmers, who do not suffer repeated blows to the head. The independent-measures t test produced $t = 2.11$ with $df = 11$ and a value of $r^2 = 0.288$ (28.8%).
a. Convert the data from this problem into a form suitable for the point-biserial correlation (use 1 for the swimmers and 0 for the soccer players), and then compute the correlation.
b. Square the value of the point-biserial correlation to verify that you obtain the same r^2 value that was computed in Chapter 10.

23. Problem 14 in Chapter 10 described a study by Rozin, Bauer, and Cantanese (2003) comparing attitudes toward eating for male and female college students. The results showed that females are much more concerned about weight gain and other negative aspects of eating than are males. The following data represent the results from one measure of concern about weight gain.

Males	Females
22	54
44	57
39	32
27	53
35	49
19	41
	35
	36
	48

Convert the data into a form suitable for the point-biserial correlation and compute the correlation.

24. Studies have shown that people with high intelligence are generally more likely to volunteer as participants in research, but not for research that involves unusual experiences such as hypnosis. To examine this phenomenon, a researcher administers a questionnaire to a sample of college students. The survey asks for the student's grade point average (as a measure of intelligence) and whether the student would like to take part in a future study in which participants would be hypnotized. The results showed that 7 of the 10 lower-intelligence people were willing to participant but only 2 of the 10 higher-intelligence people were willing.

a. Convert the data to a form suitable for computing the phi-coefficient. (Code the two intelligence categories as 0 and 1 for the X variable, and code the willingness to participate as 0 and 1 for the Y variable.)

b. Compute the phi-coefficient for the data.