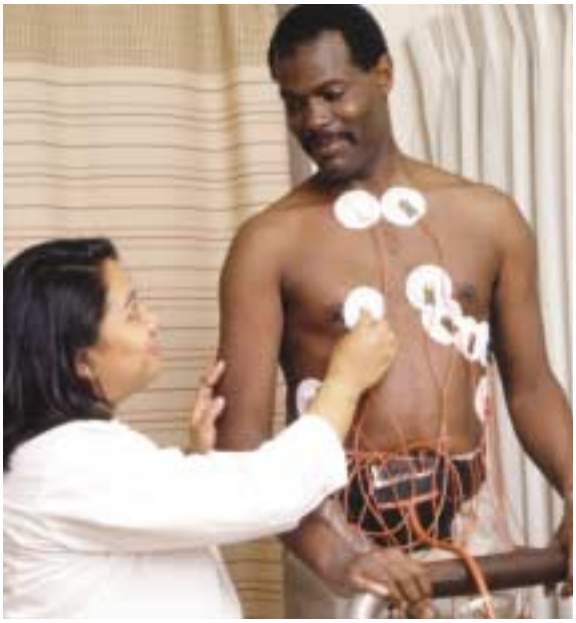
3.4 Cause and Effect



Usually, the main reason for a correlational study is to find evidence of a cause-and-effect relationship. A health researcher may wish to prove that even mild exercise reduces the risk of heart disease. A chemical company developing an oil additive would like to demonstrate that it improves engine performance. A school board may want to know whether calculators help students learn mathematics. In each of these cases, establishing a strong correlation between the variables is just the first step in determining whether one affects the other.

INVESTIGATE & INQUIRE: Correlation Versus Cause and Effect

1. List the type of correlation that you would expect to observe between the following pairs of variables. Also list whether you think the correlation is due to a cause-and-effect relationship or some other factor.
   a) hours spent practising at a golf driving range, golf drive distance
   b) hours spent practising at a golf driving range, golf score
   c) size of corn harvest, size of apple harvest
   d) score on a geometry test, score on an algebra test
   e) income, number of CDs purchased
2. Compare your list with those of your classmates and discuss any differences. Would you change your list because of factors suggested by your classmates?
3. Suggest how you could verify whether there is a cause-and-effect relationship between each pair of variables.

A strong correlation does not prove that the changes in one variable cause changes in the other. There are various types and degrees of causal relationships between variables.

Cause-and-Effect Relationship: A change in X produces a change in Y. Such relationships are sometimes clearly evident, especially in physical processes. For example, increasing the height from which you drop an object increases its impact velocity. Similarly, increasing the speed of a production line increases the number of items produced each day (and, perhaps, the rate of defects).

Common-Cause Factor: An external variable causes two variables to change in the same way. For example, suppose that a town finds that its revenue from parking fees at the public beach each summer correlates with the local tomato harvest. It is extremely unlikely that cars parked at the beach have any effect on the tomato crop. Instead good weather is a common-cause factor that increases both the tomato crop and the number of people who park at the beach.

Reverse Cause-and-Effect Relationship: The dependent and independent variables are reversed in the process of establishing causality. For example, suppose that a researcher observes a positive linear correlation between the amount of coffee consumed by a group of medical students and their levels of anxiety. The researcher theorizes that drinking coffee causes nervousness, but instead finds that nervous people are more likely to drink coffee.

Accidental Relationship: A correlation exists without any causal relationship between variables. For example, the number of females enrolled in undergraduate engineering programs and the number of "reality" shows on television both increased for several years. These two variables have a positive linear correlation, but it is likely entirely coincidental.

Presumed Relationship: A correlation does not seem to be accidental even though no cause-and-effect relationship or common-cause factor is apparent. For example, suppose you found a correlation between people's level of fitness and the number of adventure movies they watched. It seems logical that a physically fit person might prefer adventure movies, but it would be difficult to find a common cause or to prove that the one variable affects the other.

Example 1 Causal Relationships

Classify the relationships in the following situations.

a) The rate of a chemical reaction increases with temperature.

b) Leadership ability has a positive correlation with academic achievement.

c) The prices of butter and motorcycles have a strong positive correlation over many years.

d) Sales of cellular telephones had a strong negative correlation with ozone levels in the atmosphere over the last decade.

e) Traffic congestion has a strong correlation with the number of urban expressways.

Solution

a) Cause-and-effect relationship: Higher temperatures cause faster reaction rates.

b) Presumed relationship: A positive correlation between leadership ability and academic achievement seems logical, yet there is no apparent common-cause factor or cause-and-effect relationship.

c) Common-cause factor: Inflation has caused parallel increases in the prices of butter and motorcycles over the years.

d) Accidental relationship: The correlation between sales of cellular telephones and ozone levels is largely coincidental. However, it is possible that the chemicals used to manufacture cellular telephones cause a small portion of the depletion of the ozone layer.

e) Cause-and-effect relationship and reverse cause-and-effect relationship: Originally expressways were built to relieve traffic congestion, so traffic congestion did lead to the construction of expressways in major cites throughout North America. However, numerous studies over the last 20 years have shown that urban expressways cause traffic congestion by encouraging more people to use cars.

As Example 1 demonstrates, several types of causal relationships can be involved in the same situation. Determining the nature of causal relationships can be further complicated by the presence of **extraneous variables** that affect either the dependent or the independent variable. Here, *extraneous* means external rather than irrelevant.

For example, you might expect to see a strong positive correlation between term marks and final examination results for students in your class since both these variables are affected by each student's aptitude and study habits. However, there are extraneous factors that could affect the examination results, including the time each student had for studying before the examination, the individual examination schedules, and varying abilities to work well under pressure.

In order to reduce the effect of extraneous variables, researchers often compare an **experimental group** to a **control group**. These two groups should be as similar as possible, so that extraneous variables will have about the same effect on both groups. The researchers vary the independent variable for the experimental group but not for the control group. Any *difference* in the dependent variables for the two groups can then be attributed to the changes in the independent variable.

Example 2 Using a Control Group

A medical researcher wants to test a new drug believed to help smokers overcome the addictive effects of nicotine. Fifty people who want to quit smoking volunteer for the study. The researcher carefully divides the volunteers into two groups, each with an equal number of moderate and heavy smokers. One group is given nicotine patches with the new drug, while the second group uses ordinary nicotine patches. Fourteen people in the first group quit smoking completely, as do nine people in the second group.

a) Identify the experimental group, the control group, the independent variable, and the dependent variable.

b) Can the researcher conclude that the new drug is effective?

c) What further study should the researcher do?

Solution

a) The experimental group consists of the volunteers being given nicotine patches with the new drug, while the control group consists of the volunteers being given the ordinary patches. The independent variable is the presence of the new drug, and the dependent variable is the number of volunteers who quit smoking.

b) The results of the study are promising, but the researcher has not proven that the new drug is effective. The sample size is relatively small, which is prudent for an early trial of a new drug that could have unknown side-effects. However, the sample is small enough that the results could be affected by random statistical fluctuations or extraneous variables, such as the volunteers' work environments, previous attempts to quit, and the influence of their families and friends.

c) Assuming that the new drug does not have any serious side-effects, the researcher should conduct further studies with larger groups and try to select the experimental and control groups to minimize the effect of all extraneous variables. The researcher might also conduct a study with several experimental groups that receive different dosages of the new drug.

When designing a study or interpreting a correlation, you often need background knowledge and insight to recognize the causal relationships present. Here are some techniques that can help determine whether a correlation is the result of a cause-and-effect relationship.

- Use sampling methods that hold the extraneous variables constant.
- Conduct similar investigations with different samples and check for consistency in the results.
- Remove, or account for, possible common-cause factors.

The later chapters in this book introduce probability theory and some statistical methods for a more quantitative approach to determining cause-and-effect relationships.

Project Prep

In your statistics project, you may wish to consider cause-and-effect relationships and extraneous variables that could affect your study.

Key Concepts

- Correlation does not necessarily imply a cause-and-effect relationship. Correlations can also result from common-cause factors, reverse cause-and-effect relationships, accidental relationships, and presumed relationships.
- Extraneous variables can invalidate conclusions based on correlational evidence.
- Comparison with a control group can help remove the effect of extraneous variables in a study.

Communicate Your Understanding

1. Why does a strong linear correlation not imply cause and effect?
2. What is the key characteristic of a reverse cause-and-effect relationship?
3. Explain the difference between a common-cause factor and an extraneous variable.
4. Why are control groups used in statistical studies?

Practise

A

1. Identify the most likely type of causal relationship between each of the following pairs of variables. Assume that a strong positive correlation has been observed with the first variable as the independent variable.
   - **a)** alcohol consumption, incidence of automobile accidents
   - **b)** score on physics examination, score on calculus examination
   - **c)** increase in pay, job performance
   - **d)** population of rabbits, consumer price index
   - **e)** number of scholarships received, number of job offers upon graduation
   - **f)** coffee consumption, insomnia
   - **e)** funding for athletic programs, number of medals won at Olympic games

2. For each of the following common-cause relationships, identify the common-cause factor. Assume a positive correlation between each pair of variables.
   a) number of push-ups performed in one minute, number of sit-ups performed in one minute
   b) number of speeding tickets, number of accidents
   c) amount of money invested, amount of money spent

Apply, Solve, Communicate

3. A civil engineer examining traffic flow problems in a large city observes that the number of traffic accidents is positively correlated with traffic density and concludes that traffic density is likely to be a major cause of accidents. What alternative conclusion should the engineer consider?

B

4. **Communication** An elementary school is testing a new method for teaching grammar. Two similar classes are taught the same material, one with the established method and the other with the new method. When both classes take the same test, the class taught with the established method has somewhat higher marks.
   a) What extraneous variables could influence the results of this study?
   b) Explain whether the study gives the school enough evidence to reject the new method.
   c) What further studies would you recommend for comparing the two teaching methods?

5. **Communication** An investor observes a positive correlation between the stock price of two competing computer companies. Explain what type of causal relationship is likely to account for this correlation.

6. **Application** A random survey of students at Statsville High School found that their interest in computer games is positively correlated with their marks in mathematics.
   a) How would you classify this causal relationship?
   b) Suppose that a follow-up study found that students who had increased the time they spent playing computer games tended to improve their mathematics marks. Assuming that this study held all extraneous variables constant, would you change your assessment of the nature of the causal relationship? Explain why or why not.

7. a) The net assets of Custom Industrial Renovations Inc., an industrial construction contractor, has a strong negative linear correlation with those of MuchMega-Fun, a toy distributor. How would you classify the causal relationship between these two variables?
   b) Suppose that the two companies are both subsidiaries of Diversified Holdings Ltd., which often shifts investment capital between them. Explain how this additional information could change your interpretation of the correlation in part a).

8. **Communication** Aunt Gisele simply cannot sleep unless she has her evening herbal tea. However, the package for the tea does not list any ingredients known to induce sleep. Outline how you would conduct a study to determine whether the tea really does help people sleep.

9. Find out what a *double-blind* study is and briefly explain the advantages of using this technique in studies with a control group.

10. a) The data on page 157 show a positive correlation between the size of the graduating class and the number of

Chapter Problem

graduates hired. Does this correlation mean that increasing the number of graduates causes a higher demand for them? Explain your answer.

b) A recession during the first half of the 1990s reduced the demand for business graduates. Review the data on page 157 and describe any trends that may be caused by this recession.

ACHIEVEMENT CHECK

Knowledge / Understanding	Thinking / Inquiry / Problem Solving	Communication	Application

11. The table below lists numbers of divorces and personal bankruptcies in Canada for the years 1976 through 1985.

Year	Divorces	Bankruptcies
1976	54 207	10 049
1977	55 370	12 772
1978	57 155	15 938
1979	59 474	17 876
1980	62 019	21 025
1981	67 671	23 036
1982	70 436	30 643
1983	68 567	26 822
1984	65 172	22 022
1985	61 976	19 752

a) Create a scatter plot and classify the linear correlation between the number of divorces and the number of bankruptcies.

b) Perform a regression analysis. Record the equation of the line of best fit and the correlation coefficient.

c) Identify an external variable that could be a common-cause factor.

d) Describe what further investigation you could do to analyse the possible relationship between divorces and bankruptcies.

12. Search the E-STAT, CANSIM II, or other databases for a set of data on two variables with a positive linear correlation that you believe to be accidental. Explain your findings and reasoning.

13. Use a library, the Internet, or other resources to find information on the Hawthorne effect and the placebo effect. Briefly explain what these effects are, how they can affect a study, and how researchers can avoid having their results skewed by these effects.

14. **Inquiry / Problem Solving** In a behavioural study of responses to violence, an experimental group was shown violent images, while a control group was shown neutral images. From the initial results, the researchers suspect that the gender of the people in the groups may be an extraneous variable. Suggest how the study could be redesigned to

a) remove the extraneous variable

b) determine whether gender is part of the cause-and-effect relationship

15. Look for material in the media or on the Internet that incorrectly uses correlational evidence to claim that a cause-and-effect relationship exists between the two variables. Briefly describe

a) the nature of the correlational study

b) the cause and effect claimed or inferred

c) the reasons why cause and effect was not properly proven, including any extraneous variables that were not accounted for

d) how the study could be improved