while keeping the value of the control variable X_2 essentially constant. This helps us to detect

- **Spuriousness**, in which X_2 jointly affects both Y and X_1.
- **Chain relationships**, in which X_2 is an **intervening variable**, so that X_1 affects Y indirectly through its effects on X_2.
- **Suppressor variables**, in which the X_1–Y association appears only after controlling for X_2.
- **Statistical interaction**, in which X_2 affects the X_1–Y association, the nature or strength of that association varying according to the fixed value for X_2.

Table 10.8 summarizes some possible three-variable relationships. The remainder of this text presents statistical methods for multivariate relationships.

TABLE 10.8 Some Three-Variable Relationships

Graph	Name of Relationship	What Happens After Controlling for X_2
$X_2 \nearrow X_1$ $\searrow Y$	Spurious X_1–Y association	Association between X_1 and Y disappears.
$X_1 \longrightarrow X_2 \longrightarrow Y$	Chain relationship; X_2 intervenes; X_1 indirectly causes Y	Association between X_1 and Y disappears.
X_2 $\downarrow$ $X_1 \longrightarrow Y$	Interaction	Association between X_1 and Y varies according to level of X_2.
$X_2 \searrow Y$ $X_1 \nearrow$	Multiple causes	Association between X_1 and Y does not change.
$X_1 \longrightarrow Y$ $\searrow X_2 \nearrow$	Both direct and indirect effects of X_1 on Y	Association between X_1 and Y changes, but does not disappear.

PROBLEMS

Practicing the Basics

1. For each of the three criteria for a causal relationship, describe a relationship between two variables that is not causal because that criterion would be violated.
2. An association exists between college GPA and whether one has ever used marijuana.
   a) Explain how the direction of a causal arrow might go in either direction.
   b) Explain how a third variable might be responsible for the association.
3. Give an example of two variables for which the time order for a causal connection is ambiguous, and explain why.
4. Explain clearly what it means to *control* for a variable. Describe a situation in which an association would change dramatically after controlling a certain variable.

5. Explain what is meant by a *spurious* association.
   a) Illustrate using the variables X_1 = shoe size, X_2 = age, and Y = number of books one has ever read, for a random sample of children from schools in Winnipeg, Canada.
   b) Illustrate using the variables X_1 = shoe size, X_2 = gender, and Y = annual income, for a random sample of adults. Draw a scatter diagram to illustrate this case if, overall, men tend to have larger shoes and higher income, on the average, than females.

6. Refer to Problem 9.17 (Table 9.13). Those data exhibit a strong negative correlation between birth rate and per capita television ownership. While there *could* potentially be a causal relationship, there are other plausible explanations for this association. Using another variable from that exercise, provide an explanation for which this association is spurious.

7. Table 10.9 relates occupational level (white collar, blue collar) and political party choice, controlling for income.
   a) Construct the bivariate table between occupational level and political party, ignoring income. Is there an association? If so, describe it.
   b) Do the partial tables display an association? Interpret them.
   c) Using the nature of the association between income and each of the other variables, explain why the bivariate table has such different association than the partial tables.
   d) Do the partial tables exhibit statistical interaction? Explain.
   e) Construct a chain diagram that might explain the relationships for these data, identifying the intervening variable.
   f) Show that the data are also consistent with a spurious association, and draw the corresponding diagram.
   g) Refer to parts (e) and (f). Which diagram seems more appropriate? Why?

TABLE 10.9

		High Income		Medium High Income		Medium Low Income		Low Income	
		White Collar	Blue Collar	White Collar	Blue Collar	White Collar	Blue Collar	White Collar	Blue Collar
Party	Democrat	45	5	100	25	75	300	45	405
	Republican	405	45	300	75	25	100	5	45

8. In murder trials in 20 Florida counties during 1976 and 1977, the death penalty was given in 19 out of 151 cases in which a white killed a white, in 0 out of 9 cases in which a white killed a black, in 11 out of 63 cases in which a black killed a white, and in 6 out of 103 cases in which a black killed a black (M. Radelet, *American Sociological Review*, Vol. 46, 1981, pp. 918–927).
   a) Construct partial tables to study the association between defendant's race and the death penalty verdict, controlling for victim's race.
   b) Describe the association in the partial tables, using the difference of proportions or the odds ratio.
   c) Construct the bivariate table, ignoring rather than controlling victim's race. Describe the association, and compare to (b).
   d) *Simpson's paradox* states that the associations in partial tables can all have a different direction than the association in the bivariate table. Show that these data satisfy Simpson's paradox.

9. The news media report that a study has found that children who eat breakfast get better math grades than those who do not eat breakfast. This result was based on a simple bivariate association, with X = whether eat breakfast (yes, no) and Y = grade in last math course taken. How might this result be spurious, and how could you check for that possibility?

10. For a particular Big Ten university, the mean income for male faculty is $8000 higher than the mean income for female faculty. Explain how this difference could disappear:
    a) Controlling for number of years since received highest degree, if male professors tend to be older and more experienced, relatively few female professors having been hired until recent years.
    b) Controlling for college of employment, if relatively more female faculty are in low-salary colleges (Arts and Science, Education) and relatively more male faculty are in high-salary colleges (Medical School, Law, Engineering).
    c) In (a) and (b), explain how the variables might satisfy a chain relationship.

11. Refer to the variables in Table 9.4. The number of bedrooms has a moderately strong positive correlation with selling price ($r = .59$). Controlling for size of home, however, this association diminishes greatly. Explain how this could happen, illustrating with a diagram showing potential direct and indirect effects of number of bedrooms on selling price.

12. Refer to the variables in Table 9.16. Perhaps surprisingly, a moderate positive correlation exists between crime rate and percent who are at least high school graduates ($r = .468$). Percentage living in metropolitan areas is also strongly positively correlated both with crime rate ($r = .678$) and with high school graduation rate ($r = .791$).
    a) Explain why the association between crime rate and high school graduation rate could disappear, or even change direction, when we control for percentage living in metropolitan areas.
    b) Which type of relationship is more plausible for these variables, a spurious relationship or a chain relationship? Explain.

13. Give an example of three variables for which the effect of X_1 on Y would be
    a) Spurious, disappearing when X_2 is controlled.
    b) Part of a chain relationship, disappearing when an intervening variable X_2 is controlled.
    c) Weakened, but not eliminated, when X_2 is controlled.
    d) Unaffected by controlling X_2.
    e) Different at different levels of X_2 (i.e., showing interaction).

14. Opposition to the legal availability of abortion is stronger among the very religious than the nonreligious, and it is also stronger among those with conservative sexual attitudes than those with more permissive attitudes.
    a) Draw a three-variable diagram of how these variables might be related, treating abortion attitude as the response variable. (Note: More than one diagram is plausible.)
    b) Explain how you would test your diagram, showing potential sample data that are consistent with it.

15. Table 10.10 lists the mean salary, in thousands of dollars, of full-time instructional faculty on nine-month contracts in United States institutions of higher education in 1993–1994, by gender and rank.
    a) Suppose that gender is the explanatory variable. Identify the response variable and the control variable.
    b) Describe the bivariate relationship between gender and salary.
    c) Describe the relationship between gender and salary, controlling for academic rank.