SECTION 2.5 Summary

A **two-way table** of counts organizes data about two categorical variables. Values of the **row variable** label the rows that run across the table, and values of the **column variable** label the columns that run down the table. Two-way tables are often used to summarize large amounts of data by grouping outcomes into categories.

The **joint distribution** of the row and column variables is found by dividing the count in each cell by the total number of observations.

The **row totals** and **column totals** in a two-way table give the **marginal distributions** of the two variables separately. It is clearer to present these distributions as percents of the table total. Marginal distributions do not give any information about the relationship between the variables.

To find the **conditional distribution** of the row variable for one specific value of the column variable, look only at that one column in the table. Find each entry in the column as a percent of the column total.

There is a conditional distribution of the row variable for each column in the table. Comparing these conditional distributions is one way to describe the association between the row and the column variables. It is particularly useful when the column variable is the explanatory variable. When the row variable is explanatory, find the conditional distribution of the column variable for each row and compare these distributions.

Bar graphs are a flexible means of presenting categorical data. There is no single best way to describe an association between two categorical variables.

We present data on three categorical variables in a **three-way table,** printed as separate two-way tables for each level of the third variable. A comparison between two variables that holds for each level of a third variable can be changed or even reversed when the data are **aggregated** by summing over all levels of the third variable. **Simpson's paradox** refers to the reversal of a comparison by aggregation. It is an example of the potential effect of lurking variables on an observed association.

SECTION 2.5 Exercises

For Exercise 2.115, see page 137; for 2.116, see page 138; for 2.117, see page 139; for 2.118, see page 140; for 2.119, see page 141; and for 2.120, see page 142.

2.121 Exercise and adequate sleep. A survey of 656 boys and girls who were 13 to 18 years old asked about adequate sleep and other health-related behaviors. The recommended amount of sleep is six to eight hours per night.[34] In the survey 54% of the respondents reported that they got less than this amount of sleep on school nights. An exercise scale was developed and was used to classify the students as above or below the median in this domain. Here is the 2 × 2 table of counts with students classified as getting or not getting adequate sleep and by the exercise variable: SLEEP

	Exercise	
Enough sleep	**High**	**Low**
Yes	151	115
No	148	242

(a) Find the distribution of adequate sleep for the high exercisers.

(b) Do the same for the low exercisers.

(c) Summarize the relationship between adequate sleep and exercise using the results of parts (a) and (b).

2.122 Adequate sleep and exercise. Refer to the previous exercise. SLEEP

(a) Find the distribution of exercise for those who get adequate sleep.

(b) Do the same for those who do not get adequate sleep.

(c) Write a short summary of the relationship between adequate sleep and exercise using the results of parts (a) and (b).

(d) Compare this summary with the summary that you obtained in part (c) of the previous exercise. Which do you prefer? Give a reason for your answer.

2.123 Which hospital is safer? Insurance companies and consumers are interested in the performance of hospitals. The government releases data about patient outcomes in hospitals that can be useful in making informed health care decisions. Here is a two-way table of data on the survival of patients after surgery in two hospitals. All patients undergoing surgery in a recent time period are included. "Survived" means that the patient lived at least 6 weeks following surgery. HOSPITALS

	Hospital A	Hospital B
Died	63	16
Survived	2037	784
Total	2100	800

What percent of Hospital A patients died? What percent of Hospital B patients died? These are the numbers one might see reported in the media.

2.124 Patients in "poor" or "good" condition. Refer to the previous exercise. Not all surgery cases are equally serious, however. Patients are classified as being in either "poor" or "good" condition before surgery. Here are the data broken down by patient condition. The entries in the original two-way table are just the sums of the "poor" and "good" entries in this pair of tables. HOSPITALS

	Good Condition		Poor Condition	
	Hospital A	Hospital B	Hospital A	Hospital B
Died	6	8	57	8
Survived	594	592	1443	192
Total	600	600	1500	200

(a) Find the death rate for Hospital A patients who were classified as "poor" before surgery. Do the same for Hospital B. In which hospital do "poor" patients fare better?

(b) Repeat (a) for patients classified as "good" before surgery.

(c) What is your recommendation to someone facing surgery and choosing between these two hospitals?

(d) How can Hospital A do better in both groups, yet do worse overall? Look at the data and carefully explain how this can happen.

2.125 Full-time and part-time college students. The Census Bureau provides estimates of numbers of people in the United States classified in various ways.[35] Let's look at college students. The following table gives us data to examine the relation between age and full-time or part-time status. The numbers in the table are expressed as thousands of U.S. college students. USCOLLEGESTUDENTS

U.S. college students by age and status

	Status	
Age	**Full-time**	**Part-time**
15–19	3388	389
20–24	5238	1164
25–34	1703	1699
35 and over	762	2045

(a) What is the U.S. Census Bureau estimate of the number of full-time college students aged 15 to 19?

(b) Give the joint distribution of age and status for this table.

(c) What is the marginal distribution of age? Display the results graphically.

(d) What is the marginal distribution of status? Display the results graphically.

2.126 Condition on age. Refer to the previous exercise. Find the conditional distribution of status for each of the four age categories. Display the distributions graphically and summarize their differences and similarities. USCOLLEGESTUDENTS

2.127 Condition on status. Refer to the previous two exercises. Compute the conditional distribution of age for each of the two status categories. Display the distributions graphically and write a short paragraph describing the distributions and how they differ. USCOLLEGESTUDENTS

2.128 Complete the table. Here are the row and column totals for a two-way table with two rows and two columns:

a	b	200
c	d	200
200	200	400

offered this service.[22] Here are the results classified by the asset size (in millions of dollars) of the bank: RDCASSET

	Offer RDC	
Asset size	**Yes**	**No**
Under $100	63	309
$101–200	59	132
$201 or more	112	85

(a) Summarize the results of this survey question numerically and graphically.

(b) Test the null hypothesis that there is no association between the size of a bank, measured by assets, and whether or not they offer RDC. Report the test statistic, the P-value, and your conclusion.

9.23 Exercise and adequate sleep. A survey of 656 boys and girls who were 13 to 18 years old asked about adequate sleep and other health-related behaviors. The recommended amount of sleep is six to eight hours per night.[23] In the survey 54% of the respondents reported that they got less than this amount of sleep on school nights. An exercise scale was developed and was used to classify the students as above or below the median in this domain. Here is the 2 × 2 table of counts with students classified as getting or not getting adequate sleep and by the exercise variable: SLEEP

	Exercise	
Enough sleep	**High**	**Low**
Yes	151	115
No	148	242

Note that you answered parts (a) through (c) of this exercise if you completed Exercise 2.121 (page 146).

(a) Find the distribution of adequate sleep for the high exercisers.

(b) Do the same for the low exercisers.

(c) Summarize the relationship between adequate sleep and exercise using the results of parts (a) and (b).

(d) Perform the significance test for examining the relationship between exercise and getting enough sleep. Give the test statistic, the P-value (with a sketch similar to the one on page 518) and summarize your conclusion. Be sure to include numerical and graphical summaries.

9.24 Lying to a teacher. One of the questions in a survey of high school students asked about lying to teachers.[24] The following table gives the numbers of students who said that they lied to a teacher at least once during the past year, classified by gender. LYING

	Gender	
Lied at least once	**Male**	**Female**
Yes	3,228	10,295
No	9,659	4,620

(a) Add the marginal totals to the table.

(b) Calculate appropriate percents to describe the results of this question.

(c) Summarize your findings in a short paragraph.

(d) Test the null hypothesis that there is no association between gender and lying to teachers. Give the test statistic, the P-value and summarize your conclusion. Be sure to include numerical and graphical summaries.

9.25 Trust and honesty in the workplace. The students surveyed in the study described in the previous exercise were also asked whether they thought trust and honesty were essential in business and the workplace. Here are the counts classified by gender: TRUST

	Gender	
Trust and honesty are essential	**Male**	**Female**
Agree	11,724	14,169
Disagree	1,163	746

Answer questions given in the previous exercise for this question.

9.26 Why not use a chi-square test? As part of the study on ongoing fright symptoms due to exposure to horror movies at a young age, the following table was created based on the written responses from 119 students.[25] Explain why a chi-square test is not appropriate for this table.

Percent of students who reported each problem

	Type of Problem			
	Bedtime		Waking	
Movie or video	**Short term**	**Enduring**	**Short term**	**Enduring**
Poltergeist ($n = 29$)	68	7	64	32
Jaws ($n = 23$)	39	4	83	43
Nightmare on Elm Street ($n = 16$)	69	13	37	31
Thriller (music video) ($n = 16$)	40	0	27	7
It ($n = 24$)	64	0	64	50
The Wizard of Oz ($n = 12$)	75	17	50	8
E.T. ($n = 11$)	55	0	64	27

9.27 Age and time status of U.S. college students. The Census Bureau provides estimates of numbers of people in the United States classified in various ways.[26] Let's look at college students. The following table gives us data to examine the relation between age and full-time or part-time status. The numbers in the table are expressed as thousands of U.S. college students. USCOLLEGE1

U.S. college students by age and status: October 2004

	Status	
Age	**Full-time**	**Part-time**
15–19	3553	329
20–24	5710	1215
25–34	1825	1864
35 and over	901	1983

(a) Give the joint distribution of age and status for this table.

(b) What is the marginal distribution of age? Display the results graphically.

(c) What is the marginal distribution of status? Display the results graphically.

(d) Compute the conditional distribution of age for each of the two status categories. Display the results graphically.

(e) Write a short paragraph describing the distributions and how they differ.

9.28 Time status versus gender for the 20–24 age category. Refer to Exercise 9.27. The table below breaks down the 20–24 age category by gender. USCOLLEGE2

	Gender		
Status	**Male**	**Female**	**Total**
Full-time	2719	2991	5710
Part-time	535	680	1215
Total	3254	3671	6925

(a) Compute the marginal distribution for gender. Display the results graphically.

(b) Compute the conditional distribution of status for males and for females. Display the results graphically and comment on how these distributions differ.

(c) If you wanted to test the null hypothesis that there is no difference between these two conditional distributions, what would the expected cell counts be for the full-time status row of the table?

(d) Computer software gives $X^2 = 5.17$. Using Table F, give an appropriate bound for the P-value and state your conclusions at the 5% level.

9.29 Waking versus bedtime symptoms. As part of the study on ongoing fright symptoms due to exposure to horror movies at a young age, the following table was presented to describe the lasting impact these movies have had during bedtime and waking life: FRIGHTSYMPTOMS

	Waking symptoms	
Bedtime symptoms	**Yes**	**No**
Yes	36	33
No	33	17

(a) What percent of the students have lasting waking-life symptoms?

(b) What percent of the students have both waking-life and bedtime symptoms?

(c) Test whether there is an association between waking-life and bedtime symptoms. State the null and alternative hypotheses, the X^2 statistic, and the P-value.

9.30 Find the degrees of freedom and P-value. For each of the following situations give the degrees of freedom and an appropriate bound on the P-value (give the exact value if you have software available) for the X^2 statistic for testing the null hypothesis of no association between the row and column variables.

(a) A 2 by 2 table with $X^2 = 1.25$.

(b) A 4 by 4 table with $X^2 = 18.34$.

(c) A 2 by 8 table with $X^2 = 24.21$.

(d) A 5 by 3 table with $X^2 = 12.17$.

9.31 Can you construct the joint distribution from the marginal distributions? Here are the row and column totals for a two-way table with two rows and two columns:

a	b	50
c	d	150
100	100	200

Find *two different* sets of counts a, b, c, and d for the body of the table. This demonstrates that the relationship between two variables cannot be obtained solely from the two marginal distributions of the variables.

9.32 Construct a table with no association. Construct a 3 × 2 table of counts where there is no apparent association between the row and column variables.

9.33 Gender versus motivation for volunteer service. A study examined patterns and characteristics of volunteer service for young people from high school

Education level	Pet ownership status: Non-pet owners	Dog owners	Cat owners
Less than high school	421	93	28
High school graduate	666	100	40
Postsecondary	845	135	99

Note that "Dog owners" and "Cat owners" designate individuals who own a dog only or a cat only, respectively. Individuals who own both a dog and a cat are not included in this table. Analyze the data. Include numerical and graphical summaries as well as a significance test. Summarize your results and conclusions.

9.44 Pet ownership and gender. Refer to the previous exercise. Here are similar data giving the relationship between pet ownership status and gender:
PETSANDGENDER

Gender	Pet ownership status: Non-pet owners	Dog owners	Cat owners
Female	1024	157	85
Male	915	171	82

Analyze the data. Include numerical and graphical summaries as well as a significance test. Summarize your results and conclusions.

9.45 Changing majors. A task force set up to examine retention of students in the majors that they chose when starting college examined data on transfers to other majors.[35] Here are some data giving counts of students classified by initial major and the area that they transferred to:

Initial major	Area transferred to: Engineering	Management	Liberal arts	Other	Total
Biology	13	25	158		398
Chemistry	16	15	19		114
Mathematics	3	11	20		72
Physics	9	5	14		61

Complete the table by computing the values for the "Other" column. Write a short paragraph explaining what conclusions you can draw about the relationship between initial major and area transferred to. Be sure to include numerical and graphical summaries as well as the details of your significance test.

9.46 Cracks in veneer. Many furniture pieces are built with veneer, a thin layer of fine wood that is fastened to less expensive wood products underneath. Face checks are cracks that sometimes develop in the veneer. When face checks appear, the furniture needs to be reconstructed. Because this is a fairly expensive process, researchers seek ways to minimize the occurrence of face checks by controlling the manufacturing process. In one study, the type of adhesive used was one of the factors examined.[36] Because of the way that the veneer is cut, it has two different sides, called loose and tight, either of which can face out. Here is a table giving the numbers of veneer panels with and without face checks for two different adhesives, PVA and UF. Separate columns are given for the loose side and the tight side.
VENEER

Adhesive	Loose side: Face checks: No	Yes	Tight side: Face checks: No	Yes
PVA	10	54	44	20
UF	21	43	37	27

Analyze the data. Write a summary of your results concerning the relationship between the adhesive and the occurrence of face checks. Be sure to include numerical and graphical summaries as well as the details of your significance tests.

9.47 Are Mexican Americans less likely to be selected as jurors? Refer to Exercise 8.99 (page 509) concerning *Castaneda v. Partida*, the case where the Supreme Court review used the phrase "two or three standard deviations" as a criterion for statistical significance. Recall that there were 181,535 persons eligible for jury duty, of whom 143,611 were Mexican Americans. Of the 870 people selected for jury duty, 339 were Mexican Americans. We are interested in finding out if there is an association between being a Mexican American and being selected as a juror. Formulate this problem using a two-way table of counts. Construct the 2×2 table using the variables Mexican American or not and juror or not. Find the X^2 statistic and its P-value. Square the z statistic that you obtained in Exercise 8.81 and verify that the result is equal to the X^2 statistic.

9.48 Evaluation of an herbal remedy. A study designed to evaluate the effects of the herbal remedy *Echinacea purpurea* randomly assigned healthy children who were 2 to 11 years old to receive either *echinacea* or a placebo.[37] Each time a child had an upper respiratory infection (URI), treatment with *echinacea* or the placebo was given for the duration of the URI. The dose for the *echinacea* was based on the age of the child according to the recommendation of the manufacturer. The *echinacea* children had 337 URIs, while the placebo children had 370 URIs. For each URI many variables were measured. One of these was the parental assessment of the illness severity. Here are the data:
HERBALASSESS

Parental assessment	Group: *Echinacea*	Placebo
Mild	153	170
Moderate	128	157
Severe	48	40

They also recorded the presence or absence of various types of *adverse events*. Here is a summary:
HERBALEVENTS

Adverse event	Group: *Echinacea*	Placebo
Itchiness	13	7
Rash	24	10
"Hyper" behavior	30	23
Diarrhea	38	34
Vomiting	22	21
Headache	33	24
Stomachache	52	41
Drowsiness	63	48
Other	63	48
Any adverse event	152	146

(a) Analyze the parental assessment data. Write a summary of your analysis and conclusion. Be sure to include graphical and numerical summaries.

(b) Analyze each adverse event. Display the results graphically in a single graph. Make a table of the relevant descriptive statistics.

(c) Use a statistical significance test to compare the *echinacea* URIs with those of the placebo URIs for each type of adverse event. Summarize the results in a table and write a short report giving your conclusions about the effect of *echinacea* on URIs in healthy children who are 2 to 11 years old. Explain why you need to analyze each type of adverse event separately rather than performing a chi-square test on the preceding 10×2 table.

(d) One concern about analyzing several outcome variables in situations like this is that we may be able to find statistical significance by chance if we look at a sufficiently large number of outcome variables. Explain why this is a concern in general but is not a concern that is important for the interpretation of the results of your analysis here.

(e) The authors of the paper describing these results note that the unit of analysis for their computations is the URI and not the child. They state that similar results were found using more sophisticated statistical methods. Based on the descriptive statistics that you have computed, are you inclined to agree or to disagree with this state[illegible] the authors? Explain

(f) This study was published in the *Journal of the American Medical Association* (*JAMA*) and was criticized in an article that appeared in *Alternative & Complementary Therapies*.[38] Three herbalists gave responses to the original article. Among their criticisms were (*i*) the dose of *echinacea* was too low; (*ii*) the treatment should have been given before the URI, not at the onset of symptoms; (*iii*) we should be skeptical of any positive trials on pharmaceuticals or negative trials on natural remedies that are published in *JAMA*; and (*iv*) *Echinacea angustifolia* (not *E. purpurea*) should have been used in combination with other herbs. Discuss these criticisms and write a summary of your opinions regarding *echinacea*.

The following exercises concern the optional material on goodness of fit discussed in Section 9.3.

9.49 Use of academic assistance services. The 2005 National Survey of Student Engagement reported on the use of campus services during the first year of college.[39] In terms of academic assistance (for example tutoring, writing lab), 43% never used the services, 35% sometimes used the services, 15% often used the services, and 7% very often used the services. You decide to see if your large university has this same distribution. You survey first-year students and obtain the counts 79, 83, 36, and 12, respectively. Use a goodness of fit test to examine how well your university reflects the national average.

9.50 Goodness of fit to a standard Normal distribution. Computer software generated 500 random numbers that should look as if they are from the standard Normal distribution. They are categorized into five groups: (1) less than or equal to -0.6; (2) greater than -0.6 and less than or equal to -0.1; (3) greater than -0.1 and less than or equal to 0.1; (4) greater than 0.1 and less than or equal to 0.6; and (5) greater than 0.6. The counts in the five groups are 139, 102, 41, 78, and 140, respectively. Find the probabilities for these five intervals using Table A. Then compute the expected number for each interval for a sample of 500. Finally, perform the goodness of fit test and summarize your results.

9.51 More on the goodness of fit to a standard Normal distribution. Refer to the previous exercise. Use software to generate your own sample of 500 standard Normal random variables, and perform the goodness of fit test. Choose a different set of intervals from the ones used in the previous exercise.

9.52 Goodness of fit to the uniform distr[illegible] Computer soft[illegible]