From Table 2.9, the estimated conditional odds ratio between defendant's race and the death penalty in the first partial table, for which victims' race is white, equals $\hat{\theta}_{XY(1)} = (53 \times 37)/(414 \times 11) = 0.43$. The sample odds for white defendants receiving the death penalty were 43% of the sample odds for black defendants. In the second partial table, for which victim's race is black, $\hat{\theta}_{XY(2)} = (0 \times 139)/(16 \times 4) = 0.0$, because the death penalty was never given to white defendants having black victims.

The conditional odds ratios can be quite different from the marginal odds ratio. The marginal odds ratio for defendant's race and the death penalty uses the 2×2 marginal table in Table 2.9, ignoring instead of controlling victims' race. The estimate equals $(53 \times 176)/(430 \times 15) = 1.45$. The sample odds of the death penalty were 45% higher for white defendants than for black defendants. Yet, we just observed that those odds were smaller for a white defendant than for a black defendant, for each victims' race. This reversal in the association when we control for victims' race illustrates Simpson's paradox.

If the population has X and Y independent in each partial table, then X and Y are said to be *conditionally independent, given* Z. All conditional odds ratios between X and Y then equal 1. Conditional independence of X and Y, given Z, does not imply marginal independence of X and Y. That is, when odds ratios between X and Y equal 1 at each category of Z, the marginal odds ratio may differ from 1. An association can exist between two variables but completely disappear when we adjust for another variable. Exercise 2.28 shows an example.

2.7.5 Homogeneous Association

Two binary variables X and Y satisfy *homogeneous association* when

$$\theta_{XY(1)} = \theta_{XY(2)} = \cdots,$$

that is, when all $\{\theta_{XY(k)}\}$ are identical. Conditional independence of X and Y is the special case in which each conditional odds ratio equals 1.0.

For X = smoking (yes, no), Y = lung cancer (yes, no), and Z = age (<45, 45-65, >65), suppose $\theta_{XY(1)} = 1.2$, $\theta_{XY(2)} = 4.8$, and $\theta_{XY(3)} = 12.2$. Then, smoking has a weak effect on lung cancer for young people, but the effect strengthens considerably with age. The *XY* association is not homogeneous, because the *XY* conditional odds ratio changes across levels of Z, but if Z is gender and $\theta_{XY(1)} = \theta_{XY(2)} = 4.0$, then homogeneous association occurs.

In a three-way table, homogeneous XY association means that any conditional odds ratio formed using two categories of X and two categories of Y is the same at each category of Z. Inference about associations in multi-way contingency tables is best handled in the context of models. Sections 4.3 and 7.1 introduce models that have the property of homogeneous association. We will see there and in Section 5.2 how to use the data to judge whether conditional independence or homogeneous association are plausible.

EXERCISES

2.1 The PSA blood test is designed to detect prostate cancer. Suppose that of men who have this disease, the test fails to detect prostate cancer in 1 in 4, and of men who do not have it, 1 in 10 have positive test results (so-called false-positive results).

Let C ($\bar{C}$) denote the event of having (not having) prostate cancer and let $+$ ($-$) denote a positive (negative) test result.

a. Which is true: $P(- \mid C) = 1/4$ or $P(C \mid -) = 1/4$? $P(\bar{C} \mid +) = 1/10$ or $P(+ \mid \bar{C}) = 1/10$?

b. Find the sensitivity and specificity of this test.

c. Of men who take the PSA test, suppose $P(C) = 0.04$. Find the cell probabilities in the 2×2 table for the joint distribution that cross-classifies Y = diagnosis with X = true disease status.

d. Using (**c**), find the marginal distribution for the diagnosis and show that $P(C \mid +) = 0.238$. (In fact, the National Cancer Institute estimates that only about 25% of men who have a slightly elevated PSA level, 4–10 ng/mL, actually have prostate cancer.[18])

2.2 For diagnostic testing, let X = true status (1 = disease, 2 = no disease) and Y = diagnosis (1 = positive, 2 = negative). Let $\pi_1 = P(Y = 1 \mid X = 1)$ and $\pi_2 = P(Y = 1 \mid X = 2)$. Let γ denote the probability that a subject has the disease.

a. Given that the diagnosis is positive, use Bayes' Theorem to show that the probability a subject truly has the disease is

$$P(X = 1 \mid Y = 1) = \pi_1\gamma/[\pi_1\gamma + \pi_2(1 - \gamma)].$$

b. For mammograms for detecting breast cancer, suppose $\gamma = 0.01$, sensitivity $= \pi_1 = 0.86$, and specificity $1 - \pi_2 = 0.88$. Find the positive predictive value.

c. To better understand the answer in (**b**), find the joint probabilities for the 2×2 cross-classification of X and Y. Discuss their relative sizes in the two cells that refer to a positive test result.

2.3 According to recent UN figures, the annual gun homicide rate is 62.4 per one million residents in the US and 1.3 per one million residents in Britain. Compare the proportion of residents killed annually by guns using the (**a**) difference of proportions, (**b**) relative risk. Which measure is more useful for describing the strength of association? Why?

2.4 The opening 2018 World Cup odds against being the winning team specified by `espn.com` were $9/2$ for Germany, $5/1$ for Brazil, $11/2$ for France, $20/1$ for England, and $7/1$ for Spain. Find the corresponding prior probabilities of winning for these five teams.

2.5 Consider the following two studies reported in the *New York Times*:

a. A British study reported that of smokers who get lung cancer, "women were 1.7 times more vulnerable than men to get small-cell lung cancer." Is 1.7 an odds ratio or a relative risk?

b. A National Cancer Institute study about tamoxifen and breast cancer reported that the women taking the drug were 45% less likely to experience invasive breast cancer compared to the women taking placebo. Find the relative risk for (i) those taking the drug compared to those taking placebo, (ii) those taking placebo compared to those taking the drug.

[18] See `www.cancer.gov/types/prostate/psa-fact-sheet`.

2.6 Finding and interpreting measures comparing proportions:

a. An observational study[19] of patients hospitalized for gunshot wounds in the US between 2004 and 2013 classified the intent of the gunshot using categories (assault, unintentional, suicide, undetermined, law enforcement, terrorism), with counts (179,793, 65,502, 24,624, 17,401, 5266, 10). Compare proportions for the categories *unintentional* and *terrorism* by a difference and by a ratio. Interpret.

b. According to a 2015 study by the Pew Research Center (`www.people-press.org`), the percentage of Americans who favor allowing gays and lesbians to marry legally was 81% for Democrats who identified themselves as liberal and 22% for Republicans who identified themselves as conservative. Identify the two variables and find the odds ratio between them.

2.7 For adults who sailed on the Titanic on its fateful voyage, the odds ratio[20] between gender (female, male) and survival (yes, no) was 11.4.

a. What is wrong with the interpretation, "The probability of survival for females was 11.4 times that for males?" Give the correct interpretation.

b. The odds of survival for females equaled 2.9. For each gender, find the proportion who survived. Find the value of RR in the interpretation, "The probability of survival for females was RR times that for males."

2.8 A research study estimated that under a certain condition, the probability a subject would be referred for heart catheterization was 0.906 for whites and 0.847 for blacks.

a. A press release about the study stated that the odds of referral for cardiac catheterization for blacks are 60% of the odds for whites. Explain how they obtained 60% (more accurately, 57%).

b. An Associated Press story[21] that described the study stated "Doctors were only 60% as likely to order cardiac catheterization for blacks as for whites." What is wrong with this interpretation? Give the correct percentage for this interpretation. (In stating results to the general public, it is better to use the relative risk than the odds ratio. It is simpler to understand and less likely to be misinterpreted.)

2.9 A 20-year study of British male physicians[22] noted that the proportion who died from lung cancer was 0.00140 per year for cigarette smokers and 0.00010 per year for non-smokers. The proportion who died from heart disease was 0.00669 for smokers and 0.00413 for nonsmokers. Describe the association of smoking with lung cancer and with heart disease, using the difference of proportions and the odds ratio. Interpret. Which response (lung cancer or heart disease) is more strongly related to cigarette smoking, in terms of the reduction in deaths that could occur with an absence of smoking?

2.10 Table 2.10 shows fatality results for drivers and passengers in auto accidents in Florida in 2015, according to whether the person was wearing a shoulder and lap belt restraint versus not using one. Find and interpret the odds ratio.

2.11 Table 2.11 cross-classifies votes in the 2008 and 2012 US Presidential elections. Estimate and find a 95% confidence interval for the population odds ratio. Interpret.

[19] By A. Cook, T. Osler, D. Hosmer, et al., *Injury* **48**: 621–627 (2017).
[20] For data, see R. Dawson, *J. Statist. Educ.* **3** (1995).
[21] For details, see *N. Engl. J. Medic.* **341**: 279–283 (1999).
[22] By R. Doll and R. Peto, *British Med. J.* **2**: 1525–1536 (1976).

Table 2.10 Data for exercise 2.10 on auto accidents.

	Injury		
Restraint Use	Fatal	Nonfatal	Total
No	433	8049	8482
Yes	570	554,883	555,453

Source: Florida Department of Highway Safety and Motor Vehicles.

Table 2.11 Data on presidential votes, for exercise 2.11.

	Vote in 2012	
Vote in 2008	Obama	Romney
Obama	802	53
McCain	34	494

Source: 2014 General Social Survey.

2.12 Data posted at the FBI website (`www.fbi.gov`) indicated that of all blacks slain in 2015, 92% were slain by blacks, and of all whites slain in 2015, 93% were slain by whites. Let Y denote race of victim and X denote race of murderer. Which conditional distribution do these statistics refer to, Y given X or X given Y? Find and interpret the odds ratio.

2.13 Refer to Table 2.1 about belief in an afterlife. Conduct a test of statistical independence. Report the P-value and interpret.

2.14 A poll by Louis Harris and Associates of 1249 adult Americans indicated that 36% believe in ghosts and 37% believe in astrology. Can you compare the proportions using inferential methods for independent binomial samples? If yes, do so. If not, explain why not.

2.15 An article[23] summarized results from the Nurses' Health Study and the Health Professionals Follow-Up Study. The article reported (with RR = relative risk) that "Compared with nonregular use, regular aspirin use was associated with lower risk of overall cancer (RR 0.97; 95% CI 0.94, 0.99), which was primarily due to a lower incidence of gastrointestinal cancers, especially colorectal cancers (RR 0.81; 95% CI 0.75, 0.88)."

a. Identify the response variables and the explanatory variable for these two results. Explain how to interpret the confidence interval about colorectal cancers.

b. Would the association with overall cancer be considered (i) significant or nonsignificant? (ii) strong or weak? Explain.

2.16 Table 2.12 shows data from a General Social Survey cross-classifying a person's perceived happiness with their family income. The table displays the observed and expected cell counts and the standardized residuals for testing independence.

a. For testing independence, $X^2 = 73.4$. Report the *df* value and the P-value, and interpret.

b. Interpret the standardized residuals in the corner cells having (i) counts 21 and 83, (ii) counts 110 and 94.

[23] By Y. Cao et al., *JAMA Oncology* **2**: 762–769 (2016).

Table 2.12 Data for Exercise 2.16, with estimated expected frequencies and standardized residuals.

	Happiness		
Income	Not Too Happy	Pretty Happy	Very Happy
Above average	21 (35.8) −2.973	159 (166.1) −0.947	110 (88.1) 3.144
Average	53 (79.7) −4.403	372 (370.0) 0.224	221 (196.4) 2.907
Below average	94 (52.5) 7.368	249 (244.0) 0.595	83 (129.5) −5.907

2.17 Table 2.13 is based on data from the 2016 General Social Survey.

a. Test the null hypothesis of independence between political party identification and race. Interpret.

b. Use standardized residuals to describe the evidence.

c. Partition chi-squared into two components, the first of which compares the races on the (Democrat, Republican) choice. Interpret the quite different results for the two cases.

Table 2.13 Data for Exercise 2.17.

	Political Party Identification		
Race	Democrat	Republican	Independent
White	871	821	336
Black	347	42	83

2.18 Each subject in a sample of 100 men and 100 women is asked to indicate which of the following factors (one or more) are responsible for increases in teenage crime: A – the increasing gap in income between the rich and poor, B – the increase in the percentage of single-parent families, C – insufficient time that parents spend with their children. A cross-classification of the responses by gender is

```
-----------------------
Gender   A    B    C
Men      60   81   75
Women    75   87   86
-----------------------
```

Is it valid to apply the chi-squared test of independence to this table? Why or why not? Explain how this table actually provides information needed to cross-classify gender with each of three variables. Construct the contingency table relating gender to opinion about whether factor A is responsible for increases in teenage crime.

2.19 Table 2.14 is from a recent General Social Survey. For these data, $X^2 = 69.2$. Write a short report summarizing inference. In your report, mention an alternative test of independence that is relevant for these data.

Table 2.14 Table for Exercise 2.19, with standardized residuals.

	Religious Beliefs		
Highest Degree	Fundamentalist	Moderate	Liberal
Less that High School	178 (4.5)	138 (−2.6)	108 (−1.9)
High School or Junior College	570 (2.6)	648 (1.3)	442 (−4.0)
Bachelor or Graduate	138 (−6.8)	252 (0.7)	252 (6.3)

2.20 Formula (2.3) has alternative formula $X^2 = n\sum(\hat{\pi}_{ij} - \hat{\pi}_{i+}\hat{\pi}_{+j})^2/\hat{\pi}_{i+}\hat{\pi}_{+j}$. Explain why, for particular $\{\hat{\pi}_{ij}\}$, X^2 is large when n is sufficiently large, regardless of whether the association is practically important. Hence, chi-squared tests merely indicate the degree of evidence against a hypothesis and do not describe the strength of association.

2.21 A GSS that cross-classified income in thousands of dollars (<5, 5–15, 15–25, >25) by job satisfaction (very dissatisfied, a little satisfied, moderately satisfied, very satisfied) for black Americans produced a 4×4 table having counts, by row, (2, 4, 13, 3 / 2, 6, 22, 4 / 0, 1, 15, 8 / 0, 3, 13, 8).

a. Test independence of job satisfaction and income using X^2. Interpret and explain the deficiency of this test for these data. Find the standardized residuals. Do they suggest any association pattern?

b. Conduct a test that treats the variables in a quantitative manner, using scores (3, 10, 20, 35) for income and (1, 3, 4, 5) for job satisfaction. Explain why results differ so much from part (**a**).

2.22 A study (B. Kristensen et al., *J. Intern. Med.* **232**: 237–245 (1992)) considered the effect of prednisolone on severe hypercalcaemia in women with metastatic breast cancer. Of 30 patients, 15 were randomly selected to receive prednisolone and the other 15 formed a control group. Normalization in their level of serum-ionized calcium was achieved by 7 of the 15 prednisolone-treated patients and by 0 of the 15 patients in the control group. Use Fisher's exact test to find a P-value for testing whether results were significantly better for treatment than control. Interpret.

2.23 Table 2.15 contains results of a study comparing radiation therapy with surgery in treating cancer of the larynx. Some R output follows:

```
---------------------------------------------------------------------------
> fisher.test(matrix(c(21,2,15,3),ncol=2,byrow=TRUE),alternative="two.sided")
p-value = 0.6384
> fisher.test(matrix(c(21,2,15,3),ncol=2,byrow=TRUE),alternative="greater")
p-value = 0.3808
> fisher.test(matrix(c(21,2,15,3),ncol=2,byrow=TRUE),alternative="less")
p-value = 0.8947
---------------------------------------------------------------------------
```

Table 2.15 Data for Exercise 2.23.

	Cancer Controlled	Cancer Not Controlled
Surgery	21	2
Radiation therapy	15	3

Source: W. Mendenhall et al., *Int. J. Radiat. Oncol. Biol. Phys.* **10**: 357–363 (1984), with permission from Elsevier Science Ltd.

a. Report and interpret the P-value for Fisher's exact test with (i) H_a: $\theta > 1$, and (ii) H_a: $\theta \neq 1$.

b. Obtain and interpret the mid P-value for H_a: $\theta \neq 1$ and find the corresponding confidence interval based on mid P-values. Give advantages of this type of P-value, compared to the ordinary one.

2.24 a. At each age level, the death rate is higher in South Carolina than in Maine, but overall the death rate is higher in Maine.[24] Explain how this could be possible.

b. Smith and Jones are baseball players. Smith had a higher batting average than Jones in 2005 and 2006. Is it possible that for the combined data for these two years, Jones had the higher batting average? Explain, and illustrate using data.

2.25 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.[25]

a. Exhibit the data as a three-way contingency table. Construct the partial tables needed to study the conditional association between defendant's race and the death penalty verdict. Find and interpret the sample conditional odds ratios.

b. Find and interpret the sample marginal odds ratio between defendant's race and the death penalty verdict. Do these data exhibit Simpson's paradox? Explain its cause.

2.26 Give an example of three variables X, Y, and Z, for which you expect X and Y to be marginally associated but conditionally independent, controlling for Z.

2.27 Based on murder rates in the United States, the Associated Press reported that the probability a newborn child has of eventually being a murder victim is 0.0263 for nonwhite males, 0.0049 for white males, 0.0072 for nonwhite females, and 0.0023 for white females. Find the conditional odds ratios between race and whether a murder victim, given gender. Interpret.

2.28 The expected frequencies in Table 2.16 show a hypothetical relationship among three variables: Y = response, X = drug treatment, and Z = clinic. Show that X and

Table 2.16 Expected frequencies illustrating that conditional independence does not imply marginal independence.

Clinic	Drug Treatment	Response: Success	Response: Failure
1	A	18	12
	B	12	8
2	A	2	8
	B	8	32

[24] For data, see H. Wainer, *Chance* **12**: 44 (1999).
[25] M. Radelet, *Amer. Sociol. Rev.* **46**: 918–927 (1981).

Y are conditionally independent, given Z, but marginally associated. Explain how the marginal XY association can be so different from its conditional association, using the values of the conditional XZ and YZ odds ratios. Explain why it would be misleading to study only the marginal table and conclude that successes are more likely with treatment A than with treatment B.

2.29 Refer to Table 2.1 about belief in an afterlife.

a. Treating the data as independent binomial samples and using beta(0.5, 0.5) prior distributions, find the posterior mean estimates for the probabilities of believing in an afterlife.

b. Find and interpret (i) 95% posterior intervals for the difference of proportions and the odds ratio; (ii) the posterior probability that belief in an afterlife is more probable for women than for men.

2.30 True or false?

a. A 95% confidence interval for the odds ratio between MI (yes, no) and treatment (placebo, aspirin) is (1.44, 2.33). If we form the table with aspirin in the first row (instead of placebo), the confidence interval is $(1/2.33, 1/1.44) = (0.43, 0.69)$.

b. A survey of college students analyzes the association between opinion about whether it should be legal to (1) use marijuana, (2) drink alcohol if you are 18 years old. We may get a different odds ratio value if we treat marijuana use as the response variable than if we treat alcohol use as the response variable.

c. Interchanging two rows or interchanging two columns in a contingency table has no effect on the value of the X^2 or G^2 chi-squared statistics. Thus, these tests treat both the rows and the columns of the contingency table as nominal scale, and if either or both variables are ordinal, the test ignores that information.

d. Suppose that income (high, low) and gender are conditionally independent, given the type of job (secretarial, construction, service, professional, ...). Then, income and gender are also independent in the 2×2 marginal table.

e. According to the Pew Research Center (`www.people-press.org`), when adults in the US were asked in 2010 whether there is solid evidence that the average temperature on Earth has been getting warmer over the past few decades, the estimated odds of a *yes* response for a Democrat was 2.96 times higher than for an Independent, and it was 2.08 times higher for an Independent than for a Republican. The estimated odds ratio between opinion on global warming and whether one is a Democrat or a Republican equals $2.96 \times 2.08 = 6.2$.