Chapter 6 – Two-Way Tables

6.1 (a) This table describes 736 + 450 + 193 + 205 + 144 + 80 = 1808 people, and 736 + 450 + 193 = 1379 have played video games. **(b)** (736 + 205)/1808 = 0.5205 = 52.05%. We do this for all three grade levels. The complete marginal distribution for grades is:

Grade	Percent
A's and B's	52.05%
C's	32.85%
D's and F's	15.10%

Of all boys, 32.85% + 15.10% = 47.95% received a grade of C or lower.

6.2 (a) The sum of all entries is 19,174. This table describes 19,174,000 (19,174 thousand) college students. **(b)** The marginal distribution for ages of undergraduates follows:

Age group	Percent
15 to 19 years	22.27%
20 to 24 years	42.22%
25 to 34 years	20.80%
35 years or older	14.71%

There are 4198 thousand + 3897 thousand = 8095 thousand undergraduates in the 20–24 age group. This is 8095/19,174 = 0.4222 = 42.22% of all undergraduates.

6.3 There are 736 + 450 + 193 = 1379 players. Of these, 736/1379 = 53.37% earned A's or B's. Similarly, there are 205 + 144 + 80 = 429 nonplayers. Of these, 205/429 = 47.79% earned A’s or B’s. Continuing, the conditional distributions of grades follow:

Grades	Players	Nonplayers
A's and B's	53.37%	47.79%
C's	32.63%	33.57%
D's and F's	14.00%	18.65%

If anything, players have slightly higher grades than nonplayers, but this could be due to chance.

6.4 PLAN: Starting with the two-way table from Exercise 6.2, find and compare the conditional distributions of sex for each age group. SOLVE: For example, for the 25 to 34 years age group, the proportion of women is 2197/(2197 + 1791) = 0.5509, or 55.09%. CONCLUDE: The data support our suspicion—the percent of women in the 25 to 34 years age group is, indeed, larger than the percent of women in the 20 to 24 years age group.

Age group	Percent
15 to 19 years	55.40%
20 to 24 years	51.86%
25 to 34 years	55.09%
35 years or older	63.26%

6.5 Two examples are shown. In general, choose *a* to be any number from 10 to 50, and then all the other entries can be determined.

30	20
30	20

50	0
10	40

6.6 (a) Seth made (5 + 3)/(5 + 4 + 3 + 3) = 0.533, or 53.3% of his field goal attempts. Roberto made 63 of 117 (53.8%) of his field goal attempts. **(b)** The table below describes the percent of field goals made for each type of field goal for each player. For example, Seth made 5/9 = 0.5556, or 55.56%, of two-point attempts.

	Seth	Roberto
Two-pointers	55.56%	55.36%
Three-pointers	50.00%	20.00%

(c) Roberto's overall field goal percent is higher than Seth's, but Seth's percent is higher for both types of field goals. This is an example of Simpson's paradox—the comparison that holds for both field goal groups is reversed when the groups are combined into one group. Notice that Roberto took an extremely large number of two-point shots (especially as compared to Seth).

6.7 (a) For Rotorua district, 79/8889 = 0.0089, or 0.9%, of Maori are in the jury pool, while 258/24,009 = 0.0107, or 1.07%, of non-Maori are in the jury pool. For Nelson district, the corresponding percents are 0.08% for Maori and 0.17% for non-Maori. Hence, in each district, the percent of non-Maori in the jury pool exceeds the percent of Maori in the jury pool. **(b)** Combining the regions into one table:

	Maori	Non-Maori
In jury pool	80	314
Not in jury pool	10,138	56,353
Total	10,218	56,667

For Maori, the overall percent in the jury pool is 80/10,218 = 0.0078, or 0.78%, while for non-Maori, the overall percent in the jury pool is 314/56,667 = 0.0055, or 0.55%. Hence, overall Maori have a larger percent in the jury pool, but in each region, they have a lower percent in the jury pool. **(c)** The reason for Simpson's paradox occurring with this example is that Maori constitute a large proportion of Rotorua's population, while in Nelson they are a small minority community.

6.8 (b) 1918. (228 + 112 + 281 + 281 + 106 + 481 + 21 + 408 = 1918.)

6.9 (b) 340. (228 + 112 = 340.)

6.10 (a) about 18%. (340/1918 = 0.177, or 17.7%.)

6.11 (a) the marginal distribution of age.

6.12 (b) about 36%. (228/636 = 0.358, or 35.8%.)

6.13 (c) the conditional distribution of age among those who use social networking on their phone.

6.14 (c) about 67%. (228/340 = 0.671, or 67.1%.)

6.15 (b) the conditional distribution of whether one uses social networking on his/her phone among individuals age 18–29.

6.16 (b) four bars.

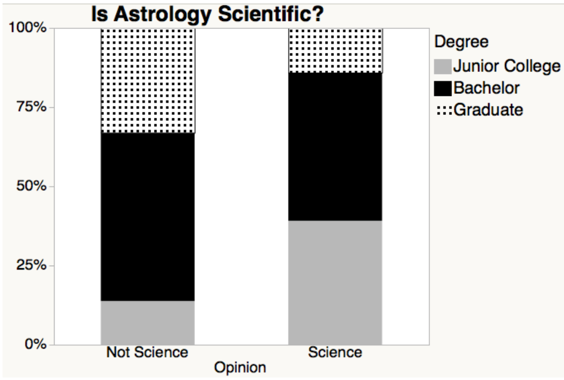
6.17 (b) an example of Simpson's paradox: full-time students do better in both kinds of courses but worse overall because they take more science courses.

6.18 The two distributions are given below.

	Not at all scientific	Very or sort of scientific
Junior college	13.78%	39.13%
Bachelor	53.08%	46.74%
Graduate	33.14%	14.13%

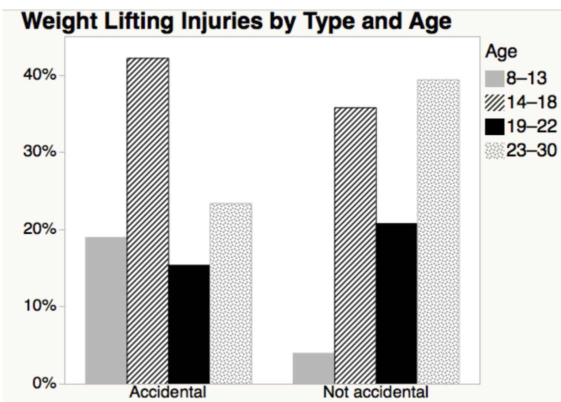
For example, in the distribution of people who feel astrology is not at all scientific, the percent with junior college degrees is $\frac{47}{47 + 181 + 113} = 0.1378$, or 13.78%.

The table and bar chart reveal that, loosely, adults who believe astrology is not at all scientific tend to have relatively more college education than adults who believe astrology is very or sort of scientific.



6.19 For each type of injury (accidental, not accidental), the distribution of ages is produced below.

	Accidental	Not accidental
8–13	19.0%	4.0%
14–18	42.2%	35.8%
19–22	15.4%	20.8%
23–30	23.4%	39.4%



We see that, among accidental weight-lifting injuries, the percent of relatively younger lifters is larger. Among the injuries that are not accidental, the percent of relatively older lifters is larger.

6.20 The tables below give the two marginal distributions. The marginal distribution of marital status is found by taking, for example, 36,590/117,048 = 0.3126, or 31.26%, for the single group. The marginal distribution of income is found by taking, for example, 8843/117,048 = 0.0756, or 7.56%, for no income.

Single	Married	Divorced	Widowed
31.26%	56.60%	9.35%	2.79%

No income	$1–$49,999	$50,000–$99,999	$100,000 and over
7.56%	57.63%	22.87%	11.95%

As rounded here, the marital status marginal distribution sums to 100%, but the income marginal distribution sums to 100.01%.

6.21 The percent of single men with no income is 6013/36,590 = 0.1643, or 16.43%. The percent of men with no income who are single is 6013/8843 = 0.6800, or 68.00%.

6.22 Divide the entries in the first column by the first column total. For example, from Exercise 6.21, 16.43% is 0.1643 = 6013/36,590.

Income	None	$1– $49,999	$50,000– $99,999	$100,000 and over
Percent of single men	16.43%	66.63%	12.89%	4.04%

These should add to 100%, except for possible roundoff error (as rounded here, the distribution adds to 99.99%).

6.23 (a) We need to compute percents to account for the fact that the study included many more married men than single men, so we would expect their numbers to be higher in every job grade (even if marital status had no relationship with income). **(b)** A table of percents is provided; descriptions of the relationship may vary. Divorced and widowed men had similar percents between the two income groups studied. Single men had much higher percents of no income; married men had much higher percents of incomes at and over $100,000.

	Single	Married	Divorced	Widowed
No income	68.00%	22.80%	7.94%	1.27%
$100,000 and up	10.57%	80.52%	7.14%	1.77%

6.24 One example would be that men who are married, widowed, or divorced may be more "invested" in their careers than men who are single. There is still a feeling of societal pressure for a man to "provide" for his family.

6.25 (a) The two-way table of race (white, black) versus death penalty (death penalty, no death penalty) follows:

	White defendant	Black defendant
Death penalty	19	17
No death penalty	141	149

(b) For black victims: The percent of white defendants given the death penalty is 0/9 = 0, or 0%. The percent of black defendants given the death penalty is 6/103 = 0.058, or 5.8%.

For white victims: The percent of white defendants given the death penalty is 19/151 = 0.126, or 12.6%. The percent of black defendants given the death penalty is 11/63 = 0.175, or 17.5%. Hence, for both victim races, black defendants are given the death penalty relatively more often than white defendants. However, overall, referring to the table in part (a), 19/160 = 0.119, or 11.9%, of white defendants got the death penalty, while 17/166 = 0.102, or 10.2%, of black defendants got the death penalty. This illustrates Simpson's paradox. **(c)** For white defendants, (19 + 132)/(19 + 132 + 0 + 9) = 0.9438 = 94.4% of victims were white. For black defendants, only (11 + 52)/(11 + 52 + 6 + 97) = 0.3795, or 37.95%, of victims were white. Meanwhile, the death penalty was predominantly assigned to cases involving white victims; a death penalty was assigned to the defendant in 14.0% of all cases with a white victim but in only 5.4% of all cases with a black victim. Hence, because most white defendants' victims are white and cases with white victims carry additional risk of a death penalty, white defendants are being assigned the death penalty more often overall.

6.26 Examples will vary. Here is one very simple possibility. The key is to be sure that the three-way table has a lower percent of overweight people among the smokers than among the nonsmokers.

Smoker	Early death	
	Yes	No
Obese	1	0
Not obese	4	2

Nonsmoker	Early death	
	Yes	No
Obese	3	6
Not obese	1	3

Combined	Early death	
	Yes	No
Obese	4	6
Not obese	5	5

6.27 PLAN: From the given two-way table of results, find and compare the conditional distributions of outcome (success, no success) for each treatment (Chantix, Bupropion, and placebo). SOLVE: The percents for each column are provided in the table. For example, for Chantix, the percent of successes (no smoking in weeks 9–12) is 155/(155 + 197) = 0.4403, or 44.0%. Because we're comparing success rates, we'll leave off the row for those who smoked in weeks 9–12, as this is just 100—the percent who did not smoke in weeks 9–12. CONCLUDE: Clearly, a larger percent of subjects using Chantix were not smoking during weeks 9–12, compared with results for either of the other treatments. In fact, as we'll learn later, this result is statistically significant … random chance doesn't easily explain this difference, and we might conclude that Chantix use increases the chance of success.

	Chantix	Bupropion	Placebo
Percent not smoking in weeks 9–12	44.0%	29.5%	17.7%

6.28 PLAN: From the given two-way table of response by sex, find and compare the conditional distributions of response for men alone and women alone. SOLVE: The table represents the responses of 516 men and 636 women. To find the conditional distributions, divide each entry in the table by its column total. These percents are given in that table; for example, 76/516 = 0.1473, or 14.73%. CONCLUDE: Men are more likely to view animal testing as justified if it might save human lives: Over two-thirds of men agree or strongly agree with this statement, compared to slightly less than half of women. The percent who disagree or strongly disagree tell a similar story: 16% of men versus 30% of women.

Response	Male	Female
Strongly agree	14.7%	9.3%
Agree	52.3%	38.8%
Neither	16.9%	21.9%
Disagree	11.8%	19.3%
Strongly disagree	4.3%	10.7%

6.29 PLAN: Calculate and compare the conditional distributions of sex for each degree level. SOLVE: We compute, for example, the percent of women earning associate's degrees: 646/(646 + 383) = 0.6278, or 62.78%. The table shows the percent of women at each degree level, which is all we need for comparison. CONCLUDE: Women constitute a substantial majority of associate's, bachelor's, and master's degrees, and a small majority of doctor's and professional degrees.

Degree	Female
Associate's	62.78%
Bachelor's	57.88%
Master's	61.94%
Professional or Doctor's	53.54%

6.30 PLAN: Find and compare the conditional distributions of type of complication for each of the three treatments. SOLVE: The table provides the percents of subjects with various complications for each treatment. For example, for subjects with gastric banding, 81/5380 = 0.0151, or 1.5%, had non-life-threatening complications. CONCLUDE: Without question, gastric bypass surgery carries the greatest risk for both non-life-threatening and serious complications. Gastric banding seems to be the safest procedure, with the lowest rates for both types of complications.

	Non-life-threatening	Serious	None
Gastric banding	1.5%	0.9%	97.6%
Sleeve gastrectomy	3.6%	2.2%	94.1%
Gastric bypass	6.7%	3.6%	89.7%

6.31 PLAN: Find and compare the conditional distributions for health (self-reported) for each group (smokers and nonsmokers). SOLVE: The table provides the percent of subjects with various health outlooks for each group. CONCLUDE: Clearly, the outlooks of current smokers are generally bleaker than those of current nonsmokers. Much larger percents of nonsmokers reported being in excellent or very good health, while much larger percents of smokers reported being in fair or poor health.

	Health outlook				
	Excellent	Very good	Good	Fair	Poor
Current smoker	6.2%	28.5%	35.9%	22.3%	7.2%
Current nonsmoker	12.4%	39.9%	33.5%	14.0%	0.3%

6.32 (a) The two-way table is provided.

	Saw shadow	Did not see shadow
Above average temp.	49	10
At or below average temp.	51	6

(b) PLAN: Compare the conditional distributions of actual temperatures for Phil's "forecasts." SOLVE: In the 100 years in which Phil saw his shadow, the temperature was at or below average temperature 51/100 = 51% of the time. In the 16 years in which he did not see his shadow, the temperature was above average 10/16 = 62.5% of the time. CONCLUDE: Phil is slightly better than flipping a coin as a forecaster. He is correct more than half of the time whether seeing his shadow or not, but he actually agrees with the actual weather better in years he does not see his shadow. Unfortunately, he sees his shadow in most years.

6.33 PLAN: Because the numbers of students who use (or do not use) medications are different, we find the conditional distributions of those who do and do not use medications. SOLVE: The table provides the percent of subjects with various levels of sleep quality for each drug-use group. CONCLUDE: Those who use medications are less likely to have optimal sleep quality and more likely to have poor sleep quality than those who do not use medications. This is certainly a case where one would not want to ascribe causation: Do those who use medications to stay awake have poor sleep quality because they use the medications, or do they use the medications to stay awake because they had poor sleep quality before using them?

	Sleep quality		
	Optimal	Borderline	Poor
Use medications	21.3%	30.5%	48.3%
Do not use medications	38.2%	26.7%	35.2%

6.34 to **6.36** are Web-based exercises.