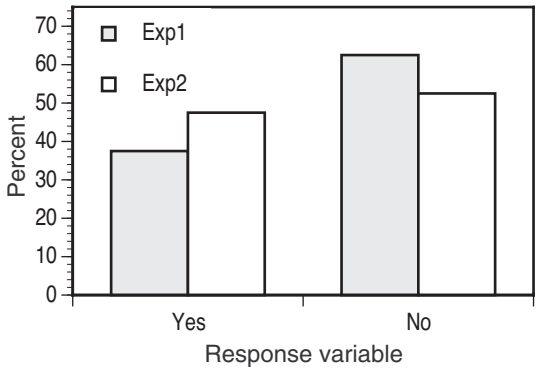
Chapter 9 Solutions

9.1. **(a)** The conditional distributions are given in the table below. For example, given Explanatory $= 1$, the distribution of the response variable is $\frac{75}{200} = 37.5\%$ Yes and $\frac{125}{200} = 62.5\%$ No. **(b)** The graphical display might take the form of a bar graph like the one shown below, but other presentations are possible. **(c)** One notable feature is that when Explanatory $= 1$, "No" is more common, but "Yes" and "No" are nearly evenly split when Explanatory $= 2$.

Response variable	Explanatory variable 1	2
Yes	37.5%	47.5%
No	62.5%	52.5%



9.2. **(a)** The expected cell count for the first cell is $\frac{(170)(200)}{400} = 85$. **(b)** This X^2 statistic has df $= (2-1)(2-1) = 1$. **(c)** Because $3.84 < X^2 < 5.02$, the P-value is between 0.025 and 0.05.

9.3. The table below summarizes the bounds for the P-values, and also gives the exact P-values (given by software). In each case, df $= (r-1)(c-1)$.

	X^2	Size of table	df	Crit. values (Table F)	Bounds for P	Actual P
(a)	2.5	2 by 2	1	$2.07 < X^2 < 2.71$	$0.10 < P < 0.15$	0.1138
(b)	6.5	2 by 2	1	$5.41 < X^2 < 6.63$	$0.01 < P < 0.02$	0.0108
(c)	16.3	3 by 5	8	$15.51 < X^2 < 17.53$	$0.025 < P < 0.05$	0.0382
(d)	16.3	5 by 3	8	$15.51 < X^2 < 17.53$	$0.025 < P < 0.05$	0.0382

9.4. The Minitab output shown on the right gives $X^2 \doteq 54.307$, df $= 1$, and $P < 0.0005$, indicating significant evidence of an association. In Example 8.11, the test statistic was $z \doteq 7.37$; the square of this statistic is $z^2 \doteq 54.3169$, which agrees with X^2 except for rounding.

Minitab output

```
             Men    Women    Total
Yes         1392     1748     3140
         1215.19  1924.81

No          3956     6723    10679
         4132.81  6546.19

Total       5348     8471    13819

ChiSq = 25.726 + 16.241 +
         7.564 +  4.776 = 54.307
df = 1, p = 0.000
```

9.5. The chi-square goodness of fit statistic is $X^2 \doteq 15.2$ with df $= 5$, for which $0.005 < P < 0.01$ (software gives 0.0096). The details of the computation are given in the table below; note that there were 475 M&M's in the bag.

	Expected frequency	Expected count	Observed count	$O - E$	$\frac{(O-E)^2}{E}$
Brown	0.13	61.75	61	−0.75	0.0091
Yellow	0.14	66.5	59	−7.5	0.8459
Red	0.13	61.75	49	−12.75	2.6326
Orange	0.20	95	77	−18	3.4105
Blue	0.24	114	141	27	6.3947
Green	0.16	76	88	12	1.8947
			475		15.1876

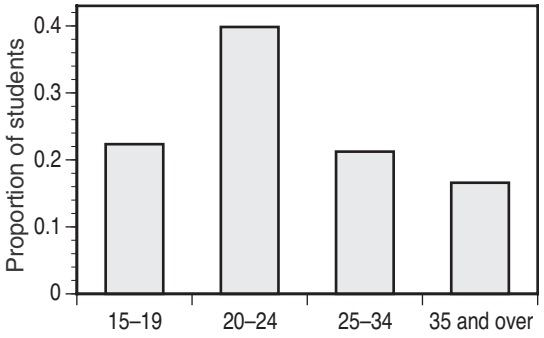
9.6. The main problem is that this is not a two-way table. Specifically, each of the 119 students might fall into several categories: They could appear on more than one row if they saw more than one of the movies and might even appear more than once on a given row (for example, if they have both bedtime and waking symptoms arising from the same movie).

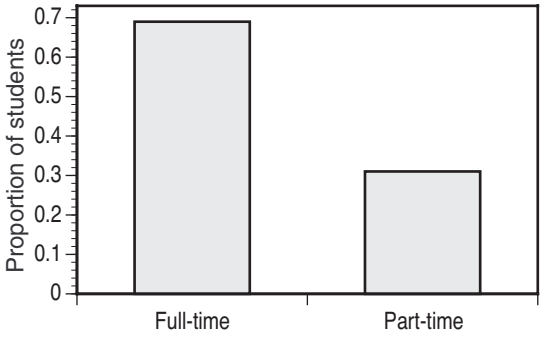
Another potential problem is that this is a table of *percents* rather than *counts*. However, because we were given the value of n for each movie title, we could use that information to determine the counts for each category; for example, it appears that 20 of the 29 students who watched *Poltergeist* had short-term bedtime problems because $\frac{20}{29} \doteq 68.96\%$ (perhaps the reported value of 68% was rounded incorrectly). If we determine all of these counts in this way (and note several more apparent rounding errors in the process), those counts add up to 200, so we see that students really were counted more than once.

If the values of n had not been given for each movie, then we could not do a chi-squared analysis even if this were a two-way table.

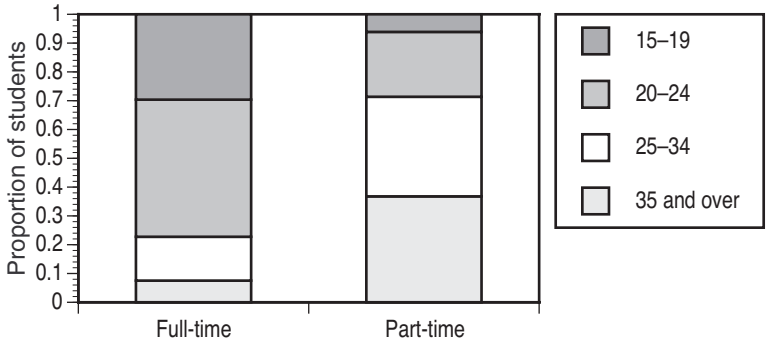
9.7. (a) The joint distribution is found by dividing each number in the table by 17,380 (the total of all the numbers). These proportions are given in italics on the right. For example, $\frac{3553}{17380} \doteq 0.2044$, meaning that about 20.4% of all college students are full-time and aged 15 to 19. **(b)** The marginal distribution of age is found by dividing the *row* totals by 17,380; they are in the right margin of the table and the graph on the left below. For example, $\frac{3882}{17380} \doteq 0.2234$, meaning that about 22.3% of all college students are aged 15 to 19. **(c)** The marginal distribution of status is found by dividing the *column* totals by 17,380; they are in the bottom margin of the table and the graph on the right below. For example, $\frac{11989}{17380} \doteq 0.6898$, meaning that about 69% of all college students are full-time. **(d)** The conditional distributions are given in the table below. For each status category, the conditional distribution of age is found by dividing the counts in that column by that column total. For example, $\frac{3553}{11989} \doteq 0.2964$, $\frac{5710}{11989} \doteq 0.4763$, etc., meaning that of all full-time college students, about 29.64% are aged 15 to 19, 47.63% are 20 to 24, and so on. Note that each set of four numbers should add to 1 (except for rounding error). Graphical presentations may vary; one possibility is shown below. **(e)** We see that full-time students are dominated by younger ages, while part-time students are more likely to be older.

	FT	PT	
15–19	3553	329	3882
	0.2044	*0.0189*	0.2234
20–24	5710	1215	6925
	0.3285	*0.0699*	0.3984
25–34	1825	1864	3689
	0.1050	*0.1072*	0.2123
35+	901	1983	2884
	0.0518	*0.1141*	0.1659
	11989	5391	17380
	0.6898	0.3102	

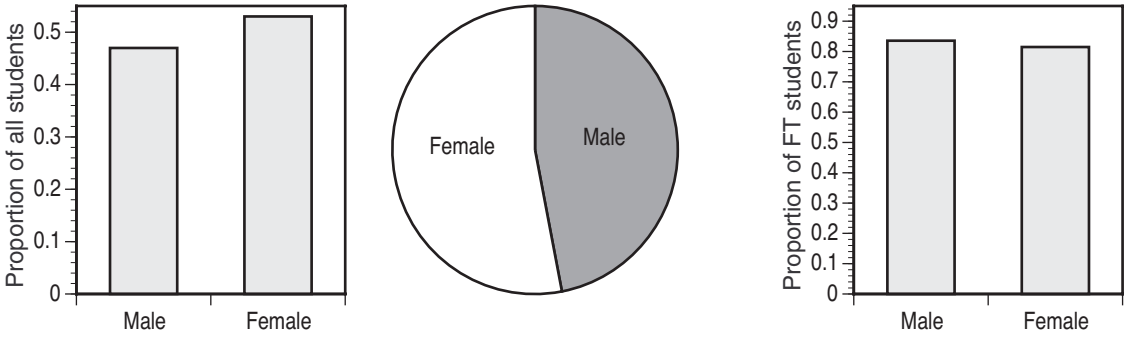




	Full-time	Part-time
15–19	0.2964	0.0610
20–24	0.4763	0.2254
25–34	0.1522	0.3458
35+	0.0752	0.3678



9.8. **(a)** Of all students aged 20 to 24 years, $\frac{3254}{6925} \doteq 46.99\%$ are men and the rest ($\frac{3671}{6925} \doteq 53.01\%$) are women. Shown below are two possible graphical displays. In the bar graph on the left, the bars represent the proportion of all students (in this age range) in each gender. Alternatively, because the two percents represent parts of a single whole, we can display the distribution as a pie chart like that in the middle. **(b)** Among male students, $\frac{2719}{3254} \doteq 83.56\%$ are full-time and the rest ($\frac{535}{3254} \doteq 16.44\%$) are part-time. Among female students, those numbers are $\frac{2991}{3671} \doteq 81.48\%$ and $\frac{680}{3671} \doteq 18.52\%$. Men in this age range are (very slightly) more likely to be full-time students. The bar graph below on the right shows the proportions of full-time students side by side; note that a pie graph would *not* be appropriate for this display because the two proportions represent parts of two different wholes. **(c)** For the full-time row, the expected counts are $\frac{(5710)(3254)}{6925} \doteq 2683.08$ and $\frac{(5710)(3671)}{6925} \doteq 3026.92$. **(d)** Using df $= 1$, we see that $X^2 = 5.17$ falls between 5.02 and 5.41, so $0.02 < P < 0.025$ (software gives 0.023). This is significant evidence (at the 5% level) that there is a difference in the conditional distributions.



9.9. **(a)** Yes, this seems to satisfy the assumptions for test of a two-way table. Specifically, this is the first model (comparing two populations); the populations are "adult Americans asked about Hillary Rodham Clinton," and "adult Americans asked about Hillary Clinton." The original sample was divided in half to create (independent) samples of these two populations. **(b)** This statistic has df $= (r-1)(c-1) = (1)(3) = 3$. **(c)** $X^2 = 4.23$ falls between the first two critical values in Table F, so $0.20 < P < 0.25$. (Software gives $P = 0.2377$.)

9.10. **(a)** The percent who have lasting waking symptoms is the total of the first column divided by the grand total: $\frac{69}{119} \doteq 57.98\%$. **(b)** The percent who have both waking and bedtime symptoms is the count in the upper left divided by the grand total: $\frac{36}{119} \doteq 30.25\%$. **(c)** To test H_0: There is no relationship between waking and bedtime symptoms vs. H_a: There is a relationship, we find $X^2 \doteq 2.275$ (df $= 1$) and $P \doteq 0.132$. We do not have enough evidence to conclude that there is a relationship.

Minitab output

```
        WakeYes    WakeNo     Total
BedYes       36        33        69
          40.01     28.99

 BedNo       33        17        50
          28.99     21.01

 Total       69        50       119

ChiSq =  0.402 +  0.554 +
         0.554 +  0.765 = 2.275
df = 1, p = 0.132
```

9.11. (a) $\frac{14}{24} \doteq 58.33\%$ of desipramine users did not have a relapse, while $\frac{6}{24} = 25\%$ of lithium users and $\frac{4}{24} \doteq 16.67\%$ of those who received placebos succeeded in breaking their addictions. Desipramine seems to be effective. Note that use of percents is not as crucial here as in other cases because each drug was given to 24 addicts. **(b)** Yes; this seems to satisfy the assumptions. Specifically, this is the first model (comparing three populations); the populations are drug users given each of three treatments. **(c)** The test statistic is $X^2 = 10.5$ (df $= 2$), for which $P = 0.005$. We have good evidence that there is a relationship between treatment and relapse.

Minitab output

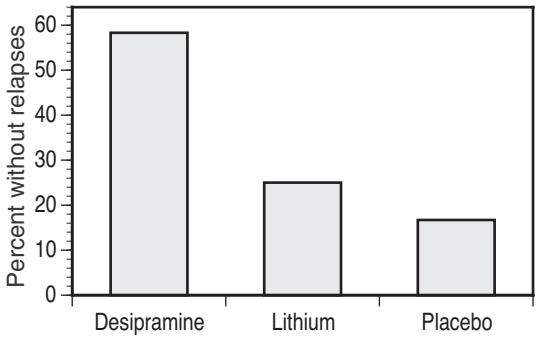
```
             Yes        No     Total
Desip         10        14        24
           16.00      8.00

 Lith         18         6        24
           16.00      8.00

Plcbo         20         4        24
           16.00      8.00

Total         48        24        72

ChiSq =  2.250 +  4.500 +
         0.250 +  0.500 +
         1.000 +  2.000 = 10.500
df = 2, p = 0.005
```



9.12. The table below gives df $= (r-1)(c-1)$, bounds for P, and software P-values.

	X^2	Size of table	df	Crit. values (Table F)	Bounds for P	Software P
(a)	1.25	2 by 2	1	$X^2 < 1.32$	$P > 0.25$	0.2636
(b)	18.34	4 by 4	9	$16.92 < X^2 < 19.02$	$0.025 < P < 0.05$	0.0314
(c)	24.21	2 by 8	7	$22.04 < X^2 < 24.32$	$0.001 < P < 0.0025$	0.0010
(d)	12.17	5 by 3	8	$12.03 < X^2 < 13.36$	$0.10 < P < 0.15$	0.1438

9.13. Two examples are shown on the right. In general, choose a to be any number from 0 to 50, and then all the other entries can be determined.

30	20
70	80

10	40
90	60

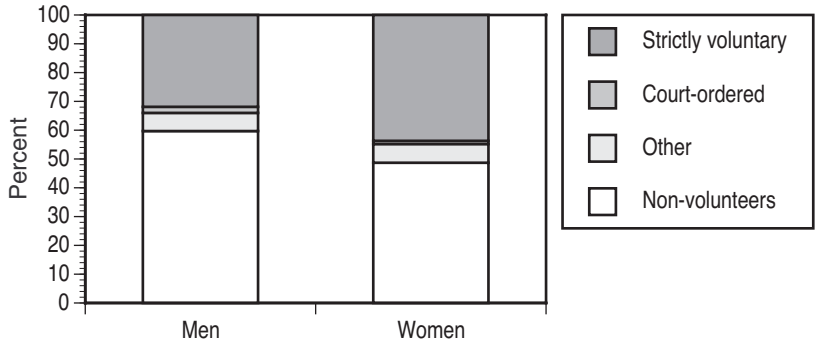
Note: *This is why we say that such a table has "one degree of freedom": We can make one (nearly) arbitrary choice for the first number, and then have no more decisions to make.*

9.14. To construct such a table, we can start by choosing values for the row and column sums r_1, r_2, r_3, c_1, c_2, as well as the grand total N. Note that the $N = r_1 + r_2 + r_3 = c_1 + c_2$, so we only have four choices to make. Then find each count a, b, c, d, e, f by taking the corresponding *row* total, times the corresponding *column* total, divided by the *grand* total. For example, $a = r_1 \times c_1/N$ and $d = r_2 \times c_2/N$. Of course, these counts should be whole numbers, so it may be necessary to make adjustments in the row and column totals to meet this requirement.

a	b	r_1
c	d	r_2
e	f	r_3
c_1	c_2	N

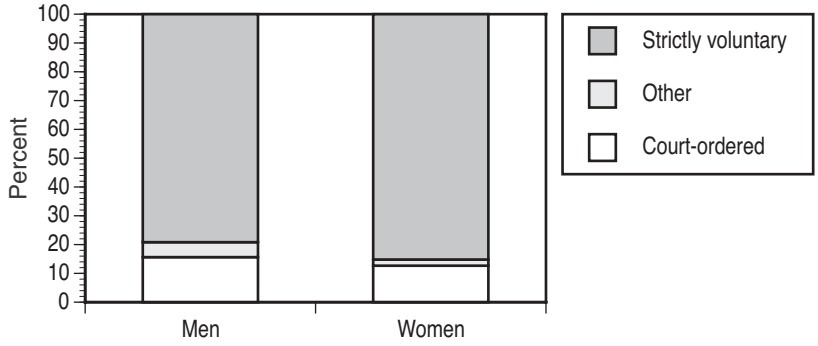
The simplest such table would have all six counts a, b, c, d, e, f equal to one another (which would arise if we start with $r_1 = r_2 = r_3$ and $c_1 = c_2$).

9.15. **(a)** Different graphical presentations are possible; one is shown below. More women perform volunteer work; the notably higher percent of women who are "strictly voluntary" participants accounts for the difference. (The "court-ordered" and "other" percents are similar for men and women.) **(b)** Either by adding the three "participant" categories or by subtracting from 100% the non-participant percentage, we find that 40.3% of men and 51.3% of women are participants. The relative risk of being a volunteer is therefore $\frac{51.3\%}{40.3\%} \doteq 1.27$.

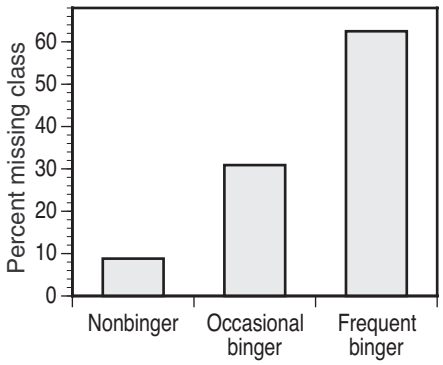


9.16. Table shown on the right; for example, $\frac{31.9\%}{40.3\%} \doteq 79.16\%$. The percents in each row sum to 100%, with no rounding error for up to four places after the decimal. Both this graph and the graph in the previous exercise show that women are more likely to volunteer, but in this view we cannot see the difference in the rate of non-participation.

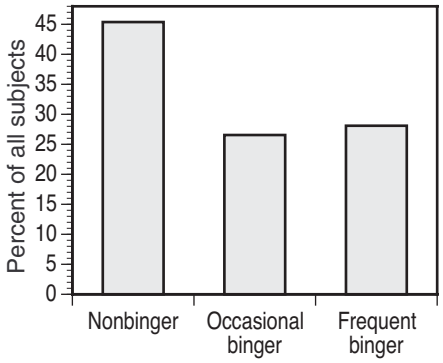
Gender	Strictly voluntary	Court-ordered	Other
Men	79.16%	5.21%	15.63%
Women	85.19%	2.14%	12.67%



9.17. (a) A numerical summary might take different forms, but the most logical way to examine the effect of drinking on missing class is to compute the conditional distribution of missing class for each type of drinker. That is, we look at the distribution of the response variable (missing class) for each value of the explanatory variable (drinking status). We find that only $\frac{4617}{5063} \doteq 8.81\%$ of nonbingers have missed a class because of drinking, while $\frac{2047}{2962} \doteq 30.89\%$ of occasional and $\frac{1176}{3135} \doteq 62.49\%$ of frequent bingers have missed class. One possible graphical display is shown on the right (above).



(b) Among all subjects, $\frac{5063}{11160} \doteq 45.37\%$ were nonbingers, $\frac{2962}{11160} \doteq 26.54\%$ were occasional bingers, and $\frac{3135}{11160} \doteq 28.09\%$ were frequent bingers. A bar graph of these percents is shown on the right (below); a pie graph could also be used. **(c)** To find relative risk, compute the ratios of the percents in (a). For occasional bingers versus nonbingers: $\frac{30.89}{8.81} \doteq 3.5068$. For frequent bingers versus nonbingers: $\frac{62.49}{8.81} \doteq 7.0937$. Compared to nonbingers, occasional bingers are about 3.5 times more likely—and frequent bingers about 7 times more likely—to miss class because of drinking. **(d)** The test statistic is $X^2 \doteq 2672$ (df $= 2$), which has a tiny P-value. The evidence of a relationship between drinking habits and missing class is overwhelming.



Minitab output

```
        Nonbinge     Occas      Freq     Total
   No       4617      2047      1176      7840
         3556.80   2080.83   2202.37

  Yes        446       915      1959      3320
         1506.20    881.17    932.63

Total       5063      2962      3135     11160

ChiSq= 316.019 +  0.550 +478.316 +
       746.262 +  1.299 +1.1E+03 = 2671.963
df = 2, p = 0.000
```

9.18. **(a)** The best numerical summary would note that we view target audience ("magazine readership") as explanatory, so we should compute the conditional distribution of model dress for each audience. This table and graph are shown below. **(b)** Minitab output is shown on the right: $X^2 \doteq 80.9$, df $= 2$, and P is very small. We have very strong evidence that target audience affects model dress. **(c)** The sample is not an SRS: A set of magazines were chosen, and then all ads in three issues of those magazines were examined. It is not clear how this sampling approach might invalidate our conclusions, but it does make them suspect.

Minitab output

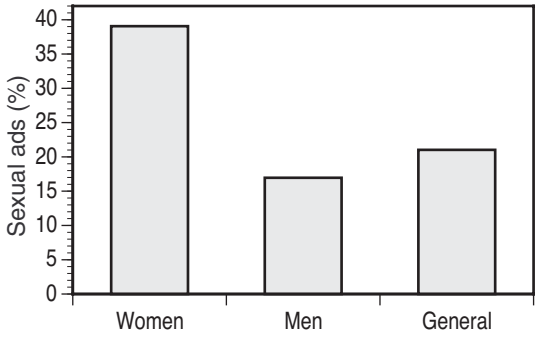
```
          Women      Men     Genl    Total
    1       351      514      248     1113
          424.84   456.56   231.60

    2       225      105       66      396
          151.16   162.44    82.40

Total       576      619      314     1509

ChiSq = 12.835 +  7.227 +  1.162 +
        36.074 + 20.312 +  3.265 = 80.874
df = 2, p = 0.000
```

	Magazine readership		
Model dress	Women	Men	General
Not sexual	60.94%	83.04%	78.98%
Sexual	39.06%	16.96%	21.02%



9.19. **(a)** As the conditional distribution of model dress for each age group has been given to us, it only remains to display this distribution graphically. One such presentation is shown below. **(b)** In order to perform the significance test, we must first recover the counts from the percents. For example, there were $(0.723)(1006) \doteq 727$ non-sexual ads in young adult magazines. The remainder of these counts can be seen in the Minitab output below, where we see $X^2 \doteq 2.59$, df $= 1$, and $P \doteq 0.108$—not enough evidence to conclude that age group affects model dress.

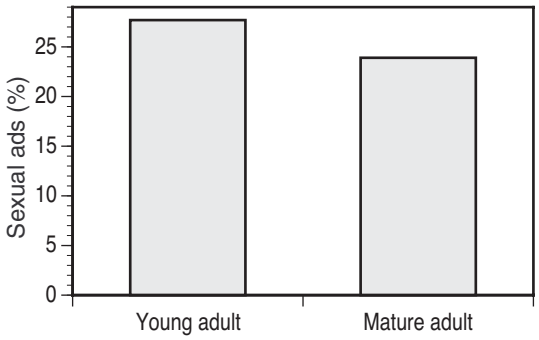
Minitab output

```
          Young    Mature     Total
    1       727       383      1110
          740.00    370.00

    2       279       120       399
          266.00    133.00

Total      1006       503      1509

ChiSq =  0.228 +  0.457 +
         0.635 +  1.271 = 2.591
df = 1, p = 0.108
```



9.20. (a) The missing entries (shown shaded on the right) are found by subtracting the mutation counts from the totals. **(b)** $\frac{30}{96} = 31.25\%$ of steel-mill rats and $\frac{23}{150} \doteq 15.33\%$ of rural rats showed mutation at the *Hm-2* gene locus. **(c)** The Minitab output below shows $X^2 \doteq 8.773$, df $= 1$, $P = 0.003$—strong evidence that location and mutation occurrence are related.

	Location	
Mutation	Steel-mill air	Rural air
Yes	30	23
No	66	127
Total	96	150

Minitab output

```
            Mill     Rural     Total
   1          30        23        53
           20.68     32.32

   2          66       127       193
           75.32    117.68

Total         96       150       246

ChiSq =   4.197 +   2.686 +
          1.153 +   0.738 = 8.773
df = 1, p = 0.003
```



9.21. (a) The missing entries (shown shaded on the right) are found by subtracting the number who have tried low-fat diets from the given totals. **(b)** Viewing gender as explanatory, compute the conditional distributions of low-fat diet for each gender: $\frac{35}{181} \doteq 19.34\%$ of women and $\frac{8}{105} \doteq 7.62\%$ of men have tried low-fat diets. **(c)** The test statistic is $X^2 = 7.143$ (df $= 1$), for which $P = 0.008$. We have strong evidence of an association; specifically, women are more likely to try low-fat diets.

	Gender	
Low-fat diet?	Women	Men
Yes	35	8
No	146	97
Total	181	105

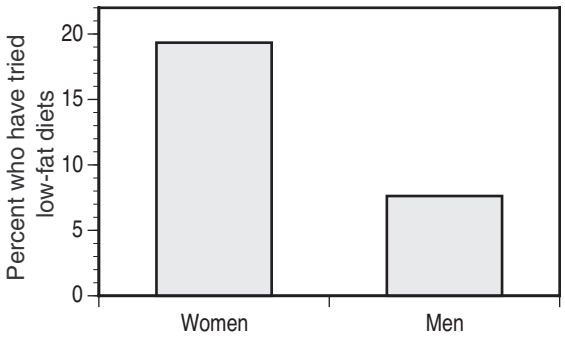
Minitab output

```
            Women       Men     Total
  Yes          35         8        43
            27.21     15.79

   No         146        97       243
          153.79     89.21

Total         181       105       286

ChiSq =   2.228 +   3.841 +
          0.394 +   0.680 = 7.143
df = 1, p = 0.008
```



9.22. **(a)** Subtract the "agreed" counts from the sample sizes to get the "disagreed" counts. The table is in the Minitab output on the right. (The output has been slightly altered to have more descriptive row and column headings.) We find $X^2 \doteq 2.67$, df $= 1$, and $P = 0.103$, so we cannot conclude that students and non-students differ in the response to this question. **(b)** For testing H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$, we have $\hat{p}_1 \doteq 0.3607$, $\hat{p}_2 \doteq 0.5085$, $\hat{p} \doteq 0.4333$, $\text{SE}_{D_p} \doteq 0.09048$, and $z = -1.63$. Up to rounding, $z^2 = X^2$ and the P-values are the same. **(c)** The statistical tests in (a) and (b) assume that we have two SRSs, which we clearly do not have here. Furthermore, the two groups differed in geography (northeast/West Coast) in addition to student/non-student classification. These issues mean we should not place too much confidence in the conclusions of our significance test—or, at least, we should not generalize our conclusions too far beyond the populations "upper level northeastern college students taking a course in Internet marketing" and "West Coast residents willing to participate in commercial focus groups."

Minitab output

```
        Students   Non-st    Total
Agr           22       30       52
           26.43    25.57

Dis           39       29       68
           34.57    33.43

Total         61       59      120

ChiSq =  0.744 +  0.769 +
         0.569 +  0.588 = 2.669
df = 1, p = 0.103
```

9.23. **(a)** First we must find the counts in each cell of the two-way table. For example, there were about $(0.172)(5619) \doteq 966$ Division I athletes who admitted to wagering. These counts are shown in the Minitab output on the right, where we see that $X^2 \doteq 76.7$, df $= 2$, and $P < 0.0001$. There is very strong evidence that the percent of athletes who admit to wagering differs by division. **(b)** Even with much smaller numbers of students (say, 1000 from each division), P is still very small. Presumably the estimated numbers are reliable enough that we would not expect the true counts to be less than 1000, so we need not be concerned about the fact that we had to estimate the sample sizes. **(c)** If the reported proportions are wrong, then our conclusions may be suspect—especially if it is the case that athletes in some division were more likely to say they had not wagered when they had. **(d)** It is difficult to predict exactly how this might affect the results: Lack of independence could cause the estimated percents to be too large, or too small, if our sample included several athletes from teams which have (or do not have) a "gambling culture."

Minitab output

```
          Div1     Div2     Div3    Total
1          966      621      998     2585
        1146.87   603.54   834.59

2         4653     2336     3091    10080
        4472.13  2353.46  3254.41

Total     5619     2957     4089    12665

ChiSq = 28.525 +  0.505 + 31.996 +
         7.315 +  0.130 +  8.205 = 76.675
df = 2, p = 0.000
```

9.24. In Exercise 9.17, we test for independence (model 2) between drinking status and missing classes because of drinking. In Exercise 9.18, we are comparing three populations (model 1): advertisements in magazines targeting women, those in magazines aimed at men, and those in general-interest magazines. In Exercise 9.20, we are comparing two populations (model 1): mice exposed to steel-mill air and those exposed to rural air. In Exercise 9.23, one could argue for either answer. If we chose three separate random samples from each division, then we are comparing three populations (model 1). If a single random sample

of student athletes was chosen, and then we classified each student by division and by gambling response, this is a test for independence (model 2).

Note: *For some of these problems, either answer may be acceptable, provided a reasonable explanation is given. The distinctions between the models can be quite difficult to make since the difference between several populations might, in fact, involve classification by a categorical variable. In many ways, it comes down to how the data were collected. Of course, the difficulty is that the method of collecting data may not always be apparent, in which case we have to make an educated guess. One question we can ask to educate our guess is whether we have data that can be used to estimate the (population) marginal distributions.*

9.25. The Minitab output on the right shows both the two-way table (column and row headings have been changed to be more descriptive) and the results for the significance test: $X^2 \doteq 12.0$, df $= 1$, and $P = 0.001$, so we conclude that gender and flower choice are related. The count of 0 does not invalidate the test: Our smallest expected count is 6, while the text says that "for 2×2 tables, we require that all four expected cell counts be 5 or more."

Minitab output

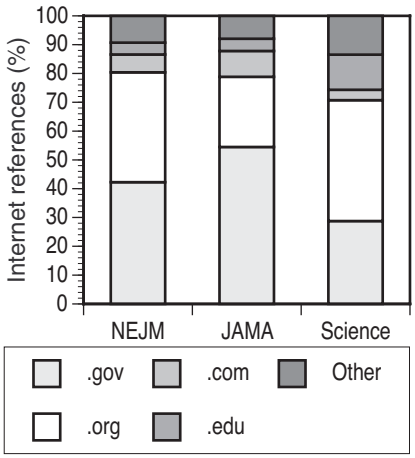
```
          Female      Male     Total
bihai         20         0        20
           14.00      6.00

  no          29        21        50
           35.00     15.00

Total         49        21        70

ChiSq =  2.571 +  6.000 +
         1.029 +  2.400 = 12.000
df = 1, p = 0.001
```

9.26. The graph below depicts the conditional distribution of domain type for each journal; for example, in *NEJM*, $\frac{41}{97} \doteq 42.27\%$ of Internet references were to .gov domains, $\frac{37}{97} \doteq 38.14\%$ were to .org domains, and so on. The Minitab output shows the expected counts, which tell a story similar to the bar graph, and show that the relationship between journal and domain type is significant ($X^2 \doteq 56.12$, df $= 8$, $P < 0.0005$).



Minitab output

```
          NEJM      JAMA   Science     Total
.gov        41       103       111       255
         36.81     71.72    146.47

.org        37        46       162       245
         35.36     68.91    140.73

.com         6        17        14        37
          5.34     10.41     21.25

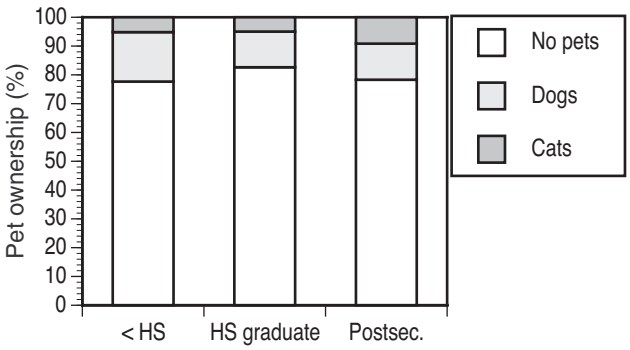
.edu         4         8        47        59
          8.52     16.59     33.89

other        9        15        52        76
         10.97     21.37     43.65

Total       97       189       386       672

ChiSq = 0.477 + 13.644 +  8.591 +
        0.076 +  7.615 +  3.215 +
        0.081 +  4.178 +  2.475 +
        2.395 +  4.451 +  5.072 +
        0.354 +  1.901 +  1.595 = 56.12
df = 8, p = 0.000
```

9.27. The graph on the right depicts the conditional distribution of pet ownership for each education level; for example, among those who did not finish high school, $\frac{421}{542} \doteq 77.68\%$ owned no pets, $\frac{93}{542} \doteq 17.16\%$ owned dogs, and $\frac{28}{542} \doteq 5.17\%$ (the rest) owned cats. (One could instead compute column percents—the conditional distribution of education for each pet-ownership group—but education level makes more sense as the explanatory variable here.) The (slightly altered) Minitab output shows that the relationship between education level and pet ownership is significant ($X^2 \doteq 23.15$, df $= 4$, $P < 0.0005$). Specifically, dog owners have less education, and cat owners more, than we would expect if there were no relationship between pet ownership and educational level.

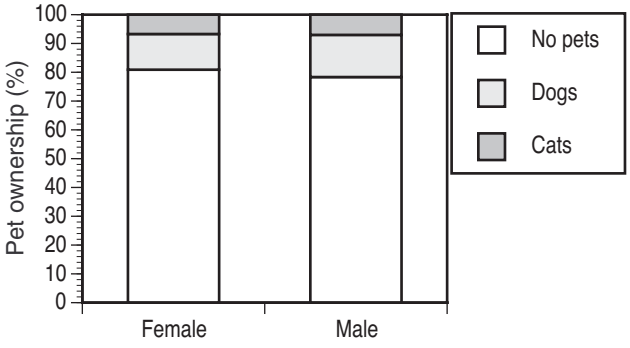


Minitab output

	None	Dogs	Cats	Total
<HS	421	93	28	542
	431.46	73.25	37.29	
HS	666	100	40	806
	641.61	108.93	55.46	
>HS	845	135	99	1079
	858.93	145.82	74.25	
Total	1932	328	167	2427

```
ChiSq =  0.253 +  5.326 +  2.316 +
         0.927 +  0.732 +  4.310 +
         0.226 +  0.803 +  8.254 = 23.147
df = 4, p = 0.000
```

9.28. The graph on the right depicts the conditional distribution of pet ownership for each gender; for example, among females, $\frac{1024}{1266} \doteq 80.88\%$ owned no pets, $\frac{157}{1266} \doteq 12.40\%$ owned dogs, and $\frac{85}{1266} \doteq 6.71\%$ (the rest) owned cats. (One could instead compute column percents—the conditional distribution of gender for each pet-ownership group—but gender makes more sense as the explanatory variable here.) The (slightly altered) Minitab output shows that the relationship between education level and pet ownership is *not* significant ($X^2 \doteq 2.838$, df $= 2$, $P = 0.242$).

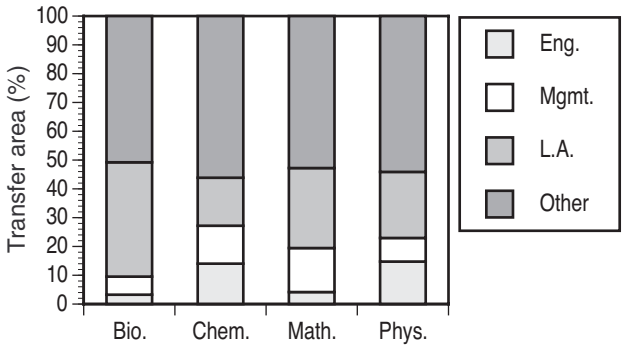


Minitab output

	None	Dogs	Cats	Total
Female	1024	157	85	1266
	1008.53	170.60	86.86	
Male	915	171	82	1168
	930.47	157.40	80.14	
Total	1939	328	167	2434

```
ChiSq =  0.237 +  1.085 +  0.040 +
         0.257 +  1.176 +  0.043 = 2.838
df = 2, p = 0.242
```

9.29. The missing entries can be seen in the "Other" column of the Minitab output below; they are found by subtracting the engineering, management, and liberal arts counts from each row total. The graph on the right shows the conditional distribution of transfer area for each initial major; for example, of those initially majoring in biology, $\frac{13}{398} \doteq 3.27\%$ transferred to engineering, $\frac{25}{398} \doteq 6.28\%$ transferred to management, and so on. The relationship is significant ($X^2 \doteq 50.53$, df $= 9$, $P < 0.0005$). The largest contributions to X^2 come from chemistry or physics to engineering and biology to liberal arts (more transfers than expected) and biology to engineering and chemistry to liberal arts (fewer transfers than expected).



Minitab output

```
          Eng      Mgmt       LA    Other    Total
Bio        13        25      158      202      398
         25.30     34.56   130.20   207.95

Chem       16        15       19       64      114
          7.25      9.90    37.29    59.56

Math        3        11       20       38       72
          4.58      6.25    23.55    37.62

Phys        9         5       14       33       61
          3.88      5.30    19.96    31.87

Total      41        56      211      337      645

ChiSq =  5.979 +  2.642 +  5.937 +  0.170 +
        10.574 +  2.630 +  8.973 +  0.331 +
         0.543 +  3.608 +  0.536 +  0.004 +
         6.767 +  0.017 +  1.777 +  0.040 = 50.527
df = 9, p = 0.000
```

9.30. **(a)** The null hypothesis is H_0: $p_1 = p_2$, where p_1 and p_2 are the proportions of women customers in each city. $\hat{p}_1 = \frac{203}{241} \doteq 0.8423$, $\hat{p}_2 = \frac{150}{218} \doteq 0.6881$, and $\hat{p} = \frac{203+150}{241+218} \doteq 0.7691$, so $\text{SE}_{D_p} \doteq 0.03939$, $z = 3.92$, and $P = 0.0001$. **(b)** $X^2 \doteq 15.334$, which equals z^2. With df $= 1$, Table F tells us that $P < 0.0005$. **(c)** For a confidence interval, we compute $\text{SE}_D \doteq 0.03919$, and the 95% confidence interval is $0.1543 \pm 0.0768 \doteq 0.0774$ to 0.2311. Using the plus four method: $\tilde{p}_1 \doteq 0.8395$ and $\tilde{p}_2 \doteq 0.6864$, $\text{SE}_D \doteq 0.03915$, and the interval is $0.1531 \pm 0.0767 \doteq 0.0764$ to 0.2299.

Minitab output

```
          City1    City2    Total
M            38       68      106
          55.66    50.34

W           203      150      353
         185.34   167.66

Total       241      218      459

ChiSq =  5.601 +  6.192 +
         1.682 +  1.859 = 15.334
df = 1, p = 0.000
```

9.31. With $X^2 = 3.955$ and df $= (5-1)(2-1) = 4$, Table F tells us that $P > 0.25$ (or, with software, we find that $P = 0.413$). There is little evidence to make us believe that there is a relationship between city and income.

9.32. Note that the given counts actually form a three-way table (classified by adhesive, side, and checks). Therefore, this analysis should *not* be done as if the counts come from a 2×4 two-way table; for one thing, no conditional distribution will answer the question of interest (how to avoid face checks). Nonetheless, many students may do this analysis, for which they will find $X^2 = 6.798$, df $= 3$, and $P = 0.079$.

	Face checks	
	No	Yes
PVA/loose	10	54
PVA/tight	44	20
UF/loose	21	43
UF/tight	37	27

A better approach is to rearrange the table as shown on the right. The conditional distributions across the rows will then give us information about avoiding face checks; the graph below illustrates this. We find $X^2 \doteq 45.08$, df $= 3$, and $P < 0.0005$, so we conclude that the appearance of face checks is related to the adhesive/side combination—specifically, we recommend the PVA/tight combination.

Another approach (not quite as good as the previous one) is to perform two separate analyses—say, one for loose side, and one for tight side. These computations show that UF is better than PVA for loose side ($X^2 \doteq 5.151$, df $= 1$, $P = 0.023$), but there is no significant difference for tight side ($X^2 \doteq 1.647$, df $= 1$, $P = 0.200$). We could also do separate analyses for PVA ($X^2 \doteq 37.029$, df $= 1$, $P < 0.0005$) and UF ($X^2 \doteq 8.071$, df $= 1$, $P = 0.005$), from which we conclude that for either adhesive, the tight side has fewer face checks.

Minitab output

```
          NoChk      Chk    Total
PVA-L        10       54       64
          28.00    36.00

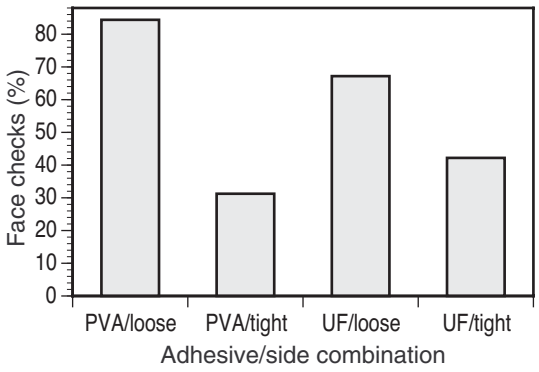
PVA-T        44       20       64
          28.00    36.00

 UF-L        21       43       64
          28.00    36.00

 UF-T        37       27       64
          28.00    36.00

Total       112      144      256

ChiSq = 11.571 +  9.000 +
         9.143 +  7.111 +
         1.750 +  1.361 +
         2.893 +  2.250 = 45.079
df = 3, p = 0.000
```



Minitab output

```
--------  Loose side  --------
         NoChk      Chk    Total
  PVA       10       54       64
         15.50    48.50

   UF       21       43       64
         15.50    48.50

Total       31       97      128

ChiSq =  1.952 +  0.624 +
         1.952 +  0.624 = 5.151
df = 1, p = 0.023
--------  Tight side  --------
         NoChk      Chk    Total
  PVA       44       20       64
         40.50    23.50

   UF       37       27       64
         40.50    23.50

Total       81       47      128

ChiSq =  0.302 +  0.521 +
         0.302 +  0.521 = 1.647
df = 1, p = 0.200
```

Minitab output

```
---------  PVA  ---------
         NoChk      Chk    Total
Loose       10       54       64
         27.00    37.00

Tight       44       20       64
         27.00    37.00

Total       54       74      128

ChiSq = 10.704 +  7.811 +
        10.704 +  7.811 = 37.029
df = 1, p = 0.000
---------  UF  ---------
         NoChk      Chk    Total
Loose       21       43       64
         29.00    35.00

Tight       37       27       64
         29.00    35.00

Total       58       70      128

ChiSq =  2.207 +  1.829 +
         2.207 +  1.829 = 8.071
df = 1, p = 0.005
```

9.33. (a) We should examine column percents because we suspect that "source" is explanatory. These are given in the table (along with expected counts for the chi-square test). The test statistic for cats is:

$$X^2 = 1.305 + 0.666 + 2.483 + 0.632 + 0.323 + 1.202 = 6.611 \ (\text{df} = 2, \ P = 0.037)$$

For dogs:

$$X^2 = 0.569 + 9.423 + 9.369 + 0.223 + 3.689 + 3.668 = 26.939 \ (\text{df} = 2, \ P < 0.0005)$$

Cats	Private	Pet store	Other	
Cases	124	16	76	216
	111.92	13.05	91.03	
	36.15%	40.00%	27.24%	32.63%
Control	219	24	203	446
	231.08	26.95	187.97	
	63.85%	60.00%	72.76%	67.37%
	343	40	279	662

Dogs	Private	Pet store	Other	
Cases	188	7	90	285
	198.63	21.10	65.27	
	26.63%	9.33%	38.79%	28.13%
Control	518	68	142	728
	507.37	53.90	166.73	
	73.37%	90.67%	61.21%	71.87%
	706	75	232	1013

The test is certainly significant for dogs, and is significant at $\alpha = 0.05$ for cats. **(b)** Dogs from pet stores are less likely to go to a shelter, while "other source" dogs are more likely to go. Private-source cats were slightly more likely, and other-source cats slightly less likely, to be taken to the shelter. **(c)** The control group data should be reasonably like an SRS because the sample was taken using a random-digit dialer. The cases data may be less like an SRS; this is difficult to judge. (For example, we would like to know, was this a sample of people who brought their pets to the shelter during a specific time period?)

9.34. Since we suspect that student loans may explain career choice, we examine column percents (in the table below, left). We observe that those with loans are *slightly* more likely to be in Agriculture, Science, and Technology fields and less likely to be in Management. However, the differences in the table are not significant: $X^2 = 6.525$, df $= 6$, $P = 0.368$.

For 9.34.

	Loan	No loan	
Agric.	32 8.7%	35 7.0%	67 7.7%
CDFS	37 10.1%	50 10.1%	87 10.1%
Eng.	98 26.6%	137 27.6%	235 27.2%
LA/Educ.	89 24.2%	124 24.9%	213 24.6%
Mgmt.	24 6.5%	51 10.3%	75 8.7%
Science	31 8.4%	29 5.8%	60 6.9%
Tech.	57 15.5%	71 14.3%	128 14.8%
	368	497	865

For 9.35.

	Low	Medium	High	
Agric.	5 13.5%	27 6.8%	35 8.2%	67 7.7%
CDFS	1 2.7%	32 8.0%	54 12.6%	87 10.1%
Eng.	12 32.4%	129 32.3%	94 22.0%	235 27.2%
LA/Educ.	7 18.9%	77 19.3%	129 30.1%	213 24.6%
Mgmt.	3 8.1%	44 11.0%	28 6.5%	75 8.7%
Science	7 18.9%	29 7.3%	24 5.6%	60 6.9%
Tech.	2 5.4%	62 15.5%	64 15.0%	128 14.8%
	37	400	428	865

9.35. For the table (above, right), $X^2 = 43.487$ (df $= 12$), so $P < 0.0005$, indicating that there is a relationship between PEOPLE score and field of study.

Science has a large proportion of low-scoring students, while liberal arts/education has a large percentage of high-scoring students. (These two table entries make the largest contributions to the value of X^2.)

9.36. **(a)** The 2×2 table is included in the Minitab output (below, left). **(b)** We find $X^2 \doteq 10.95$, df $= 1$, and $P = 0.001$, so we conclude that there is a relationship between gender and label use—specifically, women are more likely to be label users. **(c)** In Exercise 8.64, we found $z \doteq 3.31$, and (up to rounding) $z^2 = X^2$.

For 9.36.

Minitab output

```
          Women      Men    Total
User         63       27       90
          48.70    41.30

  Non       233      224      457
         247.30   209.70

Total       296      251      547

ChiSq =  4.198 +  4.950 +
         0.827 +  0.975 = 10.949
df = 1, p = 0.001
```

For 9.37.

Minitab output

```
          Mex-Am     Other     Total
Juror        339       531       870
          688.25    181.75

  Not     143272     37393    180665
       142922.75  37742.25

Total     143611     37924    181535

ChiSq =177.226 +671.122 +
         0.853 +  3.232 = 852.433
df = 1, p = 0.000
```

9.37. The Minitab output (above, right) shows the 2×2 table and significance test details: $X^2 = 852.433$, df $= 1$, $P < 0.0005$. Using $z = -29.2$, computed in the solution to Exercise 8.81(c), this equals z^2 (up to rounding).

9.38. Minitab outputs for both analyses are given below. For cats, $X^2 = 8.460$ (df $= 4$), which gives $P = 0.077$. We do not reject H_0 this time; with the 2×3 table, we had $P = 0.037$, so having more cells has "weakened" the evidence. For dogs: $X^2 = 33.208$ (df $= 4$), which gives $P < 0.0005$. The conclusion is the same as before: We reject H_0.

Minitab output

```
- - - - - - - - - - - - - - - - - - - -   Cats   - - - - - - - - - - - - - - - - - - - -
       Private    Store     Home    Stray  Shelter    Total
Cases      124       16       20       38       18      216
       111.92    13.05    18.92    50.25    21.86

 Ctrl      219       24       38      116       49      446
       231.08    26.95    39.08   103.75    45.14

Total      343       40       58      154       67      662

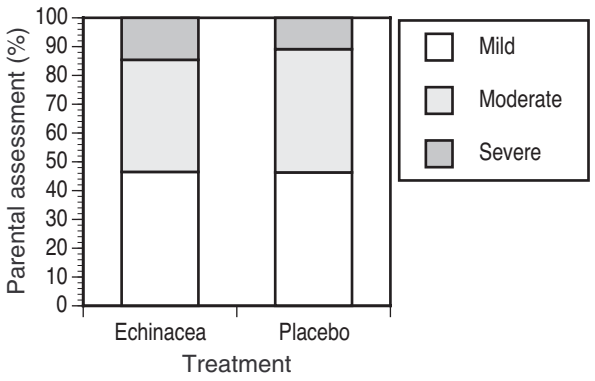
ChiSq =  1.305 +  0.666 +  0.061 +  2.985 +  0.682 +
         0.632 +  0.323 +  0.030 +  1.446 +  0.330 = 8.460
df = 4, p = 0.077
- - - - - - - - - - - - - - - - - - - -   Dogs   - - - - - - - - - - - - - - - - - - - -
       Private    Store     Home    Stray  Shelter    Total
Cases      188        7       11       23       56      285
       198.63    21.10     8.72    21.94    34.61

 Ctrl      518       68       20       55       67      728
       507.37    53.90    22.28    56.06    88.39

Total      706       75       31       78      123     1013

ChiSq =  0.569 +  9.423 +  0.595 +  0.051 + 13.228 +
         0.223 +  3.689 +  0.233 +  0.020 +  5.178 = 33.208
df = 4, p = 0.000
```

9.39. **(a)** The bar graph on the right shows how parental assessment of URIs compares for the two treatments. Note that parental assessment data were apparently not available for all URIs: We have assessments for 329 echinacea URIs and 367 placebo URIs. Minitab output gives $X^2 = 2.506$, df $= 2$, $P = 0.286$, so treatment is not significantly associated with parental assessment. **(b)** If we divide each echinacea count by 337 and each placebo count by 370, we obtain the table of proportions (below, left), and illustrated in the bar graph (below, right). **(c)** The only significant results are for rash ($z = 2.74$, $P = 0.0061$), drowsiness ($z = 2.09$, $P = 0.0366$), and other ($z = 2.09$, $P = 0.0366$). A 10×2 table would not be appropriate, because each URI could have multiple adverse events. **(d)** All results are unfavorable to echinacea, so in this situation we are not concerned that we have falsely concluded that there are differences. In general, when we perform a large number of significance tests and find a few to be significant, we should be concerned that the significant results may simply be due to chance.



Minitab output

```
          Echin  Placebo    Total
Mild        153      170      323
          152.68   170.32

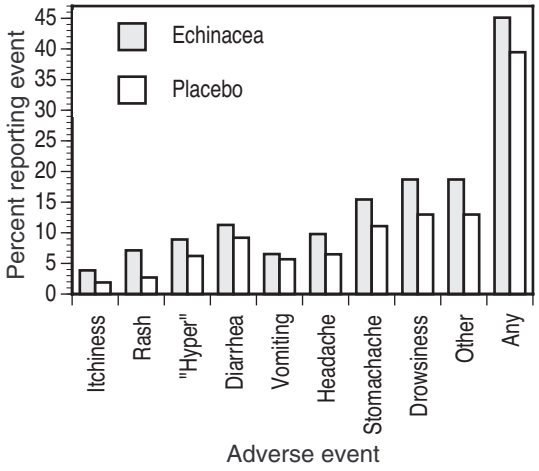
Mod         128      157      285
          134.72   150.28

Sev          48       40       88
           41.60    46.40

Total       329      367      696

ChiSq =  0.001 +  0.001 +
         0.335 +  0.300 +
         0.985 +  0.883 = 2.506
df = 2, p = 0.286
```

Event	$\hat{p}_1$	$\hat{p}_2$	z	P
Itchiness	0.0386	0.0189	1.57	0.1154
Rash	0.0712	0.0270	2.74	0.0061
"Hyper"	0.0890	0.0622	1.35	0.1756
Diarrhea	0.1128	0.0919	0.92	0.3595
Vomiting	0.0653	0.0568	0.47	0.6357
Headache	0.0979	0.0649	1.61	0.1068
Stomachache	0.1543	0.1108	1.71	0.0875
Drowsiness	0.1869	0.1297	2.09	0.0367
Other	0.1869	0.1297	2.09	0.0367
Any event	0.4510	0.3946	1.52	0.1290



(e) We would expect multiple observations on the same child to be dependent, so the assumptions for our analysis are not satisfied. Examination of the data reveals that the results for both groups are quite similar, so we are inclined to agree with the authors that there are no statistically significant differences. **(f)** Student opinions about the criticisms of this study will vary. The third criticism might be dismissed as sounding like conspiracy-theory paranoia, but the other three address the way that echinacea was administered; certainly we cannot place too much faith in a clinical trial if it turns out that the treatments were not given properly!

9.40. **(a)** Each quadrant accounts for one-fourth of the area, so we expect it to contain one-fourth of the 100 trees. **(b)** Some random variation would not surprise us; we no more expect exactly 25 trees per quadrant than we would expect to see exactly 50 heads when flipping a fair coin 100 times. **(c)** The table on the right shows the individual computations, from which we obtain $X^2 = 10.8$, df $= 3$, and $P = 0.0129$. We conclude that the distribution is not random.

Observed	Expected	$(o-e)^2/e$
18	25	1.96
22	25	0.36
39	25	7.84
21	25	0.64
100		10.8

9.41. The chi-square goodness of fit statistic is $X^2 \doteq 3.7807$ with df $= 3$, for which $P > 0.25$ (software gives 0.2861), so there is not enough evidence to conclude that this university's distribution is different. The details of the computation are given in the table below; note that there were 210 students in the sample.

	Expected frequency	Expected count	Observed count	$O-E$	$\frac{(O-E)^2}{E}$
Never	0.43	90.3	79	−11.3	1.4141
Sometimes	0.35	73.5	83	9.5	1.2279
Often	0.15	31.5	36	4.5	0.6429
Very often	0.07	14.7	12	−2.7	0.4959
			210		3.7807

9.42. The chi-square goodness of fit statistic is $X^2 \doteq 3.4061$ with df $= 4$, for which $P > 0.25$ (software gives 0.4923), so we have no reason to doubt that the numbers follow a Normal distribution. The details of the computation are given in the table below. The table entries from Table A for -0.6, -0.1, 0.1, and 0.6 are (respectively) 0.2743, 0.4602, 0.5398, and 0.7257. Then, for example, the expected frequency in the interval -0.6 to -0.1 is $0.4602 - 0.2743 = 0.1859$.

	Expected frequency	Expected count	Observed count	$O-E$	$\frac{(O-E)^2}{E}$
$z \le -0.6$	0.2743	137.2	139	1.85	0.0250
$-0.6 < z \le -0.1$	0.1859	93.0	102	9.05	0.8811
$-0.1 < z \le 0.1$	0.0796	39.8	41	1.20	0.0362
$0.1 < z \le 0.6$	0.1859	93.0	78	−14.95	2.4045
$z > 0.6$	0.2743	137.2	140	2.85	0.0592
					3.4061

9.44. The chi-square goodness of fit statistic is $X^2 = 5.50$ with df $= 4$, for which $0.20 < P < 0.25$ (software gives 0.2397), so we have no reason to doubt that the numbers follow this uniform distribution. The details of the computation are given in the table below.

	Expected frequency	Expected count	Observed count	$O - E$	$\frac{(O-E)^2}{E}$
$0 < x \leq 0.2$	0.2	100	114	14	1.96
$0.2 < x \leq 0.4$	0.2	100	92	−8	0.64
$0.4 < x \leq 0.6$	0.2	100	108	8	0.64
$0.6 < x \leq 0.8$	0.2	100	101	1	0.01
$0.8 < x < 1$	0.2	100	85	−15	2.25
					5.50

9.46. A P-value of 0.999 is suspicious because it means that there was an almost-perfect match between the observed and expected counts. (The table on the right shows how small X^2 must be in order to have a P-value of 0.999; recall that X^2 is small when the observed and expected counts are close.) We expect a certain amount of difference between these counts due to chance, and become suspicious if the difference is too small. In particular, when H_0 is true, a match like this would occur only once in 1000 attempts; if there were 1000 students in the class, that might not be too surprising.

df	X^2
1	2×10^{-6}
2	0.0020
3	0.0243
4	0.0908
5	0.2102
6	0.3810
7	0.5985
8	0.8571
9	1.1519
10	1.4787