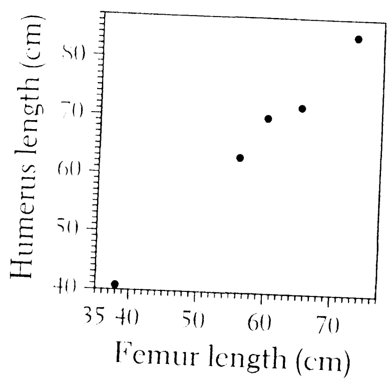
14

Inference for Regression

14.1 (a) See also the solution to Exercise 3.19. The correlation is $r = 0.994$, and linear regression gives $\hat{y} = -3.660 + 1.1969x$. The scatterplot below shows a strong, positive, linear relationship, which is confirmed by r.

(b) β represents how much we can expect the humerus length to increase when femur length increases by 1 cm, b (the estimate of β) is 1.1969, and the estimate of α is $a = -3.660$.

(c) The residuals are −0.8226, −0.3668, 3.0425, −0.9420, and −0.9110; the sum is −0.0001 (but carrying a different number of digits might change this). Squaring and summing the residuals gives 11.79, so that $s = \sqrt{11.79/3} \doteq 1.982$.



14.2 (a) HEIGHT = 71.950 + 0.38333(AGE). The intercept is 71.950 and the slope is 0.38333. (b) The estimate for α is the intercept of the least-squares line, that is, 71.950. The estimate for β is the slope of the least-squares line, that is, 0.38333. (c) The residuals are 0.25012, −0.34984, −0.49983, 0.35018, 0.20019, 0.0502. The formula for s yields $s = \sqrt{\dfrac{.6}{6-2}} = \sqrt{.15} = 0.3873$.

14.3 (a) HEIGHT = 11.547 + 0.84042(ARMSPAN). (b) The least-squares line is an appropriate model for the data because the residual plot shows no obvious pattern. (c) $a = 11.547$ estimates the true intercept, α; $b = 0.84042$ estimates the true slope, β. (d) $s = 1.6128$ estimates σ.

14.4 (a) See Exercise 3.71 for scatterplot. $r = 0.9990$ and the equation of the least-squares line is $\hat{y} = 1.766 + 0.080284x$. The scatterplot shows a strong linear relationship, which is confirmed by r.

13.39 (a) Yes, the evidence is *very* strong that a higher proportion of men die ($X^2 = 332.205$, df = 1). Possibly many sacrificed themselves out of a sense of chivalry ("women and children first").
(b) For women, $X^2 = 103.767$ (df = 2)—a very significant difference. Over half of the lowest-status women died, but this percentage drops sharply when we look at middle-status women, and it drops again for high-status women.
(c) For men, $X^2 = 34.621$ (df = 2)—another very significant difference (though not quite so strong as the women's value). Men with the highest status had the highest proportion surviving (over one-third). The proportion for low-status men was only about half as big, while middle-class men fared worst (only 12.8% survived).

13.40 Answers vary.

13.38 (a) See the Minitab output below for the two-way table. We find $X^2 = 24.243$ with $df = 3$— a very significant result ($P < 0.0005$). The most effective treatment is "Both"; the largest contributions to the X^2 statistic come from the first and last rows of the first column. Subjects taking a placebo had many more strokes than expected, while those taking both drugs had fewer strokes.

(b) See the Minitab output below for the two-way table. We find $X^2 = 1.418$ with $df = 3$, which gives $P = 0.701$; the various drug treatments had no significant effect on deaths.

(c) The combination of both drugs is effective at decreasing the risk of stroke, but no drug treatment had a significant impact on death rate.

Minitab output

```
            Stroke    NoStroke    Total
Placebo        250        1399     1649
            205.81     1443.19

Aspirin        206        1443     1649
            205.81     1443.19

Dipyr.         211        1443     1654
            206.44     1447.56

Both           157        1493     1650
            205.94     1444.06

Total          824        5778     6602

ChiSq =  9.487 + 1.353 +
         0.000 + 0.000 +
         0.101 + 0.014 +
        11.629 + 1.658 = 24.243
df = 3, p = 0.000
```

Minitab output

```
            Death     NoDeath    Total
Placebo       202        1447     1649
           189.08     1459.92

Aspirin       182        1467     1649
           189.08     1459.92

Dipyr.        188        1466     1654
           189.65     1464.35

Both          185        1465     1650
           189.19     1460.81

Total         757        5845     6602

ChiSq =  0.883 + 0.114 +
         0.265 + 0.034 +
         0.014 + 0.002 +
         0.093 + 0.012 = 1.418
df = 3, p = 0.701
```

	Temperature		
	Cold	Neutral	Hot
Hatched	16	38	75
Did not hatch	11	18	29
Total	27	56	104

Minitab output

```
        Cold     Neutral      Hot    Total
  1       16          38       75      129
       18.63       38.63    71.74

  2       11          18       29       58
        8.37       17.37    32.26

Total     27          56      104      187

ChiSq =  0.370 + 0.010 + 0.148 +
         0.823 + 0.023 + 0.329 = 1.703

df = 2, p = 0.427
```

13.36 We wish to test H_0: The hypothesized model is correct against H_a: The hypothesized model is not correct. The expected number of green-seeded plants according to the model is $(^1/_4)(880) = 220$. The $(^3/_4)(880) = 660$, while the expected number of yellow-seeded plants is value of the chi-square statistic is $X^2 = \frac{(639-660)^2}{660} + \frac{(241-220)^2}{220} = 0.668 + 2.0045 = 2.6725$. With df = 1, the P-value is $P(X_1^2 > 2.6725) = 0.1021$. There is no reason to doubt the model.

13.37 (a) No: No treatment was imposed.

(b) See the column percents in the table below. Pet owners seem to have better survival rates.

	No Pet	Pet	
Alive	28 33.07 71.8%	50 44.93 94.3%	78 84.8%
Dead	11 5.93 28.2%	3 8.07 5.7%	14 15.2%
	39	53	92

(c) $X^2 = 0.776 + 0.571 + 4.323 + 3.181 = 8.851$ (df = 1), so $0.0025 < P < 0.005$ (in fact, $P = 0.003$).

(d) Provided we believe that there are no confounding or lurking variables, we reject H_0 and conclude that owning a pet improves survival.

(e) We used a χ^2 test. In a z test, we would test H_0: $p_1 = p_2$ vs. H_a: $p_1 < p_2$, where p_1 = the proportion of non-pet owners who survived and p_2 = the proportion of pet owners who survived. For this test, $\hat{p}_1 = .718, \hat{p}_2 = .943, \hat{p} = .848, z = -2.975$, and the P-value = 0.0015. As in the χ^2 test, we reject H_0 and conclude that there is a significant difference in the survival rates. The p-value is half that obtained in (c).

13.33 (a) $H_0: p_1 = p_2$ vs. $H_a: p_1 < p_2$. The z test must be used because the chi-square procedure will not work for a one-sided alternative. (b) $z = -2.8545$ and $P = 0.0022$. Reject H_0; there is strong evidence in favor of H_a.

13.34 (a) Minitab output below. There is strong evidence of a relationship: $X^2 = 11.141$ (df = 2), which means that $0.0025 < P < 0.005$. The second row of the table accounts for most of this; those two cells contribute 3.634 + 4.873 = 8.507 to the value of X^2.

(b) Minitab output (including the two-way table) below. The relationship is still significant ($X^2 = 10.751$, df = 1, $0.0005 < P < 0.001$). A lower-than-expected number of identical twins had different behavior, and a (slightly) higher-than-expected number of identical twins had the same behavior.

Minitab output

```
              Ident      Frat     Total
Neither         443       301       744
             426.17    317.83

One             102       113       215
             123.16     91.84

Both             45        26        71
              40.67     30.33

Total           590       440      1030

ChiSq =  0.664 + 0.891 +
         3.634 + 4.873 +
         0.461 + 0.618 = 11.141
df = 2, p = 0.004
```

Minitab output

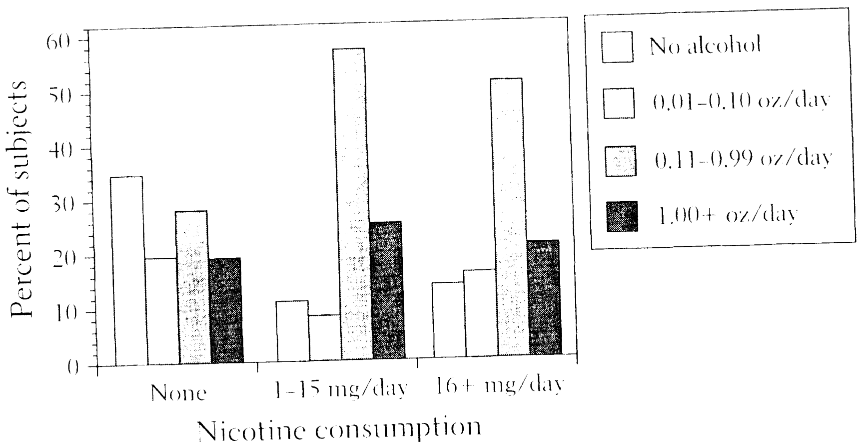
```
              Ident      Frat     Total
Same            488       327       815
             466.84    348.16

Diff            102       113       215
             123.16     91.84

Total           590       440      1030

ChiSq =  0.959 + 1.285 +
         3.634 + 4.873 = 10.751
df = 1, p = 0.001
```

13.35 (a) On the facing page. (b) Cold water: $\hat{p}_1 = \frac{16}{27} \doteq 59.3\%$; Neutral: $\hat{p}_2 = \frac{38}{56} \doteq 67.9\%$; Hot water: $\hat{p}_3 = \frac{75}{104} \doteq 72.1\%$. The percentage hatching increases with temperature; the cold water did not prevent hatching, but made it less likely. (c) The differences are not significant: $X^2 = 1.703$, df = 2, and $P = 0.427$.

13.32 There is no reason to consider one of these variables as explanatory, but a conditional distribution is useful to determine the nature of the association. Each cell in the table below contains a pair of percentages; the first is the column percent, and the second is the row percent. For example, among nonsmokers, 34.5% were nondrinkers; among nondrinkers, 85.4% were nonsmokers. The percentages in the right margin gives the distribution of alcohol consumption (the overall column percent), while the percentages in the bottom margin are the distribution of smoking behavior.

	0 mg	1–15 mg	16+ mg	
0 oz	105 82.73 34.5% 85.4%	7 17.69 10.8% 5.7%	11 22.59 13.3% 8.9%	123 27.2%
0.01–0.10 oz	58 51.12 19.1% 76.3%	5 10.93 7.7% 6.6%	13 13.96 15.7% 17.1%	76 16.8%
0.11–0.99 oz	84 109.63 27.6% 51.5%	37 23.44 56.9% 22.7%	42 29.93 50.6% 25.8%	163 36.1%
1.00+ oz	57 60.53 18.8% 63.3%	16 12.94 24.6% 17.8%	17 16.53 20.5% 18.9%	90 19.9%
	304 67.3%	65 14.4%	83 18.4%	452 100%

$X^2 = 42.252$ (df $= 6$) so $P < 0.0005$; we conclude that alcohol and nicotine consumption are not independent. The chief deviation from independence (based on comparison of expected and actual counts) is that nondrinkers are more likely to be nonsmokers than we might expect, while those drinking 0.11 to 0.99 oz/day are less likely to be nonsmokers than we might expect. One possible graph is below.

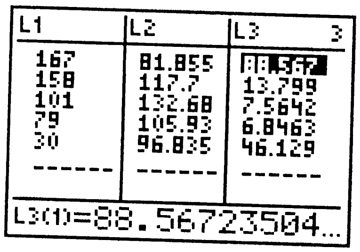


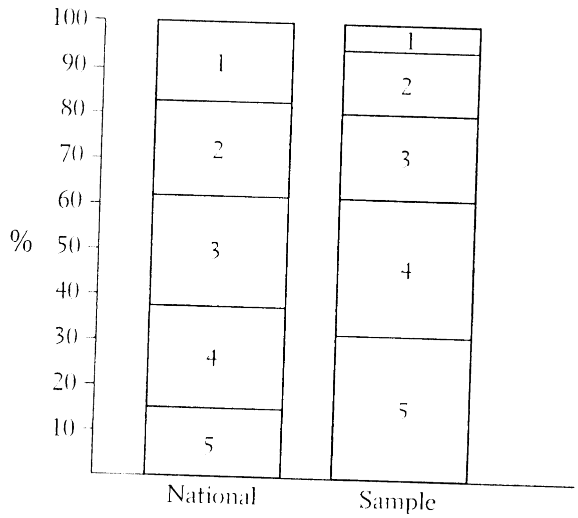
13.30 4 degrees of freedom; $P > 0.25$ (in fact, $P = 0.4121$). There is not enough evidence to reject H_0 at any reasonable level of significance; the difference in the two income distributions is not statistically significant.

13.31 The observed frequencies of scores in this sample, their marginal percents, and the expected numbers were:

Score	5	4	3	2	1
Frequency	167	158	101	79	30
Percent	31.2	29.5	18.9	14.8	5.6
Expected	81.855	117.7	132.68	105.93	96.835

H_0: The distribution of scores in this sample is the same as the distribution of scores for all students who took this inaugural exam. H_a: The distribution of scores in this sample is different from the national results. The degrees of freedom are $n - 1 = 4$, and the chi-square statistic is $X^2 = 88.57 + 13.80 + 7.56 + 6.85 + 46.13 = 162.9$. The P-value is $P(X_4^2 > 162.9) = 3.49 \times 10^{-34} = .0000$. Reject H_0 and conclude that the distribution of AP Statistics exam scores in this sample is different from the national distribution.





Postscript: As soon as the exam grades were sent to the students and their schools, in July 1997, several AP Statistics teachers who were subscribers to an AP Statistics discussion group on the Internet posted their grades in the spirit of sharing the results with their fellow teachers. While this was of interest to many of the pioneering AP Statistics teachers in their first year teaching this new course, this sample was a voluntary response sample, not a random sample. It should come as no surprise that these self-reported results were weighted toward the higher scores.

Minitab output

```
            Yes        No     Total
  1          58        58       116
          46.94     69.06

  2          84       129       213
          86.19    126.81

  3         169       294       463
         187.36    275.64

  4          98       135       233
          94.29    138.71

  5          77        99       176
          71.22    104.78

Total       486       715      1201

ChiSq =  2.605 +  1.771 +
         0.056 +  0.038 +
         1.799 +  1.223 +
         0.146 +  0.099 +
         0.469 +  0.319 =  8.525

df = 4, p = 0.075
```

13.27 H_0: all proportions are equal vs. H_a: some proportions are different. Table below. $X^2 = 10.619$ with 2 df; and $P = 0.0049$—good evidence against H_0, so we conclude that contact method makes a difference in response.

	Yes	No
Phone	168	632
One-on-one	200	600
Anonymous	224	576

13.28 (a) 7.01%, 14.02%, and 13.05%. (b) and (c) Table below—actual counts above, expected counts below. Expected counts are all much bigger than 5, so the chi-square test is safe. H_0: there is no relationship between worker class and race vs. H_a: there is some relationship. (d) df = 2; $P < 0.0005$ (basically 0). (e) Black female child-care workers are more likely to work in non-household or preschool positions.

	Black	Other
Household	172 242.36	2283 2212.64
Nonhousehold	167 117.58	1024 1073.42
Teachers	86 65.06	573 593.94

13.29 (a) H_0: $p_1 = p_2$, where p_1 and p_2 are the proportions of women customers in each city. $\hat{p}_1 = 0.8423, \hat{p}_2 = 0.6881$, $z = 3.9159$, and $P = 0.00009$. (b) $X^2 = 15.334$, which equals z^2. With df = 1, Table E tells us that $P < 0.0005$; a statistical calculator gives $P = 0.00009$. (c) 0.0774 to 0.2311.

13.24 (a) H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$. $z = -0.5675$ and $P = 0.5704$. (b) Table of expected counts is given below. $X^2 = 0.322$, which equals z^2. With 1 df, Table E tells us that $P > 0.25$; a statistical calculator gives $P = 0.5704$. (c) Gastric freezing is not significantly more (or less) effective than a placebo treatment.

	Improved	Did not improve
Gastric freezing	28 29.73	54 52.28
Placebo	30 28.27	48 49.72

13.25 (a)

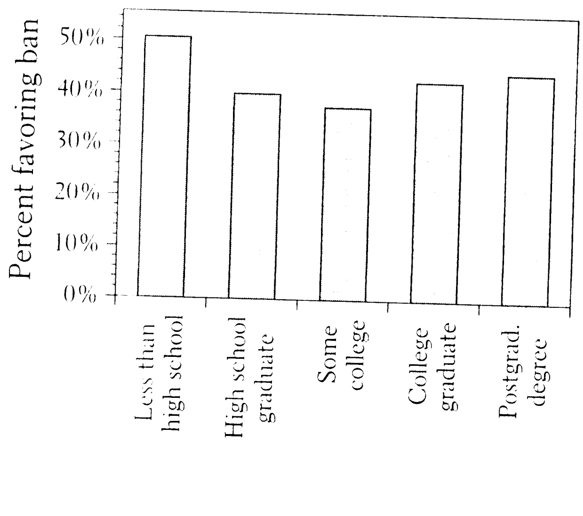
Group	Cardiac event? Yes	No	TOTAL
Stress management	3	30	33
Exercise	7	27	34
Usual care	12	28	40
TOTAL	22	85	107

(b) Success rates (% of "No"s): 90.91%, 79.41%, 70%.
(c)

Group	Cardiac event? Exp. Yes	Exp. No
Stress management	6.785	26.215
Exercise	6.991	27.009
Usual care	8.224	31.776

All expected cell counts exceed 5, so the chi-square test can be used.
(d) $X^2 = 4.84$ (df = 2), P-value = 0.0889. Though the success rate for the stress management group is slightly higher than for the other two groups, there does **not** appear to be a significant difference among the success rates.

13.26 (a) The proportions in favor are $\frac{58}{116} = 0.5$, $\frac{84}{213} \doteq 0.3944$, $\frac{169}{463} \doteq 0.3650$, $\frac{98}{233} \doteq 0.4206$, and $\frac{77}{176} = 0.4375$. Those who did not complete high school and those with a college or graduate degree appear to be more likely to favor a ban. (b) With $X^2 = 8.525$ and 4 degrees of freedom, we have $P = 0.075$, so we cannot reject the null hypothesis; we do not have enough evidence to conclude that the proportion favoring a handgun ban varies significantly with level of education.

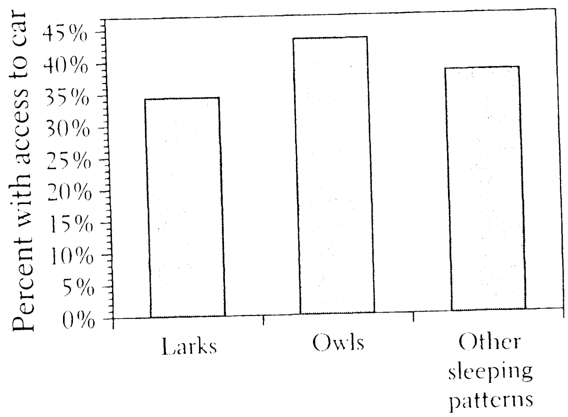


(d) No—this study demonstrates association, not causation. Certain types of students may tend to spend a moderate amount of time in extracurricular activities and also work hard on their classes—one does not necessarily cause the other.

13.22 (a) H_0: There is no association between smoking by parents and smoking by high school students. H_a: There is an association between smoking by parents and smoking by high school students. Expected counts: given in Exercise 13.20(e). $X^2 = 37.566$, df = 2. The P-value is approximately 7×10^{-9}—essentially 0.

By rejecting H_0, we conclude that there is a relationship between parents' smoking habits and those of their children.
(b) The highest contributions come from C1 R1 ("both parents smoke, student smokes") and C1 R3 ("neither parent smokes, student smokes"). When both parents smoke, their student is much more likely to smoke; when neither parent smokes, their student is unlikely to smoke.
(c) No—this study demonstrates association, not causation. There may be other factors (heredity or environment, for example) that cause *both* students and parent(s) to smoke.

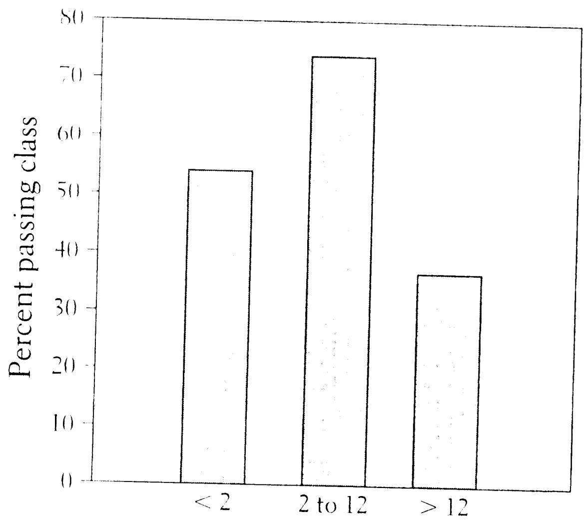
13.23 The proportions with access to a car were $\frac{122}{356} \doteq 0.3427$, $\frac{138}{318} \doteq 0.4340$, and $\frac{213}{555} \doteq 0.3838$ (graph below). To test the null hypothesis (sleep patterns do not affect car access), we find $X^2 = 5.915$ (df = 2), and $P = 0.052$. While the data suggest that owls are more likely to have access to a car, we find that the evidence is not quite significant (at the $\alpha = 0.05$ level).



Minitab output

```
         Yes       No     Total
  1      122      234       356
      137.01   218.99

  2      138      180       318
      122.39   195.61

  3      213      342       555
      213.60   341.40

Total    473      756      1229

ChiSq =  1.645 + 1.029 +
         1.992 + 1.246 +
         0.002 + 0.001 = 5.915
df = 2, p = 0.052
```

(c)



(d) H_0: There is no association between amount of time spent on extracurricular activities and grades earned in the course.
H_a: There is an association.
(e)

	<2	2–12	>12
C or better	13.78	62.71	5.51
D or F	6.22	28.29	2.49

(f) The first and last columns have lower numbers than we expect in the "passing" row (and higher numbers in the "failing" row), while the middle column has this reversed—more passed than we would have expected if the proportions were all equal.

13.20 (a) 3 × 2. (b) 22.5%, 18.6%, and 13.9%. A student's likelihood of smoking increases when one parent smokes, and increases even more when both smoke. (c) See Exercise 4.53. (d) The null hypothesis says that parents' smoking habits have no effect on their children. (e) Below. (f) In column 1, row 1, the expected count is much smaller than the actual count; meanwhile, the actual count is lower than expected in the lower left. This agrees with what we observed before: Children of non-smokers are less likely to smoke.

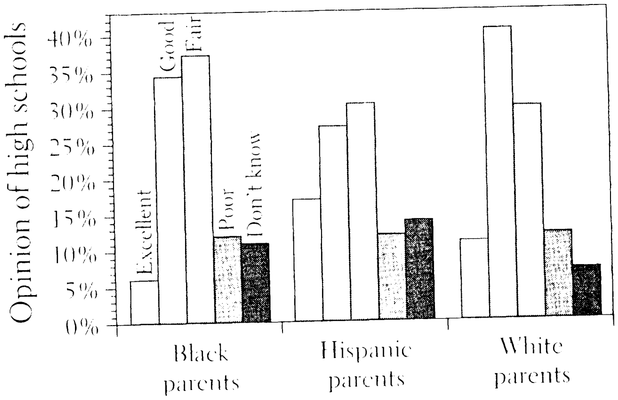
	Student smokes	Student does not smoke
Both parents smoke	332.49	1447.51
One parent smokes	418.22	1820.78
Neither parent smokes	253.29	1102.71

13.21 (a) Missing entries in table of expected counts (row by row): 62.71, 5.51, 6.22. Missing entries in components of X^2: 0.447, 0.991.

(b) df = 2, $P = 0.0313$. Rejecting H_0 means that we conclude that there is a relationship between hours spent in extracurricular activities and performance in the course.

(c) The highest contribution comes from column 3, row 2 ("> 12 hours of extracurricular activities, D or F in the course"). Too much time spent on these activities seems to hurt academic performance.

13.18 Various graphs can be made; one possibility is shown below. For the null hypothesis "There is no relationship between race and opinions about schools," we find $X^2 = 22.426$ (df = 8) and $P = 0.004$ (Minitab output below). We have evidence that there *is* a relationship; specifically, blacks are less likely, and Hispanics more likely, to consider schools "excellent," while Hispanics and whites differ in percentage considering schools "good" (whites are higher) and percentage who "don't know" (Hispanics are higher). Also, a higher percentage of blacks rated schools as "fair."



Minitab output

```
           Black    Hispanic    White    Total
  1           12          34       22       68
           22.70       22.70    22.59

  2           69          55       81      205
           68.45       68.45    68.11

  3           75          61       60      196
           65.41       65.44    65.12

  4           24          24       24       72
           24.04       24.04    23.92

  5           22          28       14       64
           21.37       21.37    21.26

Total        202         202      201      605

ChiSq =  5.047 + 5.620 + 0.015 +
         0.004 + 2.642 + 2.441 +
         1.396 + 0.301 + 0.402 +
         0.000 + 0.000 + 0.000 +
         0.019 + 2.058 + 2.481 = 22.426

df = 8, p = 0.004
```

13.19 (a) $r = 2$, $c = 3$.
(b) 55.0%, 74.7%, and 37.5%. Some (but not too much) time spent in extracurricular activities seems to be beneficial.

categories (LSC-HM and LSC-LM). It appears that males and females have distinctly different goals when they play sports.

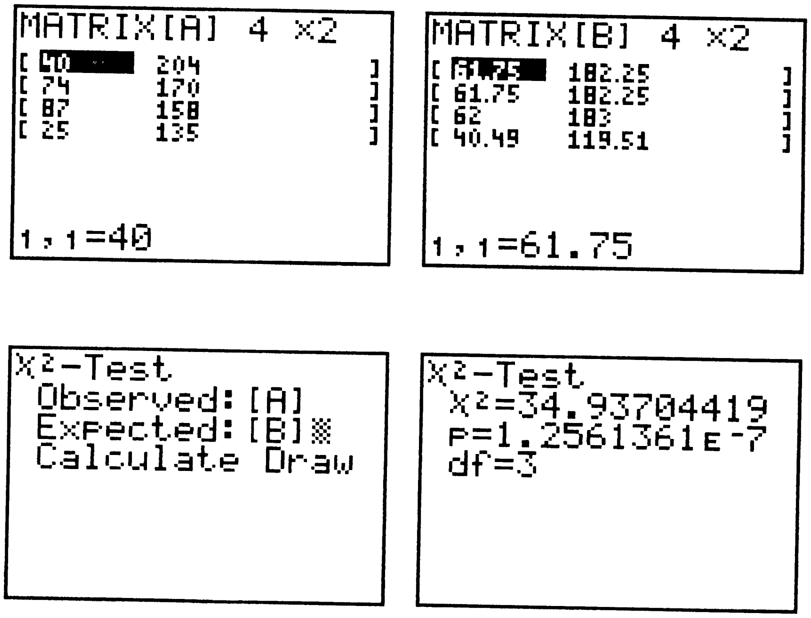
13.16 (a) Reading row by row, the components are 7.662, 2.596, 2.430, 0.823, 10.076, 3.414, 5.928, 2.008. $X^2 = 34.937$ (df = 3).

(b) According to Table E, P-value = $P(X_3^2 > 34.937) < 0.0005$. Since the largest critical value for df = 3 is 17.73, the P-value is actually quite a bit smaller than 0.0005. A P-value of this size indicates that it is extremely unlikely that such a result occurred due to chance; it represents very strong evidence against H_0.

(c) 10.076, corresponding to the "patch plus drug/success" category, is the largest contributor. This is not surprising because, according to Exercise 13.14(f), the "patch plus drug" group contains a higher than expected number of successful quitters.

(d) Treatment is strongly associated with success; specifically, the drug, or the patch together with the drug, seem to be most effective.

(e) See displays below.



13.17 (a) Reading row by row, the components are 3.211, 3.211, 2.420, 2.420, 4.923, 4.923, 1.895, 1.895. $X^2 = 24.898$ (df = 3).

(b) From Table E, $P(X_3^2 > 24.898) < 0.0005$. A P-value of this size indicates that it is extremely unlikely that such a result occurred due to chance; it represents very strong evidence against H_0.

(c) The terms corresponding to HSC-HM and LSC-HM (for both sexes) provide the largest contributions to X^2. This reflects the fact that males are more likely to have "winning" (social comparison) as a goal, while females are more concerned with "mastery."

(d) The terms and results are identical. The P-value of 0.000 in the MINITAB output reflects the fact that the true P-value in part (b) was actually considerably smaller than 0.0005.

(d) The success rate (proportion of those who quit) is the same for all four treatments.

(e) Using the formula for expected counts, we obtain the following table:

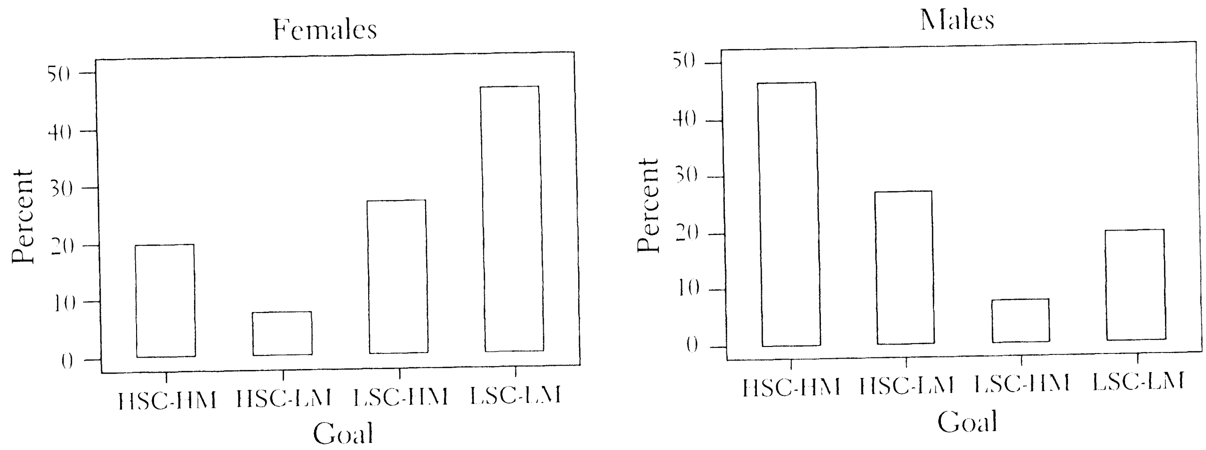
Treatment	Exp. Succ.	Exp. Fail.
Nicotine patch	61.75	182.25
Drug	61.75	182.25
Patch plus drug	62	183
Placebo	40.49	119.51

(f) A higher percentage than expected of the "patch plus drug" and "drug" subjects successfully quit, and a lower percentage than expected of the other two groups quit. This reflects the fact that the "patch plus drug" and "drug" success rates were considerably higher than the success rates for the other two groups, as seen in (b) and (c).

13.15 (a) r = the number of rows in the table, c = the number of columns in the table. In this case $r = 4$ and $c = 2$.

(b) Females: 20.9% HSC-HM, 10.4% HSC-LM, 31.3% LSC-HM, 37.3% LSC-LM. Males: 46.3% HSC-HM, 26.9% HSC-LM, 7.5% LSC-HM, 19.4% LSC-LM.

(c)



(d)

Goal	Exp. Counts (Females)	Exp. Counts (Males)
HSC-HM	22.5	22.5
HSC-LM	12.5	12.5
LSC-HM	13	13
LSC-LM	19	19

(e) Males are more likely to fall into the HSC categories (HSC-HM and HSC-LM) than their expected counts would predict. Likewise, females are more likely to fall into the LSC

13.11 The observed and expected values are:

Flavor	Grape	Lemon	Lime	Orange	Strawberry
Observed	530	470	420	610	585
Expected	523	523	523	523	523

H_0: Trix flavors are uniformly distributed. H_a: The flavors are not uniformly distributed. df = 5 − 1 = 4, and X^2 = .09369 + 5.3709 + 20.285 + 14.472 + 7.3499 = 47.57. $P(X_4^2 > 47.57) = 1.16 \times 10^{-9}$ $\doteq$ .0000. Reject H_0 and conclude that either the Trix flavors are *not* uniformly distributed, or our box of Trix is not a *random* sample.

13.12 Answers will vary.

13.13 Since the wheel is divided into four equal parts, if it is in balance, then the four outcomes should occur with approximately equal frequency. Here are the observed and expected values:

Parts	I	II	III	IV
Observed	95	105	135	165
Expected	125	125	125	125

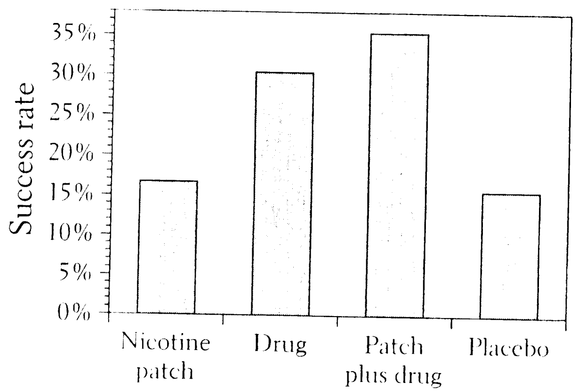
H_0: The wheel is balanced (the four outcomes are uniformly distributed). H_a: The wheel is not balanced. df = 3 and X^2 = 7.2 + 3.2 + 0.8 + 12.8 = 24. The P-value is $P(X_3^2 > 24) = 2.5 \times 10^{-5}$ = .000025. Reject H_0 and conclude that the wheel is not balanced. Since "Part IV: Win nothing" shows the greatest deviation from the expected result, there may be reason to suspect that the carnival game operator may have tampered with the wheel to make it harder to win.

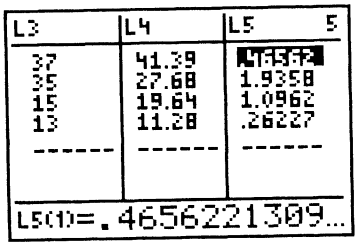
13.14 (a)

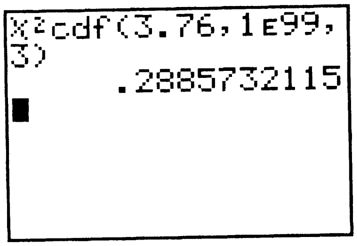
Treatment	Successes	Failures
Nicotine patch	40	204
Drug	74	170
Patch plus drug	87	158
Placebo	25	135

(b) The success rates are $\frac{40}{244} \doteq 0.1639$, $\frac{74}{244} \doteq 0.3033$, $\frac{87}{245} \doteq 0.3551$, and $\frac{25}{160} = 0.15625$.

(c)







Note that none of the difference terms is very large. The test statistic is $X^2 = 3.76$, df = 3, and the *P*-value = .2886. There is insufficient evidence to conclude that the distributions are different. The results of this simulation differ from those in the text; the reason may be due to different sample sizes or simply chance.

13.7 (a) $H_0: p_0 < p_1 = p_2 = \ldots = p_9 = 0.1$ vs. H_a: At least one of the *p*'s is not equal to 0.1. (b) Using `randInt (0, 9, 200)` → L_4, we obtained these counts for digits 0 to 9: 19, 17, 23, 22, 19, 20, 25, 12, 27, 16. (e) $X^2 = 8.9$, df = 9, *P*-value = .447. There is no evidence that the sample data were generated from a distribution that is different from the uniform distribution.

13.8 Answers will vary. You should be surprised if you get a significant *P*-value. H_0: The die is fair ($p_1 = p_2 = \ldots = p_6 = 1/6$). H_a: The die is not fair. Use the command `randInt (1, 6, 300)` → L_1 to simulate rolling a fair die 300 times. In our simulation, we obtained the following frequency distribution:

Side	1	2	3	4	5	6
Freq.	57	46	55	54	45	43

The expected counts under H_0: (300)(1/6) = 50 for each side. The test statistic is $X^2 = .98 + .32 + .5 + .32 + .5 + .98 = 3.6$, and the degrees of freedom are $n - 1 = 5$. The *P*-value is $P(X_5^2 > 3.6) = .608$. Since the *P*-value is large, we fail to reject H_0. There is no evidence that the die is not fair.

13.9 (a)

Outcome	H	T
Frequency	78	122
Expected	100	100

H_0: The distribution of heads and tails from spinning a 1982 penny shows equally likely outcomes. H_a: Heads and tails are not equally likely. df = 1 and $X^2 = 4.84 + 4.84 = 9.68$. The *P*-value is $P(X_1^2 > 9.68) = .00186$. Reject H_0 and conclude that spinning a 1982 penny does not produce equally likely results.

(b) We will test H_0: $p = 0.5$ vs. H_a: $p \neq 0.5$, where p = probability of getting tails when the coin is spun. Since $np_0 = n(1 - p_0) = 100 > 10$, the *z*-test for a single proportion may be used. Test statistic $z = 3.111$, *P*-value of test = 0.00186. Reject H_0; heads and tails are clearly not equally likely.

(c) The *p*-values are identical.

13.10 Let $p_1, p_2, \ldots, p_6$ denote the probability of getting a 1, 2, 3, . . ., 6. If the die is fair, then $p_1 = p_2 = \ldots = p_6$. H_0: $p_1 = p_2 = \ldots = p_6$ (die is fair). H_a: The die is "loaded"/unfair. The observed counts for sides 1 – 6 are: 26, 36, 39, 30, 38, 32. The expected counts are (200)(1/6) = 33.33 for each side. df = 6 – 1 = 5, and $X^2 = 1.612 + 0.214 + 0.965 + 0.333 + 0.654 + 0.053 = 3.831$. The *P*-value is $P(X_5^2 > 3.831) = .574$. Since this *P*-value is rather large, we fail to reject H_0, and conclude that there is no evidence that the die is "loaded."

13

Inference for Tables

13.1 (a) (i) $0.20 < P < 0.25$. (ii) $P = 0.235$. (b) (i) $0.02 < P < 0.025$. (ii) $P = 0.0204$. (c) (i) $P > 0.25$. (ii) $P = 0.3172$.

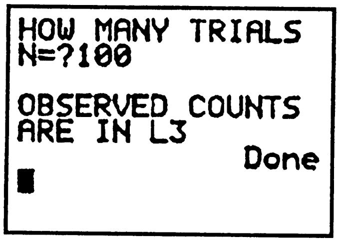
13.2 H_0: The marital-status distribution of 25- to 29-year-old U.S. males is the same as that of the population as a whole. H_a: The marital-status distribution of 25- to 29-year-old U.S. males is different from that of the population as a whole. Expected counts: 140.5, 281.5, 32, 46. $X^2 = 161.77$, df = 3. P-value $= 7.6 \times 10^{-35} \approx 0.0000$. Reject H_0. The two distributions are different.

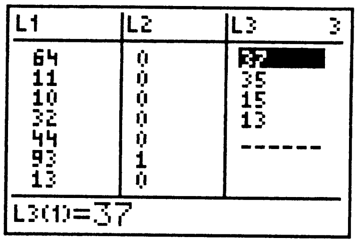
13.3 H_0: The genetic model is valid (the different colors occur in the stated ratio of 1:2:1). H_a: The genetic model is not valid. Expected counts: 21 GG, 42 Gg, 21 gg. $X^2 = 5.43$, df = 2. P-value $= P(X_2^2 > 5.43) = 0.0662$. There is no compelling reason to reject H_0 (though the P-value is a little on the low side).

13.4 H_0: The ethnicity distribution of the Ph.D. degree in 1994 is the same as it was in 1981. H_a: The ethnicity distribution of the Ph.D. degree in 1994 is different from the distribution in 1981. Expected counts = 300 × (1981 percents) = 237, 12, 4, 8, 1, 38. $X^2 = 61.98$, df = 5. P-value $= P(X_5^2 > 61.98) = 4.734 \times 10^{-12} \approx 0.0000$. We reject H_0 and conclude that the ethnicity distribution of the Ph.D. degree has changed from 1981 to 1994. (b) The greatest change is that many more nonresident aliens than expected received the Ph.D. degree in 1994 over the 1981 figures. To a lesser extent, a smaller proportion of white, non-Hispanics received the Ph.D. degree in 1994.

13.5 Use a χ^2 goodness of fit test. (b) Use a one-proportion z test. (c) You can construct the interval; however, your ability to generalize may be limited by the fact that your sample of bags is not an SRS. M&M's may be packaged by weight rather than count.

13.6 H_0: The age-group distribution in 1996 is the same as the 1980 distribution. H_a: The age-group distribution in 1996 is different from the 1980 distribution. One simulation produced observed counts: 37, 35, 15, 13. The expected counts: 41.39, 27.68, 19.64, and 11.28 are stored in list L_4, and the difference terms $(O - E)^2/E$ are assigned to L_5.





used in order to ensure that all geographical areas of the state are equally likely to contain sampled individuals.

(b) $\hat{p} = .02025$, and since $n\hat{p} = 34$ and $n(1 - \hat{p}) = 1645$ are both greater than 10, the confidence interval based on z can be used. The 95% confidence interval for p is $.02025 \pm (1.96)\sqrt{((.02025)(.97975)/1679)} = .02025 \pm 0.00674$, or 0.01351 to 0.02699.

(c) Letting p_1 = the proportion of freshmen who have used steroids and p_2 = the proportion of seniors who have used steroids, we test H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$. From the data, $\hat{p}_1 = .02025, \hat{p}_2 = .01057$, and $\hat{p} = .01905$. $n_1\hat{p}$, $n_1(1 - \hat{p})$, $n_2\hat{p}$, and $n_2(1 - \hat{p})$ are all greater than 5, so a normal approximation can be used. Test statistic $z = (.02025 - .01757)/\sqrt{((.01905)(.98095)(1/1679 + 1/1366)} = 0.5382$, and the P-value = 0.59. There is no reason to reject H_0; the difference between p_1, p_2 is not significant.

12.37 (a) $\hat{p} \doteq 0.1486$ and $SE_{\hat{p}} = \sqrt{(0.1486)(0.8514)/148} \doteq 0.02923$, so the 95% confidence interval is $0.1486 \pm (1.96)(0.02923) \doteq 0.0913$ to 0.2059.

(b) $n = (\frac{1.96}{0.04})^2(0.1486)(0.8514) \doteq 303.7$—use $n = 304$. (We should not use $p^* = 0.5$ here since we have evidence that the true value of p is not in the range 0.3 to 0.7.)

(c) Aside from the 45% nonresponse rate, the sample comes from a limited area in Indiana, focuses on only one kind of business, and leaves out any businesses not in the Yellow Pages (there might be a few of these; perhaps they are more likely to fail). It is more realistic to believe that this describes businesses that match the above profile; it *might* generalize to food-and-drink establishments elsewhere, but probably not to hardware stores and other types of business.

12.38 H_0: $p_1 = p_2$ vs. H_a: $p_1 \neq p_2$; $P = 0.6981$—insufficient evidence to reject H_0.

12.39 (a) $\hat{p}_m = 0.1415, \hat{p}_w = 0.1667$; $P = 0.6981$. (b) $z = 2.12$, $P = 0.0336$. (c) From (a): -0.1056 to 0.1559. From (b): 0.001278 to 0.049036. The larger samples make the margin of error (and thus the length of the confidence interval) smaller.

12.40 For testing H_0: $p = 1/3$ versus H_a: $p > 1/3$, we have $\hat{p} \doteq 0.3786$, and the test statistic is $z = (0.3786 - 1/3)/\sqrt{\frac{(1/3)(2/3)}{803}} \doteq 2.72$. This gives $P = 0.0033$—very strong evidence that more than one-third of this group never use condoms.

12.41 (a) 0.2465 to 0.3359—since 0 is not in this interval, we would reject H_0: $p_1 = p_2$ at the 1% level (in fact, P is practically 0). (b) No: $t = -0.8658$, which gives a P-value close to 0.4.

12.42 No—the data is not based on an SRS, and thus the z procedures are not reliable in this case. In particular, a voluntary response sample is typically biased.

12.43 $\hat{p} = \frac{427}{3160} \doteq 0.1351$, and $SE_{\hat{p}} = \sqrt{\hat{p}(1 - \hat{p})/3160} \doteq 0.006081$, so the 99% confidence interval is $0.1351 \pm (2.576)(0.006081) = 0.1194$ to 0.1508.

12.44 To test H_0: $p_1 = p_2$ vs. H_a: $p_1 < p_2$, we find $\hat{p}_1 = \frac{40}{244} \doteq 0.1639$, $\hat{p}_2 = \frac{87}{245} \doteq 0.3551$, and the pooled value $\hat{p} = \frac{40 + 87}{244 + 245} \doteq 0.2597$. Then $SE = \sqrt{\hat{p}(1 - \hat{p})(\frac{1}{244} + \frac{1}{245})} \doteq 0.03966$, so $z = (\hat{p}_1 - \hat{p}_2)/SE \doteq -4.82$. This gives a tiny P-value (7.2×10^{-7}), so we conclude that bupropion increases the success rate.

12.45 (a) $p_0 = \frac{143{,}611}{181{,}535} \doteq 0.7911$. (b) $\hat{p} = \frac{339}{870} \doteq 0.3897$, $\sigma_{\hat{p}} \doteq 0.0138$, and $z = (\hat{p} - p_0)/\sigma_{\hat{p}} \doteq -29.1$, so $P \doteq 0$ (regardless of whether H_a is $p < p_0$ or $p \neq p_0$). This is very strong evidence against H_0; we conclude that Mexican Americans are underrepresented on juries. (c) $\hat{p}_1 = \frac{339}{870} \doteq 0.3897$, while $\hat{p}_2 = \frac{143{,}611 - 339}{181{,}535 - 870} \doteq 0.7930$. Then $\hat{p} \doteq 0.7911$ (the value of p_0 from part (a)), $s_p = 0.0138$, and $z \doteq -29.2$—and again, we have a tiny P-value and reject H_0.