(b)

	Band Color			Total
	Clear	Dark	Unreadable	
February	20.45	59.09	20.45	100
March	17.65	73.53	8.82	100

(c) A variety of software packages can be used, or the problem can be solved by hand. R gives the following output.

```
        Pearson's Chi-squared test

data:  Month and Color
X-squared = 2.3766, df = 2, p-value = 0.3047
```

Since the P-value is greater than $\alpha = 0.10$, there is no statistically significant evidence that distribution of band colors is different across the two months.

10.5.7 H_0: There is no association between treatment group and condition
H_A: Treatment group and condition are related

	No response	Moderate response	Marked response	Remission	Total
Fluvoxamine	15 (17.831)	7 (6.747)	3 (2.892)	15 (12.530)	40
Placebo	22 (19.169)	7 (7.253)	3 (3.108)	11 (13.470)	43
Total	37	14	6	26	83

The test statistic is $\chi^2_s = 1.83$. The degrees of freedom are df = (2 - 1)(4 - 1) = 3. From Table 9 we find that $\chi^2_{3,0.20} = 4.64$. Thus, P > 0.20, so we do not reject H_0. There is insufficient evidence (P > 0.20) to conclude that treatment group and condition are related.

10.5.8 (a) H_0: There is no association between treatment group and condition
H_A: Treatment group and condition are related

(b) The degrees of freedom are (2 – 1)*(4 – 1) = 3.

(c) We do not reject H_0. There is little or no evidence (P=0.87) to conclude that treatment group and condition are related.

10.5.9 (a) H_0: The chance of needing penicillin is the same for all four treatment groups; H_A: The chance of needing penicillin depends on which group a patient is in.

(b) $e_{11} = 12*55/210 = 3.14$.

(c) We retain H_0 because the P-value is larger than 0.05. There is no evidence (P=0.90) that the chance of needing penicillin depends on group membership.

10.5.10 (a) H_0: The probability of improvement is the same for all four treatment groups; H_A: The probability of improvement depends on which group a patient is in.

(b) The four percentages are, in order, 65%, 65%, 59%, and 50%.

(c) We reject H_0 because the P-value is smaller than 0.05. There is sufficient evidence (P=0.03) to conclude that the chance of improvement depends on group membership.

10.5.11 (a) We retain H_0 because the P-value is larger than 0.05. There is no evidence (P=0.4551) that the chance of needing penicillin depends on group membership.

(b) There is evidence that patients respond better to acupuncture than to usual care for chronic back pain. However, there is no evidence that real acupuncture works better than simulated acupuncture.

10.6.1 Yes, because the expected frequencies (8.3, 5.7, 7.7, and 5.3) all exceed 5.

• **10.6.2** This analysis is not appropriate because the observational units (mice) are nested within the units (litters) that were randomly allocated to treatments. This hierarchical structure casts doubt on the condition that the observations on the 224 mice are independent, especially in light of the investigator's comment that the response varied considerably from litter to litter.

10.6.3 Flaw 1: The sugar observations are not independent of the starch observations, because they were measured on the same people; thus, the two samples are paired. Flaw 2: The 110 sugar observations are not independent of each other because there were 11 observations on each person (hierarchical structure); similarly, the 110 starch observations are not independent of each other.

10.7.1 $\tilde{p}_1 = (139+1)/(1062+2) = 0.1316$, $\tilde{p}_2 = (92+1)/(1065+2) = 0.0872$

$$SE_{(\tilde{p}_1-\tilde{p}_2)} = \sqrt{\frac{(0.1316)(0.8684)}{1062+2} + \frac{(0.0872)(0.9128)}{1065+2}} = 0.0135.$$

$(0.1316 - 0.0872) \pm (1.96)(0.0135)$

$(0.018, 0.071)$ or $0.018 < p_1 - p_2 < 0.071$.

10.7.2 (a) $\tilde{p}_1 = 20/51 = 0.3922$, $\tilde{p}_2 = 9/15 = 0.6000$

$$SE_{(\tilde{p}_1-\tilde{p}_2)} = \sqrt{\frac{(0.3922)(0.6078)}{51} + \frac{(0.6)(0.4)}{15}} = 0.1438.$$

$(0.3922 - 0.6000) \pm (1.96)(0.1438)$

$(-0.49, 0.07)$ or $-0.49 < p_1 - p_2 < 0.07$.

(b) We are 95% confident that the probability of liver tumor under the germ-free condition is between 0.49 lower and 0.07 higher than the probability of liver tumor under the *E. coli.* condition.

• **10.7.3** $\tilde{p}_1 = 33/107 = 0.3084$, $\tilde{p}_2 = 21/109 = 0.1927$

$$SE_{(\tilde{P}_1-\tilde{P}_2)} = \sqrt{\frac{(0.3084)(0.6916)}{107} + \frac{(0.1927)(0.8073)}{109}} = 0.0585.$$

$(0.3084 - 0.1927) \pm (1.96)(0.0585)$

(0.001,0.230) or $0.001 < p_1 - p_2 < 0.230$. No; the confidence interval suggests that bed rest may actually be harmful.

10.7.4 There is hierarchical structure in the data; the observations on two babies born to a woman are not independent of one another.

• **10.7.5 (a)** $\tilde{p}_1 = 912/1657 = 0.5504$, $\tilde{p}_2 = 4579/10002 = 0.4578$.

$$SE_{(\tilde{P}_1-\tilde{P}_2)} = \sqrt{\frac{(0.5504)(0.4496)}{1657} + \frac{(0.4578)(0.5422)}{10002}} = 0.0132.$$

$(0.5504 - 0.4578) \pm (1.96)(0.0132)$

(0.067,0.118) or $0.067 < p_1 - p_2 < 0.118$.

(b) We are 95% confident that the proportion of persons with type O blood among ulcer patients is higher than the proportion of persons with type O blood among healthy individuals by between 0.067 and 0.118. That is, we are 95% confident that p_1 exceeds p_2 by between 0.067 and 0.118.

10.7.6 $\tilde{p}_1 = 31/73 = 0.4247$, $\tilde{p}_2 = 21/72 = 0.2917$.

$$SE_{(\tilde{P}_1-\tilde{P}_2)} = \sqrt{\frac{(0.4247)(0.5753)}{73} + \frac{(0.2917)(0.7083)}{72}} = 0.0788.$$

$(0.4247 - 0.2917) \pm (1.96)(0.0788)$

(-0.021,0.287) or $-0.021 < p_1 - p_2 < 0.287$.

10.7.7 We are 95% confident that taking tocilizumab causes marked improvement in juvenile arthritis for somewhere between 43% and 75% more patients than does taking a placebo.

• **10.8.1** The data are

		Case	
		No	Yes
Control	No	107	30
	Yes	13	5

Note: There is an error in the first printing of the 2nd edition, whereby the headings of "Yes" and "No" were reversed.

The hypotheses are

H_0: There is no association between oral contraceptive use and stroke ($p = 0.5$)
H_A: There is an association between oral contraceptive use and stroke ($p \neq 0.5$)

where p denotes the probability that a discordant pair will be Yes(case)/No(control).

The test statistic for McNemar's test is $\chi^2_s = \frac{(13-30)^2}{13+30} = 6.72$.

Looking in Table 9, with df = 1, we see that $\chi^2_{0.01} = 6.63$ and $\chi^2_{0.001} = 10.83$. Thus, $0.001 < P < 0.01$, so we reject H_0. There is sufficient evidence ($0.001 < P < 0.01$) to conclude that stroke victims are more likely to be oral contraceptive users ($p > 0.5$).

10.8.2 (a) H_0: The probability that gestational age is less than 38 weeks is the same for older siblings as it is for younger siblings.

(b) $\chi^2_S = \frac{(21-5)^2}{21+5} = 9.85$.

(c) We reject H0. There is sufficient evidence (P=0.0017) to conclude that the probability that gestational age is less than 38 weeks is lower for older siblings than it is for younger siblings.

10.8.3 H_0: Among discordant pairs, Pr{"yes/no"} = Pr{"no/yes"} = 0.5
H_A: Having a tonsillectomy increases the risk of Hodgkin's disease
$\chi^2_s = \frac{(7-15)^2}{7+15} = 2.91$. With df = 1, Table 9 gives $\chi^2_{0.10} = 2.71$ and $\chi^2_{0.05} = 3.84$. The data deviate from H_0 in the direction specified by H_A. Thus, $0.025 < P < 0.05$, so we reject H_0. There is sufficient evidence ($0.025 < P < 0.05$) to conclude that having a tonsillectomy increases the risk of Hogkin's disease.

10.8.4 H_0: The probability of copulating is the same for winners and losers
H_A: The probability of copulating is greater for winners
$\chi^2_s = \frac{(8-42)^2}{8+42} = 23.12$. With df = 1, Table 9 gives $\chi^2_{0.0001} = 15.14$. The data deviate from H_0 in the direction specified by H_A. Thus, $P < 0.00005$, so we reject H_0. There is very strong evidence ($P < 0.00005$) that winners copulate more often than losers.

• **10.9.1 (a) (i)** $\hat{p}_1 = 25/517 = 0.04836$ and $\hat{p}_2 = 23/637 = 0.0361$. The relative risk is $\hat{p}_1/\hat{p}_2 = 0.04836/0.03611 = 1.339$.

(ii) The odds ratio is $\frac{(25)(614)}{(23)(492)} = 1.356$.

(b) (i) $\hat{p}_1 = 12/105 = 0.11429$ and $\hat{p}_2 = 8/92 = 0.08696$. The relative risk is $\hat{p}_1/\hat{p}_2 = 0.11429/0.08696 = 1.314$.

(ii) The odds ratio is $\frac{(12)(84)}{(8)(93)} = 1.355$.

10.9.2 (a) (i) $\hat{p}_1 = 14/336 = 0.04167$ and $\hat{p}_2 = 16/428 = 0.03738$. The relative risk is $\hat{p}_1/\hat{p}_2 = 0.04167/0.03738 = 1.115$.

(ii) The odds ratio is $\frac{(14)(412)}{(16)(322)} = 1.120$.

(b) (i) $\hat{p}_1 = 15/353 = 0.04249$ and $\hat{p}_2 = 7/89 = 0.07865$. The relative risk is $\hat{p}_1/\hat{p}_2 = 0.04249/0.07865 = 0.540$.

(ii) The odds ratio is $\frac{(15)(82)}{(7)(338)} = 0.520$.

10.9.3 $\hat{p}_1 = 3995/46941 = 0.0851$ and $\hat{p}_2 = 221/5228 = 0.0423$. The relative risk is $\hat{p}_1/\hat{p}_2 = 0.0851/0.0423 = 2.0131$. Golden retrievers are about two times more likely to suffer from hip dysplasia than are border collies.

10.9.4 (a) The sample odds ratio is $\frac{(3995)(5007)}{(42946)(221)} = 2.11$.

(b) $\ln(\hat{\theta}) = 0.7467$, $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{3995} + \frac{1}{221} + \frac{1}{42946} + \frac{1}{5007}} = 0.071$. The 95% confidence interval for $\ln(\theta)$ is $0.7464 \pm (1.96)(0.071)$, which is (0.6072,0.8856). $e^{0.6072} = 1.84$ and $e^{0.8856} = 2.42$. The 95% confidence interval for θ is (1.84, 2.42).

(c) We are 95% confident that the odds of hip dysplasia for golden retrievers is between 1.84 and 2.42 times greater than the odds of hip dysplasia for border collies. Because hip dysplasia is somewhat rare we could also say that we are *approximately* 95% confident that golden retrievers are between 1.84 and 2.42 times as likely to get hip dysplasia than are border collies.

10.9.5 (a) The sample odds ratio is $\frac{(210)(421502)}{(4391)(33724)} = 0.598$.

(b) The fact that the odds ratio is less than 1.0 means that self-employed workers are less likely to be injured than are persons who work for others.

(c) $\ln(\hat{\theta}) = -0.514$, $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{210} + \frac{1}{4391} + \frac{1}{33724} + \frac{1}{421502}} = 0.071$. The 95% confidence interval for $\ln(\theta)$ is $-0.514 \pm (1.96)(0.071)$, which is (-0.653,-0.375). $e^{-0.653} = 0.52$ and $e^{-0.375} = 0.69$. The 95% confidence interval for θ is (0.52,0.69).

(d) We are 95% confident that being self-employed decreases the odds of occupational injury by a factor of between 0.52 and 0.69. Since occupational injury is rare, we can say that we are 95% confident that the probability of a self-employed person being injured is between 0.52 and 0.69 times the probability of injury for a person employed by others.

10.9.6 (a) The sample odds ratio is $\frac{(6)(1375)}{(1)(696)} = 11.85$.

(b) $\ln(\hat{\theta}) = 2.472$; $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{6} + \frac{1}{1} + \frac{1}{696} + \frac{1}{1375}} = 1.081$. The 95% confidence interval for $\ln(\theta)$ is $2.472 \pm (1.96)(1.081)$, which is (0.353,4.591). $e^{0.353} = 1.42$ and $e^{4.591} = 98.57$. The 95% confidence interval for θ is (1.24,98.57).

(c) The numbers are small, but the confidence interval takes this into account. The confidence interval does not include 1, so we can be quite confident that taking phenypropanolamine as associated with an increase in the odds of a stroke. (This increase is estimated to be by a factor of between 1.42 and 98.57. Since a stroke is fairly rare, we can say that we are 95% confident that the probability of a stroke is between 1.42 and 98.57 times higher for patients taking phenypropanolamine than for other patients.)

• **10.9.7 (a)** The sample odds ratio is $\frac{(309)(1341)}{(266)(1255)} = 1.2413$.

(b) $\ln(\hat{\theta}) = 0.2161$; $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{309} + \frac{1}{266} + \frac{1}{1255} + \frac{1}{1341}} = 0.0924$. The 95% confidence interval for $\ln(\theta)$ is $0.2161 \pm (1.96)(0.0924)$, which is (0.0350,0.3972). $e^{0.0350} = 1.036$ and $e^{0.3972} = 1.488$. The 95% confidence interval for θ is (1.036,1.488).

(c) We are 95% confident that taking heparin increases the odds of a negative response by a factor of between 1.036 and 1.488 when compared to taking enoxaparin. Since a negative outcome is fairly rare, we can say that we are 95% confident that the probability of a negative outcome is between 1.036 and 1.488 times higher for patients given heparin than for patients given enoxaparin.

10.9.8 The sample odds ratio is $\frac{(923)(92)}{(139)(973)} = 0.6279$.

$\ln(\hat{\theta}) = -0.4654$; $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{139} + \frac{1}{923} + \frac{1}{92} + \frac{1}{973}} = 0.1420$. The 95% confidence interval for $\ln(\theta)$ is $-0.4654 \pm (1.96)(0.1420)$, which is (-0.7438,-0.1870). $e^{-0.7438} = 0.475$ and $e^{-0.1870} = 0.829$. The 95% confidence interval for θ is (0.475,0.829).

10.9.9 (a) $\hat{\theta} = \frac{64 * 24084}{50 * 60978} = 0.5055$.

(b) $\log(\hat{\theta})$ is -0.682; the SE for $\log(\hat{\theta})$ is $\sqrt{\frac{1}{50}+\frac{1}{24084}+\frac{1}{64}+\frac{1}{60978}} = 0.1889$. The endpoints of the CI in log scale are -0.682 – 0.1889*1.96 = -1.052 and -0.682 + 0.1889*1.96 = -0.312. This gives a confidence interval of 0.35 to 0.73.

(c) We are 95% confident that taking folic acid during pregnancy is associated with a decreased risk of the child developing autism of a factor between 0.35 and 0.73. Roughly speaking, the odds of autism are only one-third to three-fourths as great if the women took folic acid than if she didn't. (Note that this was an observational study so we cannot attribute causation.)

10.S.1 **(a)** H_0: Pr{CHD | Intervention} = Pr{CHD | Control}; H_A: Pr{CHD | Intervention} ≠ Pr{CHD | Control}.

(b) We retain H0. There is insufficient evidence (P=0.47) to conclude that the intervention has an effect on CHD.

10.S.2 Let p_1 = Pr{CHD|intervention} and p_2 = Pr{CHD|control}. Then $\tilde{p}_1$= (1000+1)/(19541+2) = 0.0512, $\tilde{p}_2$= (1549+1)/(29294+2)= 0.0529.

$$SE_{(\tilde{p}_1-\tilde{p}_2)} = \sqrt{\frac{(0.0512)(0.9488)}{19541+2}+\frac{(0.0529)(0.9471)}{29294+2}} = 0.0020.$$

(0.0512 – 0.0529) ± (1.96)(0.0020)

(-0.0057,0.0023) or $-0.0057 < p_1 - p_2 < 0.0023$.

• **10.S.3** Let p denote the probability of female and let 1 and 2 denote warm and cold environments.

(a) H_0: The sex ratio is 1:1 in the warm environment ($p_1 = 0.5$)
H_A: Sex ratio is not 1:1 in the warm environment ($p_1 \neq 0.5$)

The expected frequencies are calculated from H_0 as follows:

Female: E = (0.5)(141) = 70.5
Male: E = (0.5)(141) = 70.5

The observed and expected frequencies (in parentheses) are

Female	Male	Total
73 (70.5)	68 (70.5)	141

The χ^2 statistic is

$$\chi^2_s = \frac{(73-70.5)^2}{70.5} + \frac{(68-70.5)^2}{70.5} = 0.18.$$

There are two categories (female and male), so we consult Table 9 with df = 2 - 1 = 1. From Table 9, we find $\chi^2_{1,0.20} = 1.64$. Because $\chi^2_s < \chi^2_{0.20}$, the P-value is bracketed as

P > 0.20.

At significance level $\alpha = 0.05$, we reject H_0 if P < 0.05. Since P > 0.20, we do not reject H_0. There is insufficient evidence (P > 0.20) to conclude that the sex ratio is not 1:1 in the warm environment.

(b) H_0: The sex ratio is 1:1 in the cold environment ($p_2 = 0.5$)
H_A: Sex ratio is not 1:1 in the cold environment ($p_2 \neq 0.5$)
The expected frequencies are calculated from H_0 as follows:

Female: E = (0.5)(169) = 84.5
Male: E = (0.5)(169) = 84.5

The observed and expected frequencies (in parentheses) are

Female	Male	Total
107 (84.5)	62 (84.5)	169

The χ^2 statistic is

$$\chi^2_s = \frac{(107-84.5)^2}{84.5} + \frac{(62-84.5)^2}{84.5} = 11.98.$$

From Table 9, we find $\chi^2_{1,0.001} = 10.83$ and $\chi^2_{1,0.0001} = 15.14$. Thus, 0.0001 < P < 0.001, so we reject H_0. There is sufficient evidence (0.0001 < P < 0.001) to conclude that the cold environment produces more females than males.

(c) The hypotheses are

H_0: Sex ratio is the same in the two environments ($p_1 = p_2$)
H_A: Sex ratio is not the same in the two environments ($p_1 \neq p_2$)

We calculate the expected frequencies under H_0 from the formula

$$E = \frac{\text{(Row total)x(Column total)}}{\text{Grand total}}$$

The following table shows the observed and expected frequencies (in parentheses):

	Environment		
	Warm	Cold	Total
Male	68 (59.13)	62 (70.87)	130
Female	73 (81.87)	107 (98.13)	180
Total	141	169	310

The χ^2 test statistic is

$$\chi^2_s = \frac{(68-59.13)^2}{59.13} + \frac{(62-70.87)^2}{70.87} + \frac{(73-81.87)^2}{81.87} + \frac{(107-98.13)^2}{98.13} = 4.20.$$

The degrees of freedom are df = (2 - 1)(2 - 1) = 1. Table 9 shows that $\chi^2_{1,0.05} = 3.84$ and $\chi^2_{1,0.02} = 5.41$. Thus, $0.02 < P < 0.05$, so H_0 is rejected. To determine directionality, we calculate

$$\hat{p}_1 = \frac{73}{141} = 0.52,$$
$$\hat{p}_2 = \frac{107}{169} = 0.63,$$

and we note that $\hat{p}_1 < \hat{p}_2$.

There is sufficient evidence ($0.02 < P < 0.05$) to conclude that the probability of a female is higher in the cold than the warm environment.

(d) Since all eggs were from one mating, there is no basis for generalizing to other *Menidia* individuals (and in fact not all individuals exhibit the phenomenon). The population can therefore be defined as all (potential) offspring of the single mating, or perhaps of matings of the same genotype. (Strictly speaking, the population should also be limited to offspring who survive long enough for their sex to be observable.)

10.S.4 No, the proposed analysis would not be valid because the observations on cilia from the same child are not independent; there is hierarchical structure in the data.

10.S.5 The second contingency table is relevant. The null and alternative hypotheses are

H_0: $p_1 = p_2$
H_A: $p_1 \neq p_2$

where p denotes the probability of guessing drug, 1 denotes receiving the drug, and 2 denotes receiving the placebo. With df = 1, Table 9 gives $\chi^2_{0.20} = 1.64$. Thus, $P > 0.20$ and we do not reject H_0. There is little or no evidence ($P > 0.20$) to conclude that the climbers received clues. [Remark: These data illustrate an interesting psychological phenomenon. Regardless of which treatment they actually received, about 65% of the people thought they received the drug ($\hat{p}_1 = 20/31 \approx 0.65$ and $\hat{p}_2 = 21/33 \approx 0.64$).]

10.S.6 The hypotheses are

H_0: Temperature does not affect survival
H_A: High temperature enhances survival

$$\frac{({}_{20}C_2)({}_{28}C_{10})}{{}_{48}C_{12}} \approx 0.03579$$

Tables of possible outcomes that more strongly support H_A are

19	1
17	11

$$\frac{({}_{20}C_1)({}_{28}C_{11})}{{}_{48}C_{12}} \approx 0.00616$$

20	0
16	12

$$\frac{({}_{20}C_0)({}_{28}C_{12})}{{}_{48}C_{12}} \approx 0.00044$$

Thus, P = 0.03579 + 0.00616 + 0.00044 = 0.04239 and we reject H_0. There is sufficient evidence (P = 0.042) to conclude that high temperature increases the probability of survival.

10.S.7 Let p denote the probability of survival, let 1 denote 38°, and let 2 denote 40°.

H_0: Temperature does not affect survival ($p_1 = p_2$)
H_A: High temperature enhances survival ($p_1 < p_2$)

	38°	40°	Total
Died	18 (15)	2 (5)	20
Survived	18 (21)	10 (7)	28
Total	36	12	48

$\chi^2_s = 4.11$; df = 1. From Table 9, we find $\chi^2_{1,0.05} = 3.84$ and $\chi^2_{1,0.02} = 5.41$. $\hat{p}_1 = 18/36 = 0.5$ and $\hat{p}_2 = 10/12 \approx 0.83$; the data deviate from H_0 in the direction specified by H_A. Thus, $0.01 < P < 0.025$ and we reject H_0. There is sufficient evidence ($0.01 < P < 0.025$) to conclude that high temperature increases the probability of survival.

10.S.8 (a) The sample proportions for Died are $\widehat{\Pr}\{Die|Chem\} = \frac{78}{84} \approx 0.51$ and $\widehat{\Pr}\{\text{Die}|\text{Radiation}\} = \frac{66}{164} \approx 0.40$.

(b) The expected cell counts under the null hypothesis are $e_{11} = 69.74$; $e_{12} = 74.26$; $e_{21} = 84.26$; $e_{22} = 89.74$

(c) We retain H_0 because the P-value, 0.0625, is larger than 0.05. We have insufficient evidence (P=0.0625) to conclude that type of treatment affects survival.

(d) In this case the P-value would have been 0.0625/2 = 0.03125 and H_0 would have been rejected at the 0.05 level. [Note from part (a) that the data point in the direction predicted by the directional alternative.]

10.S.9 (a) The sample odds ratio is $\frac{(78)(98)}{(66)(76)} = 1.524$.

(b) $\ln(\hat{\theta}) = .4213$; $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{78}+\frac{1}{66}+\frac{1}{76}+\frac{1}{98}} = 0.2266$. The 95% confidence interval for $\ln(\theta)$ is $0.4213 \pm (1.96)(0.2266)$, which is (-0.023,0.865). $e^{-0.023} = 0.98$ and $e^{0.865} = 2.38$. The 95% confidence interval for θ is (0.98,2.38) or $0.98 < \theta < 2.38$.

10.S.10 The hypotheses are

H_0: Survival and treatment are independent
H_A: Survival depends on treatment

	Zidovudine	Didanosine	Zid. and Did.	Total
Died	17 (11.29)	7 (11.50)	10 (11.21)	34
Survived	259 (264.71)	274 (269.50)	264 (262.79)	797
Total	276	281	274	831

$\chi^2_s = 4.98$. With df = (2 - 1)(3 - 1) = 2, Table 9 gives $\chi^2_{0.10} = 4.61$ and $\chi^2_{0.05} = 5.99$, so $0.05 < P < 0.10$. We do not reject H_0; there is insufficient evidence ($0.05 < P < 0.10$) to conclude that survival depends on treatment.

10.S.11 (a) The expected cell count is 476*367/1764 = 99.03.

(b) There are 3 degrees of freedom.

(c) We reject H_0 because the P-value is smaller than 0.05. We have strong evidence (P=1.26e-07) to conclude that blood type distributions are not the same for malaria patients and controls. There is strong evidence of a relationship between blood type and malaria status.

• **10.S.12** The null and alternative hypotheses are

H_0: Site of capture and site of recapture are independent ($\Pr\{RI|CI\} = \Pr\{RI|CII\}$)
H_A: Flies preferentially return to their site of capture ($\Pr\{RI|CI\} > \Pr\{RI|CII\}$)

where C and R denote capture and recapture and I and II denote the sites.

Because H_A is directional, we begin by checking the directionality of the data. We calculate

$$\hat{P}r\{RI|CI\} = \frac{78}{134} = 0.58,$$
$$\hat{P}r\{RI|CII\} = \frac{33}{91} = 0.36,$$

and we note that $\hat{P}r\{RI|CI\} > \hat{P}r\{RI|CII\}$.

Thus, the data deviate from H_0 in the direction specified by H_A.

The test statistic is $\chi^2_s = 10.44$. From Table 9 with df = 1, we find $\chi^2_{1,0.01} = 6.63$ and $\chi^2_{1,0.001} = 10.83$. We cut the column headings in half for the directional test. Thus, $0.0005 < P < 0.005$ and we reject H_0. At the 0.01 level, there is sufficient evidence ($0.0005 < P < 0.005$) to conclude that flies preferentially return to their site of capture.

10.S.13 (a)

	Y	G	Total
R	900	300	1200
W	300	100	400
Total	1200	400	1600

$\hat{P}r\{Y\} = 1200/1600 = 3/4$; $\hat{P}r\{R\} = 1200/1600 = 3/4$;
$\hat{P}r\{Y|R\} = 900/1200 = 3/4$; $\hat{P}r\{Y|W\} = 300/400 = 3/4$.

(b) There is no single correct answer. One typical answer:

	Y	G	Total
R	1000	200	1200
W	200	200	400
Total	1200	400	1600

$\hat{P}r\{Y\} = 1200/1600 = 3/4$; $\hat{P}r\{R\} = 1200/1600 = 3/4$;
$\hat{P}r\{Y|R\} = 1000/1200 = .83$; $\hat{P}r\{Y|W\} = 200/400 = .5$.

(c) There is no single correct answer. One typical answer:

	Y	G	Total
R	875	525	1400
W	125	75	200
Total	1000	600	1600

$\hat{P}r\{Y\} = 1000/1600 = .625$; $\hat{P}r\{R\} = 1400/1600 = .875$;
$\hat{P}r\{Y|R\} = 875/1400 = .625$; $\hat{P}r\{Y|W\} = 125/200 = .625$.

• **10.S.14** (a) The sample odds ratio is $\frac{(25339)(686)}{(8914)(1141)} = 1.709$.

(b) $\ln(\hat{\theta}) = .5359$; $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{25339}+\frac{1}{1141}+\frac{1}{8914}+\frac{1}{686}} = 0.0499$. The 95% confidence interval for $\ln(\theta)$ is $0.5359 \pm (1.96)(0.0499)$, which is (0.438,0.634). $e^{0.438} = 1.55$ and $e^{0.634} = 1.89$. The 95% confidence interval for θ is (1.55,1.89) or $1.55 < \theta < 1.89$.

(c) The odds ratio gives the (estimated) odds of survival for men compared to women. Another way to say this is that it gives the odds of death for women compared to men. This ratio (of 1.709) is a good approximation to the relative risk of death for women compared to men (which is 1.658), because death is fairly rare.

10.S.15 We are 95% confident that using any products containing phenypropanolamine is associated with odds of a stroke being only 0.84 times as high up to 2.64 times as high as the odds of a stroke if no products containing phenypropanolamine are used. Since a stroke is fairly rare, we can say that we are 95% confident that the probability of a stroke is between 0.84 and 2.64 times higher for patients using any products containing phenypropanolamine than for other patients. (Note that 1 is within the confidence interval, so we cannot say that using any products containing phenypropanolamine is associated with a change in the likelihood of having a stroke.)

10.S.16 (a) H_0: Atrophied villi are not associated with smoking (Pr{P|N} = Pr{P|M}= Pr{P|H})
H_A: Atrophied villi are associated with smoking (not all the probabilities are equal)

$\chi^2_s = 12.24$; df = 2. P = 0.0022, so we reject H_0. There is sufficient evidence (P = 0.0022) to conclude that atrophied villi are associated with smoking.

(b)

		Women with atrophied villi	
Smoking status	No. women	Number	Percent
N	22	4	18
M	20	9	45
H	16	12	75

(c) The chi-square test does not take into account the fact that the percentage with atrophied villi increases as smoking level increases.

10.S.17 (a) H_0: The probability of an accident is the same when using a cellular phone as it is when not using a cellular phone

(b) $\chi^2_s = \frac{(24-157)^2}{24+157} = 97.7$. Looking in Table 9, with df = 1, we see that $\chi^2_{0.0001} = 15.14$. Thus, $P < 0.0001 < \alpha$, so we reject H_0.

(c) There is very strong evidence (P < 0.0001) that the probability of an accident is higher when using a cellular phone as it is when not using a cellular phone.

[Remark: The odds ratio for these data is 1.74; some people have likened driving while using a cellular phone to driving while being drunk.]

10.S.18 The hypotheses are

H_0: The vaccine has no effect on flu prevention
(Pr{flu | vaccine} = Pr{flu | placebo})
H_A: The vaccine prevents influenza.
(Pr{flu | vaccine} < Pr{flu | placebo})

	Flu	No flu	Total
Vaccine	28 (45.0)	785 (768.0)	813
Placebo	35 (18.0)	290 (307.0)	325
Total	63	1075	1138

$\chi^2_s = 23.824$. With df = 1, Checking the directionality we observe that $\hat{\Pr}$ {flu | vaccine} = 28/813 = 0.0344 and $\hat{\Pr}$ {flu | placebo} = 35/325 = 0.1077, which is consistent with H_A. Table 9 gives $\chi^2_{0.0001} = 15.14$, so $P < 0.00005$. We reject H_0; there is very strong evidence ($P < 0.00005$) to conclude that the vaccine prevents influenza. Note that a causal claim is possible because this study was an experiment.

10.S.19 (a) The sample odds ratio is $\frac{(28)(290)}{(785)(35)} = 0.2955$.

(b) $\ln(\hat{\theta}) = -1.2189$; $SE_{\ln(\hat{\theta})} = \sqrt{\frac{1}{28}+\frac{1}{785}+\frac{1}{35}+\frac{1}{290}} = 0.2627$. The 95% confidence interval for ln(θ) is $-1.2189 \pm (1.96)(0.2627)$, which is (-1.7338,-0.0704). $e^{-1.7338} = 0.1766$ and $e^{-0.0704} = 0.4946$. The 95% confidence interval for θ is (0.18,0.49) or $0.18 < \theta < 0.49$.

10.S.20 (a) H_0: Men and women are equally good at identifying their partners (Pr{correct|M} = Pr{correct|W}); H_A: Men and women are not equally good at identifying their partners (Pr{correct|M} ≠ Pr{correct|W}).

(b) The sample proportions correct are $\widehat{\Pr}\{correct|M\} = \frac{16}{36} \approx 0.44$ and $\widehat{\Pr}\{\text{correct}|W\} = \frac{25}{36} = 0.69$. The expected cell counts under the null hypothesis are $e_{11} = 20.5$; $e_{12} = 20.5$; $e_{21} = 15.5$; $e_{22} = 15.5$.

(c) The P-value is 0.03219. We reject H_0 because the P-value is smaller than 0.05. We have evidence (P=0.032) to conclude that women are better than men at identifying their partners by touching the backs of their hands.

10.S.21 $\hat{p}_{arb} = 54/66 \approx 0.82$ and $\hat{p}_{man} = 42/72 \approx 0.58$. The chi-square statistic for a test of equality is 8.97 with one degree of freedom; the P-value is between 0.001 and 0.01 so there is strong evidence for the alternative hypothesis: the proportion of native species is higher in the arboretum than on managed land.

UNIT III SUMMARY

III.1 (a) $\tilde{p}_1 = \frac{6+1}{53+2} = 0.127$ and $\tilde{p}_2 = \frac{15+1}{56+2} = 0.276$. The SE is $\sqrt{\frac{0.127*0.873}{55} + \frac{0.276*0.734}{58}} =$ 0.074. The confidence interval (CI) is (0.276 x– .127) ± 1.96*(0.074) or (0.004, 0.294).

(b) The CI does not include zero, so we reject H_0.

(c) There is statistically significant evidence that the proportion of women who can touch their nose with their tongue is greater than the proportion of men who can touch their nose with their tongue.

• **III.2 (a)** The two sample proportions are 5/16 = 0.3125 and 18/30 = 0.60; thus, the data support H_A. We have $0.025 < P\text{-value} < 0.05$, so we reject H_0. There is evidence that men are less likely than women to be involved in community service.

(b) We can focus on the upper left cell, which contains a 5. More extreme tables would have 4, 3, 2, 1, or 0 here, so there are 6 tables to consider.

• **III.3 (a)** If germination rate is independent of type of seed, then the best estimate of germination rate is 32/57. Applying that rate to 20 Okra seeds gives 20x32/57 = 11.23 expected to germinate.

(b) $\chi^2 = 1.84$, df = 2. $P > 0.20 > 0.05$, so we retain H_0.

III.4 (a) H_0: $p_1 = p_2$

(b) We can use McNemar's test since the data are paired. The test statistic is a chi-square with 1 df and equals $\frac{(57-13)^2}{57+13} = 27.6$. The P-value is less than 0.0001, so we reject H_0.

(c) There is statistically significant evidence that the disease is associated with oral contraceptive use.

• **III.5 (a)** If $\tilde{p} = 0.25$ then we want $\sqrt{\frac{0.25*0.75}{n+4}} \le 0.06$ so $n+4 \ge \left(\frac{\sqrt{0.25*0.75}}{0.06}\right)^2$ which equals 52.08.

Thus, we want $n \ge 48.08$, so we must take $n = 49$.

(b) If we don't have a guess for $\tilde{p}$ then we use $\tilde{p} = 0.50$. This gives $\sqrt{\frac{0.5*0.5}{n+4}} \le 0.06$ so

$n+4 \ge \left(\frac{\sqrt{0.5*0.5}}{0.06}\right)^2$ which equals 69.44. Thus we want $n \ge 65.44$, so we must take $n = 66$.

III.6 (a) $\tilde{p} = \frac{19+2}{25+4} = 0.7241$ and $SE_{\tilde{p}} = \sqrt{\frac{(0.7241)(1-0.7241)}{25+4}} = 0.0830$

$0.7241 \pm 1.96\times0.0830$

$(0.561, 0.887)$

We are 95% confident that this rat's accuracy rate is between 0.561 and 0.887.

(b) Since 95% confidence interval is entirely above 0.50, we can be confident that the rat is doing better than simply guessing.

(c) 0.70 is inside the 95% confidence interval, so while the rat's accuracy rate may be above 0.70 (as high as 0.887), it could also be below (as low as 0.561).

(d)
$H_0: p = 0.70$
$H_A: p > 0.70$

	CORRECT	INCORRECT	TOTAL
Observed counts	$o_1 = 19$	$o_2 = 6$	25
Expected counts	$e_1 = 25\times0.7 = 17.5$	$e_2 = 25\times0.3 = 7.5$	25

$$\chi_S^2 = \frac{(19-17.5)^2}{17.5} + \frac{(6-7.5)^2}{7.5} = 0.4286$$

Using Table 9 with $df = 1$, the directional P-value is greater than $\frac{0.20}{2} = 0.10$. Thus, there is no statistically significant evidence that the rat's accuracy rate is above 0.7.

III.7 (a) H_0: Pr{DC|EC} = Pr{DC|EW}

(b)

	DC	DW	TOTAL
EC	$\frac{38\times33}{50} = 25.1$	$\frac{38\times17}{50} = 12.9$	38
EW	$\frac{12\times33}{50} = 7.9$	$\frac{12\times17}{50} = 4.1$	12
Total	33	17	50

(c) The observed cell count for the EC/DC cell is 27, which is greater than the expected value under the null hypothesis. Thus, the data deviate in the predicted direction. Using Table 9 with $df = 1$, the directional P-value is between 0.10 and 0.20. Thus, there is no statistically significant evidence that the decoder rat's accuracy rate improves when the encoder rat is accurate.

(d) The expected cell count for the EW/DW cell is only 4.1, which is less than 5.

III.8 (a) H_A: Pr{agree} > 1/2

(b) Under the null hypothesis, we expect 25 agreements and 25 disagreements. Thus

$$\chi_S^2 = \frac{(33-25)^2}{25} + \frac{(17-25)^2}{25} = 5.12$$

Using Table 9 with $df = 1$, the directional P-value is between 0.02 and 0.05. Thus, there is evidence that the two rats tend to agree with each other.

III.9 The 200 observations are not independent of one another. When considering the first pair of rats (as in Exercise III.7) we can think of the 50 observations as being 50 independent measurements of how well those two rats agree. But when we combine the data across four pairs of connected rats, we cannot think of the 200 observations as being 200 draws from a fixed population. Rather, we have four sets of 50 draws from (potentially) four populations.

• **III.10 (a)** True. The standard error of the estimated proportion is largest when p = 0.50.

(b) False. A goodness-of-fit test can be conducted for any null hypothesis that specifies probabilities for each of the possible categories of a response variable, but there is no need for those probabilities to be equal.

(c) False. Either a goodness-of-fit test or a test of independence can be conducted for observational data or for experimental data.

(d) True. If the observed data perfectly agree with the expected values from the null hypothesis, then each term in the chi-square calculation will be zero, so the test statistic will be zero.