9.S.18 The hypotheses are

H_0: The men are guessing (Pr{correct} = 1/3)
H_A: The men have some ability to detect their partners (Pr{correct} > 1/3)

Correct	Wrong
16 (12)	20 (24)

$\chi^2_s = 2$. With df = 1, Table 9 gives $\chi^2_{0.20} = 1.64$ and $\chi^2_{0.10} = 2.71$, so $0.05 < P < 0.10$ and we do not reject H_0. There is insufficient evidence ($0.05 < P < 0.10$) to conclude that the men have some ability to detect their partners by touching the backs of their hands; the data are consistent with guessing.

9.S.19 (a) There are 58 + 26 = 84 total plants. The expected number of resistant plants is 84*0.75 = 63 and the expected number of susceptible plants is 84*0.25 = 21.

(b) We do not reject H_0. There is little or no evidence (P=0.21) to conclude that the hypothesized 3:1 ratio is wrong.

(c) No. While the results are consistent with the 3:1 ratio, the data do not confirm the null hypothesis, they only fail to refute it.

9.S.20 $\tilde{p}$ = (103+2)/(1438+4) = 0.073; SE = $\sqrt{\frac{0.073(1-0.073)}{1438+4}}$ = 0.0069.
The 95% confidence interval is $0.073 \pm (1.96)(0.0069)$ or (0.059,0.087) or $0.059 < p < 0.087$.

9.S.21 (a) Yes, there is compelling evidence of a difference. The confidence interval excludes 0.10 so a claim that p = 10% would be rejected at the 0.05 significance level.

(b) The hypotheses are

H_0: Pr{STD} = 0.10
H_A: Pr{STD} $\neq$ 0.10

STD	No STD
103 (143.8)	1335 (1294.2)

$\chi^2_s = 12.86$. With df = 1, Table 9 gives $\chi^2_{0.001} = 10.83$ and $\chi^2_{0.0001} = 15.14$, so $0.0001 < P < 0.001$ and we reject H_0. There is strong evidence ($0.0001 < P < 0.001$) to conclude that the probability of contracting an STD is not 10% for counseled individuals.

(c) Yes, the two answers agree.

CHAPTER 10
Categorical Data: Relationships

10.2.1 (a)

	Treatment 1	Treatment 2
Success	70	140
Failure	30	60
Total	100	200

(b) $\hat{p}_1$ = 70/100 = 0.7, $\hat{p}_2$ = 140/200 = 0.7; yes, they are equal.

10.2.2 (a)

	Treatment 1	Treatment 2
Success	30	10
Failure	270	90
Total	300	100

(b) $\hat{p}_1$ = 30/300 = 0.1, $\hat{p}_2$ = 10/100 = 0.1; yes, they are equal.

• **10.2.3 (a)** To have $\chi^2_s = 0$, the columns of the table (and the rows of the table) must be proportional to each other, as in the following table:

	Treatment 1	Treatment 2
Success	5	20
Failure	10	40
Total	15	60

(b) The estimated probabilities of success are $\hat{p}_1$ = 5/15 = 1/3 and $\hat{p}_2$ = 20/60 = 1/3. Yes, these proportions are equal.

10.2.4 (a) H_0: The striped and red forms survive equally well
H_A: The red form survives better than does the striped form

(b) H_0: $p_1 = p_2$
H_A: $p_1 < p_2$

where p denotes the probability of survival, 1 denotes striped, and 2 denotes red.

(c-d)

	Striped	Red	Total
Survived	65 (70.31)	23 (17.69)	88
Died	98 (92.69)	18 (23.31)	116
Total	163	41	204

$\chi^2_s = 3.51$; $\hat{p}_1 = 65/163 \approx 0.40$, $\hat{p}_2 = 23/41 \approx 0.56$. With df = 1, Table 9 gives $\chi^2_{0.10} = 2.71$ and $\chi^2_{0.05} = 3.84$, so $0.025 < P < 0.05$.

(e) We reject H_0; there is sufficient evidence ($0.025 < P < 0.05$) to conclude that the red form survives more successfully than does the striped form.

• **10.2.5** The hypotheses are

H_0: Mites do not induce resistance to wilt
H_A: Mites do induce resistance to wilt

Letting p denote the probability of wilt and letting 1 denote mites and 2 denote no mites, the hypotheses may be stated as

H_0: $p_1 = p_2$
H_A: $p_1 < p_2$

Because H_A is directional, we begin by checking the directionality of the data. The estimated probabilities of wilt disease are

$$\hat{p}_1 = \frac{11}{26} \approx 0.42;$$
$$\hat{p}_2 = \frac{17}{21} \approx 0.81.$$

We note that

$$\hat{p}_1 < \hat{p}_2.$$

Thus, the data do deviate from H_0 in the direction specified by H_A. We proceed to the calculation of the test statistic. The expected frequency for any given cell is found 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):

	Mites	No mites	Total
Wilt disease	11 (15.49)	17 (12.51)	28
No wilt disease	15 (10.51)	4 (8.49)	19
Total	26	21	47

The χ^2 test statistic is

$$\chi^2_s = \frac{(11-15.49)^2}{15.49} + \frac{(17-12.51)^2}{12.51} + \frac{(15-10.51)^2}{10.51} + \frac{(4-8.49)^2}{8.49} = 7.21.$$

We consult Table 9 with df = 1. From Table 9, we find $\chi^2_{1,0.01} = 6.63$ and $\chi^2_{1,0.001} = 10.83$, so $\chi^2_{1,0.001} < \chi^2_s < \chi^2_{1,0.01}$. Because H_A is directional, the column headings (0.01 and 0.001) must be cut in half to bracket the P-value; thus P-value is bracketed as

$0.0005 < P < 0.005$.

At significance level 0.01, we reject H_0 if $P < 0.01$. Since $0.0005 < P < 0.005$, we reject H_0. There is sufficient evidence ($0.0005 < P < 0.005$) to conclude that mites do induce resistance to wilt.

10.2.6 The two sample proportions of left-side tumors are 14/33 = 0.424 and 28/55 = 0.509. Thus, left-side tumors are more likely when the phone is held on the right side than when held on the left side. That is, the data show a difference in the opposite direction to that predicted. Thus, the P-value is greater than 0.50 and we do not reject H_0.

10.2.7 (a) (i) H_0: $p_1=p_2$; H_A: $p_1 \neq p_2$.
(ii) $\chi^2_S = 0.1175$.
(iii) We retain H0. There is insufficient evidence (P=0.73) to conclude that the two drugs are not equally effective.

(b) No, part(a) is not conclusive on this point. We do not know whether the study had adequate power to detect a clinically important difference. A power calculation or a confidence interval would help determine whether the drugs can be considered equally effective.

10.2.8 (a) H_0: The two products are equally effective; H_A: The two products are not equally effective.

(b) $e_{11} = 11.5$; $e_{12} = 11.5$; $e_{21} = 9.5$; $e_{22} = 9.5$.

(c) We reject H_0 because the P-value is smaller than 0.05. We have sufficient evidence (P=0.03) to conclude that the two products are not equally effective. Product A is less effective than Product B in producing pregnancy.

(d) $\hat{Pr}\{\text{Pregnant } |A\} = \frac{8}{21} = 0.38$, $\hat{Pr}\{\text{Pregnant } |B\} = \frac{15}{21} = 0.71$. These data are not in agreement with the directional alterniative hypothesis that the pregnancy rate on A is higher than B, thus the P-value for this directional test is greater than 0.50. There is no evidence that A is superior to B.

10.2.9 Let p denote the probability of a tumor, let 1 denote germ-free, and let 2 denote *E. coli*.

H_0: *E. coli* does not affect tumor incidence ($p_1 = p_2$)
H_A: *E. coli* increases tumor incidence ($p_1 < p_2$)

(a)

	Germ-free	*E. coli*	Total
Tumors	19 (21.34)	8 (5.66)	27
No tumors	30 (27.66)	5 (7.34)	35
Total	49	13	62

$\chi^2_s = 2.17$; $\hat{p}_1 = 19/49 \approx 0.39$, $\hat{p}_2 = 8/13 \approx 0.62$. With df = 1, Table 9 gives $\chi^2_{.20} = 1.64$ and $\chi^2_{.10} = 2.71$, so $0.05 < P < 0.10$. We do not reject H_0; there is insufficient evidence ($0.05 < P < 0.10$) to conclude that *E. coli* increases tumor incidence.

(b) (i) $\chi^2_s = (2)(2.17) = 4.34$. With df = 1, Table 9 gives $\chi^2_{0.05} = 3.84$ and $\chi^2_{0.02} = 5.41$, so $0.01 < P < 0.025$ and we reject H_0. There is sufficient evidence ($0.01 < P < 0.025$) to conclude that *E. coli* increases tumor incidence.

(ii) $\chi^2_s = (3)(2.17) = 6.51$. With df = 1, Table 9 gives $\chi^2_{0.02} = 5.41$ and $\chi^2_{0.01} = 6.63$, so $0.005 < P < 0.01$ and we reject H_0. There is sufficient evidence ($0.005 < P < 0.01$) to conclude that *E. coli* increases tumor incidence.

• **10.2.10** Let p denote the probability of response, let 1 denote simultaneous, and let 2 denote sequential administration. The hypotheses are

H_0: The two timings are equally effective ($p_1 = p_2$)
H_A: The two timings are not equally effective ($p_1 \neq p_2$)

The expected frequency for any given cell is found 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):

	Simultaneous	Sequential	Total
Response	11 (8.30)	3 (5.70)	14
No response	5 (7.70)	8 (5.30)	13
Total	16	11	27

The χ^2 test statistic is

$$\chi^2_s = \frac{(11-8.3)^2}{8.3} + \frac{(3-5.7)^2}{5.7} + \frac{(5-7.7)^2}{7.7} + \frac{(8-5.3)^2}{5.3} = 4.48.$$

We consult Table 9 with df = 1. From Table 9, we find $\chi^2_{1,0.05} = 3.84$ and $\chi^2_{1,0.02} = 5.41$, so $\chi^2_{1,0.02} < \chi^2_s < \chi^2_{1,0.05}$. Thus P-value is bracketed as

$0.02 < P < 0.05$.

At significance level 0.05, we reject H_0 if $P < 0.05$. Since $0.02 < P < 0.05$, we reject H_0. To determine directionality, we calculate

$$\hat{p}_1 = \frac{11}{16} \approx 0.69,$$

$$\hat{p}_2 = \frac{3}{11} \approx 0.27,$$

and we note that

$\hat{p}_1 > \hat{p}_2$.

There is sufficient evidence ($0.02 < P < 0.05$) to conclude that the simultaneous timing is superior to the sequential timing.

10.2.11 Let p denote the probability of a hip fracture, let 1 denote the hip protector group, and let 2 denote the control group.

H_0: $p_1 = p_2$
H_A: $p_1 < p_2$

	Hip protector	Control	Total
Hip fracture	13 (29.006)	67 (50.994)	80
No hip fracture	640 (623.994)	1081 (1097.006)	1721
Total	653	1148	1801

$\chi^2_s = 14.5$. With df = 1, Table 9 gives $\chi^2_{0.001} = 10.83$ and $\chi^2_{0.0001} = 15.14$. The difference in sample proportions is consistent with the alternative hypothesis, so $0.0001/2 < P < 0.001/2$. We reject H_0; there is strong evidence ($0.00005 < P < 0.0005$) to conclude that hip protectors reduce the probability of a hip fracture.

10.2.12 Let p denote the probability of a cold, let 1 denote five or fewer types of social relationships, and let 2 denote six or more types of social relationships.

H_0: Number of types of social relationships does not affect cold incidence ($p_1 = p_2$)
H_A: Number of types of social relationships affects cold incidence ($p_1 \neq p_2$)

	Five or fewer	Six or more	Total
Cold	57 (48.58)	52 (60.42)	109
No cold	66 (74.42)	101 (92.58)	167
Total	123	153	276

$\chi^2_s = 4.35$. With df = 1, Table 9 gives $\chi^2_{0.05} = 3.84$ and $\chi^2_{0.02} = 5.41$, so $0.02 < P < 0.05$. We reject H_0; there is sufficient evidence ($0.02 < P < 0.05$) to conclude that being in more types of social relationships decreases the probability of getting a cold.

• **10.2.13** Let p denote the probability of hemorrhaging, let 1 denote the ancrod group, and let 2 denote the placebo group.

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

	Ancrod	Placebo	Total
Yes	13 (8.928)	5 (9.072)	18
No	235 (239.072)	247 (242.928)	482
Total	248	252	500

$\chi^2_s = 3.82$. With df = 1, Table 9 gives $\chi^2_{0.10} = 2.71$ and $\chi^2_{0.05} = 3.84$, so $0.05 < P < 0.10$. We do not reject H_0; there is insufficient evidence ($0.05 < P < 0.10$) to conclude that hemorrhaging is more likely under one treatment than under the other.

10.2.14 **(a)** H_0: There is no relationship between a woman's response to a man and her menstrual phase; H_A: Women are more likely to respond to a man during their fertile phase.

(b) The sample proportions for Yes are $\hat{p}_1 = 13/60 \approx 0.22 = 0.22$ and $\hat{p}_2 = 11/140 \approx 0.08$. The expected cell counts under the null hypothesis are $e_{11} = 7.2$; $e_{12} = 16.8$; $e_{21} = 52.8$; $e_{22} = 123.2$.

(c) The data point in the direction predicted by H_A so the P-value is 0.0059/2 = 0.00295. We reject H_0 because the P-value is smaller than 0.02. We have strong evidence (P=0.00295) to conclude that women are more receptive to men during their fertile phase of their menstrual cycle.

10.2.15 **(a)** Six-banded armadillos make up the same proportion of roadkill in both locations.

(b) Pr{Armadillo killed | Atlantic Forest roadkill} = Pr{Armadillo killed | Cerrado roadkill}

(c)

	Atlantic Forest	Cerrado
Total Animals Killed (*n*)	178	318
Armadillos	51	66
Percent	28.7%	20.8%

(d)

Count (Expected)	Armadillo	Other	Total
Atlantic Forest	51 (41.9879)	127 (136.012)	**178**
Cerrado	66 (75.0121)	252 (242.988)	**318**
Total	**117**	**379**	**496**

(e) There is statistically significant evidence (*P*-value = $0.047 < \alpha = 0.05$) that the proportion of roadkill that is Armadillo is not the same for both locations.

(f) While it is reasonable to believe that the armadillos make up different proportions of roadkill at the two locations, that doesn't necessarily mean that there is more danger in one location than the other.

Table	Probability	Table	Probability
9 1 15 22	0.00581	1 9 23 14	0.00379
10 0 14 23	0.00038	2 8 22 15	0.02613
0 10 24 13	0.00022	3 7 21 16	0.09583

The probability of the sixth table is greater than .03594, so
P-value = 0.03594 + 0.00581 + 0.00038 + 0.00022 + 0.00379 + 0.02613 = 0.0723.
We do not reject H_0, because $P > \alpha$. There is insufficient evidence to conclude that the two types of therapy are not equally effective.

10.5.1 (a) The hypotheses are

H_0: The population success rates are the same in all five treatments
H_A: The population success rates differ across the five treatments

	Success	Failure	Total
Glucosamine	192 (199.05)	125 (117.95)	317
Chondroitin	202 (199.68)	116 (118.32)	318
Both	208 (199.05)	109 (117.95)	317
Placebo	178 (196.54)	135 (116.46)	313
Celebrex	214 (199.68)	104 (118.32)	318
Total	994	589	1583

$\chi^2_s = 9.29$. With df = (5 - 1)(2 - 1) = 4, Table 9 gives $\chi^2_{0.05} = 9.49$ and $\chi^2_{0.10} = 7.78$, so $0.05 < P < 0.10$.
We do not reject H_0; there is insufficient evidence ($0.05 < P < 0.10$) to conclude that the population success rates differ across the five treatments.

(b) $\chi^2_s = \frac{(192-199.05)^2}{199.05} + \frac{(125-117.95)^2}{117.95} + \ldots + \frac{(104-118.32)^2}{118.32} = 9.29.$

10.5.2 (a) H_0: The population sex ratio is the same in all three sites; H_A: The population sex ratio differs across the three sites.

(b) $e_{11} = 61.56$; $e_{12} = 27.70$; $e_{13} = 107.73$; $e_{21} = 58.44$; $e_{22} = 26.30$; $e_{23} = 102.27$.

(c) We reject H_0 because the P-value is smaller than 0.05. We have very strong evidence (P=5.8e-11) to conclude that the population sex ratio differs across the three sites.

(d)

Location	Total number of flies	Males Number	Males Percent
Woodland Site I	120	89	74
Woodland Site II	54	34	63
Open Ground	210	74	35

• **10.5.3** (a) Letting UP denote ulcer patient and C denote control, the hypotheses are

H_0: The blood type distributions are the same for ulcer patients and controls (Pr{O|UP} = Pr{O|C}, Pr{A|UP} = Pr{A|C}, Pr{B|UP} = Pr{B|C}, Pr{AB|UP} = Pr{AB|C})
H_A: The blood type distributions are not the same

The test statistic is $\chi^2_s = 49.0$. The degrees of freedom are df = (4 - 1)(2 - 1) = 3. From Table 9 we find that $\chi^2_{3,0.0001} = 21.11$. Because $\chi^2_s < \chi^2_{0.0001}$, the P-value is bracketed as

$P < 0.0001$.

At significance level $\alpha = 0.01$, we will reject H_0 if $P < 0.01$. Since $P < 0.0001$, we reject H_0 and conclude that the blood type distribution of ulcer patients is different from that of controls.

(b) Among ulcer patients, the percent frequency distribution of blood types is as follows:

Type O: Percent frequency = $\frac{911}{1655}$ x 100% = 55.0%.

Type A: Percent frequency = $\frac{579}{1655}$ x 100% = 35.0%.

Type B: Percent frequency = $\frac{124}{1655}$ x 100% = 7.5%.

Type AB: Percent frequency = $\frac{41}{1655}$ x 100% = 2.5%.

Among controls, the percent frequency distribution of blood types is as follows:

Type O: Percent frequency = $\frac{4578}{10000}$ x 100% = 45.8%.

Type A: Percent frequency = $\frac{4219}{10000}$ x 100% = 42.2%.

Type B: Percent frequency = $\frac{890}{10000}$ x 100% = 8.9%.

Type AB: Percent frequency = $\frac{313}{10000}$ x 100% = 3.1%.

These percentages can be arranged in a table as follows:

		Percent frequency	
		Ulcer patients	Controls
Blood type	O	55.0	45.8
	A	35.0	42.2
	B	7.5	8.9
	AB	2.5	3.1
	Total	100.0	100.0

(c)

		Ulcer patients	Controls	Total
Blood type	O	911 (779.43)	4578 (4709.57)	5489
	A	579 (681.31)	4219 (4116.69)	4798
	B	124 (143.99)	890 (870.01)	1014
	AB	41 (50.27)	313 (303.73)	354
	Total	1655	10000	11655

$$\chi^2_s = \frac{(911-779.43)^2}{779.43} + \frac{(4578-4709.57)^2}{4709.57} + \ldots = 49.0.$$

10.5.4 (a) Let C_1, C_2, and C_3 denote the three claw configurations and let T_1, T_2, and T_3 denote the three treatments.

H_0: Claw configuration is not affected by treatment ($\Pr\{C_1|T_1\} = \Pr\{C_1|T_2\} = \Pr\{C_1|T_3\}$,
$\Pr\{C_2|T_1\} = \Pr\{C_2|T_2\} = \Pr\{C_2|T_3\}$, and
$\Pr\{C_3|T_1\} = \Pr\{C_3|T_2\} = \Pr\{C_3|T_3\}$)
H_A: Claw configuration is affected by treatment (at least one of the equalities in H_0 is false)

With df = 4, Table 9 gives $\chi^2_{0.0001}$ = 23.51, so P < 0.0001. We reject H_0; there is sufficient evidence (P < 0.0001) to conclude that claw configuration is affected by treatment.

(b)

	C_1	C_2	C_3	Total
T_1	8 (4.57)	9 (5.91)	1 (7.52)	18
T_2	2 (6.60)	4 (8.54)	20 (10.87)	26
T_3	7 (5.84)	9 (7.55)	7 (9.61)	23
Total	17	22	28	62

$$\chi^2_s = \frac{(8-4.57)^2}{4.57} + \frac{(9-5.91)^2}{5.91} + \ldots = 24.35.$$

(c)

	R crusher, L cutter	R cutter, L crusher	R cutter, L cutter	Total
Oyster chips	44.4%	50.0%	5.6%	100%
Smooth plastic	7.7%	15.4%	76.9%	100%
One oyster chip	30.4%	39.1%	30.4%	100%

(d) The table in part (c) shows that the smooth plastic treatment increases the probability of a lobster having two cutter claws, whereas the oyster chips treatment increases the probability of a lobster having at least one crusher claw.

• **10.5.5** (a) The hypotheses are

H_0: Change in ADAS-Cog score is independent of treatment
H_A: Change in ADAS-Cog score is related to treatment

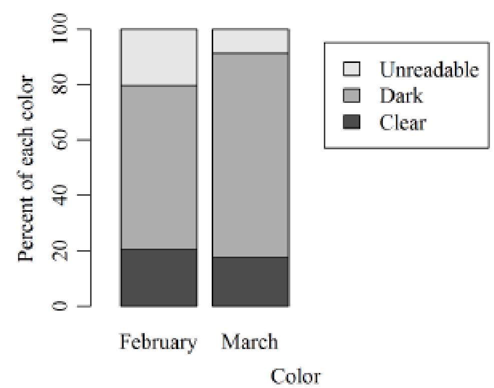
The test statistic is χ^2_s = 10.26. The degrees of freedom are df = (2 - 1)(5 - 1) = 4. From Table 9 we find that $\chi^2_{4,0.05}$ = 9.49 and $\chi^2_{4,0.02}$ = 11.67. Thus, 0.02 < P < 0.05, so we reject H_0. There is sufficient evidence (0.02 < P <0 .05) to conclude that EGb and placebo are not equally effective.

(b)

	-4 or better	-2 to -3	-1 to +1	+2 to +3	+4 or worse	Total
EGb	22(16)	18 (14.5)	12 (15.5)	7 (9)	16 (20)	75
Placebo	10 (16)	11 (14.5)	19 (15.5)	11 (9)	24 (20)	75
Total	32	29	30	18	40	150

$$\chi^2_s = \frac{(22-16)^2}{16} + \frac{(18-14.5)^2}{14.5} + \ldots = 10.26.$$

10.5.6 (a)



(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.