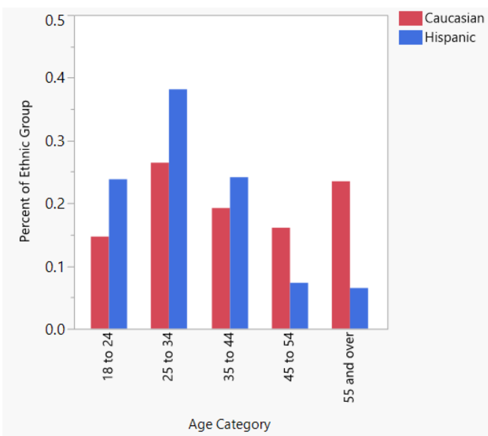
Chapter 25 – Two Categorical Variables: The Chi-Square Test

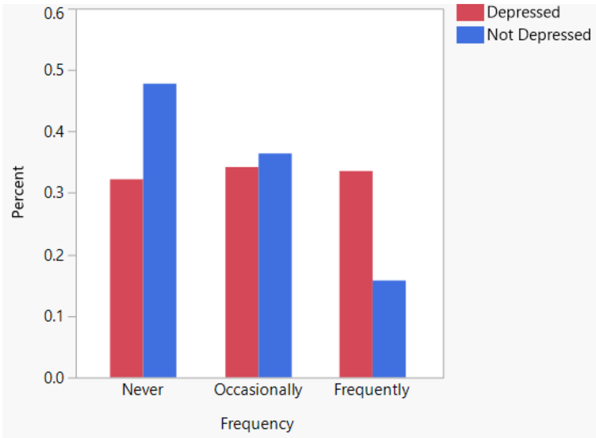
25.1 (a) The table provided gives percents in each category. As an example, there were a total of 5339 surveyed Caucasians. Of these, 785 were between 18 and 24; the proportion of Caucasians surveyed who were between 18 and 24 is 785/5339 = 0.147, which is represented as 14.7% in the table.

	Caucasian	Hispanic
18 to 24	14.7%	23.8%
25 to 34	26.5%	38.2%
35 to 44	19.2%	24.2%
45 to 54	16.1%	7.3%
55 and over	23.5%	6.5%

(b) The bar graph reveals that Hispanic visitors tend to be younger.



25.2 (a) Out of 18-year-olds suffering from depression, 101/(97 + 103 + 101) = 0.3355, or 33.55%, were bullied frequently. Out of those not depressed, 582/(1762 + 1343 + 582) = 0.1579, or 15.79%, were bullied frequently. A bar graph comparing the frequency of being bullied for those who are depressed and those who are not is given.



(b) This was an observational study and not a controlled experiment, so we cannot conclude causation.

25.3 (a) Let p_1 be the proportion of those who suffer from depression who were bullied occasionally and p_2 be the proportion of those who do not suffer from depression who were bullied occasionally. $\hat{p}_1 = 0.342$, $\hat{p}_2 = 0.364$, and the pooled sample proportion is $\hat{p} = \frac{103 + 1343}{97 + 103 + 101 + 1762 + 1343 + 582} = 0.3626$. We will test the hypotheses $H_0: p_1 = p_2$ against $H_a: p_1 \neq p_2$. We can do the test because the number of successes and failures for each group is large. The test statistic is $z = \frac{0.342 - 0.364}{\sqrt{0.3626(1-0.3626)(\frac{1}{301}+\frac{1}{3687})}} = -0.76$. and the P-value is $P = 0.4473$. There is not evidence of a significant difference between the proportions bullied occasionally for those with and without depression. **(b)** Let p_1 be the proportion of those who suffer from depression who were bullied frequently and p_2 be the proportion of those who do not suffer from depression who were bullied frequently. $\hat{p}_1 = 0.3355$, $\hat{p}_2 = 0.1579$, and the pooled sample proportion is $\hat{p} = \frac{101 + 582}{97 + 103 + 101 + 1762 + 1343 + 582} = 0.1713$. We will test the hypotheses $H_0: p_1 = p_2$ against $H_a: p_1 \neq p_2$. We can do the test because the number of successes and failures for each group is large. The test statistic is $z = \frac{0.3355 - 0.1579}{\sqrt{0.1713(1-0.1713)(\frac{1}{301}+\frac{1}{3687})}} = 7.86$, and the P-value is approximately zero. There is strong evidence that the proportions who were bullied frequently are different for those with and without depression. **(c)** The P-values only indicate the strength of evidence for a difference in a particular category of bullied frequency. This cannot tell us whether the two distributions, each with three outcomes, are significantly different. If we did three individual tests, we would not know how confident we could be in all three results when taken together.

25.4 (a) For the junior college sample, $\hat{p} = \frac{47}{47 + 36} = 0.5663$, so the standard error is $SE = \sqrt{\frac{0.5663(1-0.5663)}{47+36}} = 0.0544$. A 95% confidence interval for the proportion of all junior college graduates who think astrology is not at all scientific is $0.5663 \pm$

1.96(0.0544) = 0.4597 to 0.6729, or 45.97% to 67.29%. Following the same procedure for the other two intervals yields the given table.

Degree Held	$\hat{p}$	SE	95% Confidence Interval
Junior College	0.5663	0.0544	45.97% to 67.29%
Bachelor	0.8080	0.0263	75.65% to 85.95%
Graduate	0.8968	0.0271	84.4% to 95.0%

(b) Before we take a random sample, there's a 95% chance that the sample we collect will lead to a confidence interval that captures the true, unknown proportion of people who believe astrology is not at all scientific. As we construct more confidence intervals, each based on a different random sample, the chance that at least one of them fails to capture the parameter of interest increases. The probability, before sampling, that all three intervals contain the true proportions would be 0.95^3 = 0.8574, assuming independent samples.

25.5 (a) The expected counts for the four age categories (going from youngest to oldest) are 432.2, 741.2, 511.0, 327.5, and 446.1. These counts add up to the total observed counts of Hispanic visitors. **(b)** The observed counts for those under age 44 are larger than the expected counts, but the observed counts for those 45 and older are smaller than the counts that are expected if the distributions are the same.

25.6 (a) The expected counts are given in the table. We can see that the row and column totals agree with the observed counts.

	Never	Occasionally	Frequently	Row Totals
Depressed	140.3	109.1	51.6	301
Not Depressed	1718.7	1336.9	631.4	3687
Column Totals	1859.0	1446.0	683.0	n = 3988

(b) There are fairly large deviations between the observed and expected counts. In particular, the count for never depressed is much lower than expected under the null hypothesis, and the count for frequently depressed is much higher than expected under the null hypothesis.

25.7 (a) The null hypothesis is H_0: there is no relationship between ethnicity and age group for visitors to Monterey Bay Aquarium, and the alternative hypothesis is H_a: there is some relationship between ethnicity and age group for visits to Monterey Bay Aquarium. From Figure 25.3, we see the test statistic is $\chi^2 = 540.943$, and the P-value is $P < 0.0001$. **(b)** The cells that contribute the most to the test statistic are from the 55 and over age group. For Caucasians, the actual count for this group is larger than expected under the null hypothesis. For Hispanics, the actual count is smaller than would be expected.

25.8 (a) The chi-squared test statistic is $\chi^2 = 66.141$, and the P-value is $P < 0.0001$. Since the P-value is so small, we reject the null hypothesis. There is strong evidence that the frequency of bullying at age 13 is different for 18-year-olds who have depression versus those who do not. **(b)** Looking at the row percents, it appears the proportion of 18-year-olds with depression who were bullied frequently is larger than the proportion of those without depression who were bullied frequently.

25.9 PLAN: We want to test the hypothesis H_0: there is no relationship between the degree held and the view of astrology against H_a: there is some relationship between degree held and view of astrology. We will use a chi-square test. SOLVE: From Figure 25.5, we see the observed percent who view astrology as a science is much lower for bachelor and graduate than for junior college. Formally, we get a test statistic of $\chi^2 = 33.843$ and the P-value is $P = 0.000$. CONCLUDE: There is strong evidence that the distribution of view on astrology is related to the degree held.

25.10 (a) We know df $= (r-1)(c-1)$, where r is the number of rows and c the number of columns. Since there are five age groups and two ethnicities, we get df $= (5-1)(2-1) = 4$. **(b)** The value of the test statistic is much larger than the largest value shown in Table D for df $= 4$. From the table, we could conclude $P < 0.0005$. JMP's bound is more precise than the table's. **(c)** The degrees of freedom for this table would be df $= (r-1)(c-1) = (5-1)(4-1) = 12$.

25.11 (a) df $= (r-1)(c-1) = (2-1)(3-1) = 2$, since there are two categories for Depressed and three categories for Bullied. **(b)** The value of the test statistic is much larger than the largest value in Table D for df $= 2$, so we can conclude the P-value is $P < 0.0005$. The bound from JMP is more precise than the bound from the table. **(c)** If the null hypothesis is true, the mean of the test statistic is df $= 2$. The observed value is much larger than the mean, which is why the P-value is so small.

25.12 The smallest expected count for any cell is 51.55. Since all expected counts are at least 5, it is safe to use the chi-square test.

25.13 (a) This would be a chi-square test of homogeneity since we have two populations (high display and low display), and each individual is classified according to which sock they chose. **(b)** STATE: Use a chi-square test of homogeneity to determine if there is a relationship between which sock is chosen and whether the display was high or low. PLAN: Conduct a chi-square test of the hypotheses H_0: there is no relationship between sock location and display height versus H_a: there is some relationship between location and display height. SOLVE: The conditional distributions for the high and low display are given in the following table. In parentheses are the expected counts. From the conditional distributions, we notice the proportion who chose the middle location is a bit different for the high and low display. We can test whether the distributions are different using the chi-square test. Note that 8 of 10, or 80%, of cells have expected counts that are at least 5, and all cells have expected counts of at least one. It is safe to conduct a chi-square test. The test statistic is $\chi^2 = \frac{(14-13)^2}{13} + \frac{(12-12.5)^2}{12.5} + \cdots + \frac{(1-2.5)^2}{2.5} = 3.453$. The degrees of freedom are $df = (2-1)(5-1) = 4$. Using Table D, the P-value is $P > 0.25$. CONCLUDE: We fail to reject the null hypothesis. There is not evidence of a relationship between the location and display height. **(c)** No, this test does not answer that question. This test answers the question of whether there is a relationship between location and display height. Whether or not there are any differences among the preferences for the five locations is not questioning the relationship to display height, so the chi-square test of homogeneity is not appropriate.

	Loc 1	Loc 2	Loc 3	Loc 4	Loc 5
High Display	0.28 (13)	0.24 (12.5)	0.28 (17)	0.12 (5)	0.08 (2.5)
Low Display	0.24 (13)	0.26 (12.5)	0.40 (17)	0.08 (5)	0.02 (2.5)

25.14 (a) Assuming independence across years, we should use a chi-square test of homogeneity because each year is its own population. **(b)** STATE: Use a chi-square test of homogeneity to determine if there is a relationship between year and confidence. PLAN: Conduct a chi-square test of the hypotheses H_0: there is no relationship between year and confidence versus H_a: there is some relationship between year and confidence. A contingency table containing the relevant information was computed in JMP and is shown below. From the conditional distributions, we see the degree of confidence percents are quite different from one year to another. All expected counts are at least 5, so we can safely use a chi-square test. The test statistic is $\chi^2 = 13.5739 + 58.9510 + \cdots + 30.9106 = 252.116$. The degrees of freedom are $\text{df} = (4-1)(3-1) = 6$. The P-value is $P < 0.0001$ (Using technology). CONCLUDE: We reject the null hypothesis. There is strong evidence of a relationship between year and degree of confidence.

		Confidence			
	Count Row % Expected Cell Chi^2	Great_deal	Hardly_any	Some	Total
Year	2002	117 13.12 83.3616 13.5739	222 24.89 369.611 58.9510	553 62.00 439.028 29.5874	892
	2006	209 10.78 181.115 4.2932	703 36.27 803.033 12.4610	1026 52.94 953.852 5.4572	1938
	2010	125 9.28 125.884 0.0062	587 43.58 558.145 1.4917	635 47.14 662.971 1.1801	1347
	2014	93 5.66 153.64 23.9337	900 54.74 681.211 70.2701	651 39.60 809.15 30.9106	1644
	Total	544	2412	2865	5821

25.15 We test $H_0: p_1 = p_2 = p_3 = \frac{1}{3}$ versus H_a: not all three are equally likely. There were 53 bird strikes in all, so the expected counts are each $53 \times \frac{1}{3} = 17.67$. The chi-square statistic is then $\chi^2 = \sum \frac{(\text{observed count} - 17.67)^2}{17.67} = \frac{(31-17.67)^2}{17.67} + \frac{(14-17.67)^2}{17.67} + \frac{(8-17.67)^2}{17.67} = 10.06 + 0.76 + 5.29 = 16.11$. The degrees of freedom are df = 2. From Table D, $\chi^2 = 16.11$ falls beyond the 0.0005 critical value, so $P < 0.0005$. There is very strong evidence that the three tilts differ. The data and terms of the statistic show that more birds than expected strike the vertical window, and fewer than expected strike the 40-degree window.

25.16 (a) The population proportion with less than a high school degree is $p_{10} = \frac{69195}{461299} = 0.15$. Similarly, the proportion with a high school degree is $p_{20} = 0.547$, and the population proportion with a college degree is $p_{30} = 0.303$ We will do a chi-square test for goodness of fit to see if education levels in the sample are significantly different than the levels in the population. That is, we test the hypotheses $H_0: p_1 = 0.15, p_2 = 0.547. p_3 = 0.303$ versus H_a: at least one of the proportions is not equal to the value given in the null. Under the null hypothesis, the expected count for the less than high school degree group is $656(0.150) = 98.4$, for the high school degree group we have $656(0.547) = 358.832$, and for the college degree group we have $656(0.303) = 198.768$. These counts are all large, so we can use the chi-square test. The test statistic is $\chi^2 = \frac{(50 - 98.4)^2}{98.4} + \frac{(286 - 358.832)^2}{358.832} + \frac{(320 - 198.768)^2}{198.768} = 112.53$. There are $df = 3 - 1 = 2$ degrees of freedom. The P-value is much lower than 0.0005. Thus, there is strong evidence that the education levels in the sample differ from the population. Comparing the observed and expected counts implies those with less than a college degree are underrepresented, and those with a college degree are overrepresented. **(b)** Yes, the estimate is likely biased because the sample is not representative of the entire population. Since those with a college degree use the trail more often and are overrepresented in the sample, the estimate of 27% likely overestimates the true proportion of the population that uses the trail.

25.17 (a) The percent of subjects who chose each location (ordered from location 1 to location 5) are 26%, 25%, 34%, 10%, and 5%. **(b)** If all locations are equally likely, we would expect 100/5, or 20%, to choose each location. Thus, the expected count is 20 for each location. **(c)** PLAN: Use a chi-square goodness of fit test of the hypotheses H_0: $p_1 = p_2 = \cdots = p_5 = 0.2$ versus H_a: the probabilities are not all 0.2. SOLVE: We can perform a chi-square test since the expected count is greater than 5 for each group. The test statistic is $\chi^2 = \frac{(26 - 20)^2}{20} + \cdots + \frac{(5 - 20)^2}{20} = 29.1$. The degrees of freedom are $\text{df} = 5 - 1 = 4$. Using Table D, the P-value is less than 0.0005. CONCLUDE: We reject the null hypothesis. There is strong evidence that the locations are not all equally likely to be chosen. **(d)** PLAN: Let p denote the proportion that choose the center location. We want to know if the center location is chosen more often than it would be if all locations were chosen randomly. Test the hypotheses $H_0: p = 0.2$ against $H_a: p > 0.2$ using a significance test for a population proportion. SOLVE: The sample proportion is $\hat{p} = \frac{34}{100} = 0.34$. The test statistic is $z = \frac{0.34 - 0.2}{\sqrt{\frac{0.2(0.8)}{100}}} = 3.5$ and, using Table C, the P-value is less than 0.0002. CONCLUDE: Reject the null hypothesis. There is strong evidence that the item in the center is chosen more often than those at other locations.

25.18 Letting p_1 denote the proportion of people age 16 to 29 cited for not wearing seatbelts, and similarly defining p_2 and p_3 for the other age groups, we test H_0: $p_1 = 0.328, p_2 = 0.594, p_3 = 0.078$ versus H_a: not all proportions are equal to the population proportion of not wearing seat belts. The details of the computation are shown in the table. The expected counts are found by multiplying the expected frequencies by 803 (the total number of observations). The difference is significant: $\chi^2 = 119.84$, df = 2, and $P < 0.0005$ (using software, $P = 0.000$ to three decimal places). The largest contribution to the statistic comes from the youngest age group, which is cited more frequently than we would have expected under the hypothesis of no association.

	Expected Frequency	Observed Count	Expected Count	$O - E$	$\frac{(O-E)^2}{E}$
16 to 29	0.328	401	263.384	137.616	71.9032
30 to 59	0.594	382	476.982	−94.982	18.9139
60 or Older	0.078	20	62.634	−42.634	29.0203
		803			119.8374

25.19 STATE: Are all 12 astrological signs equally likely? PLAN: We test $H_0: p_1 = p_2 = \cdots = p_{12} = 1/12$ versus H_a: the 12 astrological-sign birth probabilities are not equally likely. SOLVE: There are 2402 subjects in this sample. Under H_0, we expect 2402/12 = 200.17 per sign, so all cells have expected counts greater than 5 and a chi-square test is appropriate. The test statistic is: $\chi^2 = \frac{(205 - 200.17)^2}{200.17} + \frac{(174 - 200.17)^2}{200.17} + \cdots + \frac{(198 - 200.17)^2}{200.17} = 18.34$. With $\text{df} = 12 - 1 = 11$, the P-value is $0.05 < P < 0.10$ (using Table D). CONCLUDE: There is little evidence that some astrological signs are more likely in birth than others. That is, there is little to no support for a conclusion that astrological signs are not equally likely.

25.20 (b) 9.37%.

25.21 (c) 22.3%. The number of current smokers in the study was 404, so we get 90/404, or 22.3%.

25.22 (b) 14.2%. 90/(90 + 545), or 14.2%.

25.23 (b) independence. The results are from classifying individuals in a single SRS.

25.24 (b) 47.7. The total number of individuals in the study was 4310, so the expected count is (25 + 484)(404)/4310 = 47.7.

25.25 (a) 10.8. $\frac{(25-47.7)^2}{47.7} = 10.8$.

25.26 (a) 4. $(r-1)(c-1) = (5-1)(2-1) = 4$.

25.27 (a) the distributions of health rating are the same for smokers and nonsmokers.

25.28 (c) the distributions of health rating are different for smokers and nonsmokers.

25.29 (c) less than 0.0005. Using Table D, we find P < 0.0005.

25.30 (c) there is a relationship between smoking status and how people rate their health.

25.31 (a) and **(b)** STATE: We compare rates of success at smoking cessation in three groups. PLAN: B denotes the bupropion group, P the placebo group, and C the Chantix® group. First, we'll construct a confidence interval for the difference in success proportions between the bupropion group and the placebo group. Then, we'll conduct a test for overall equality of success rates. SOLVE: The sample proportions are $\hat{p}_{\mathrm{B}} = 0.2948$ and $\hat{p}_{\mathrm{P}} = 0.1773$. The standard error is SE = 0.0325, so the large-sample 95% confidence interval for $p_{\mathrm{B}} - p_{\mathrm{P}}$ is $\hat{p}_{\mathrm{B}} - \hat{p}_{\mathrm{P}} \pm 1.96SE$ = 0.1175 ± 0.0637 = 0.0538 to 0.1812. The sample proportion for Chantix® is $\hat{p}_{\mathrm{C}} = 0.4403$. Now, to test $H_0 : p_{\mathrm{C}} = p_{\mathrm{B}} = p_{\mathrm{P}}$ versus H_a: not all proportions are equal, we perform a chi-square test. Note that all of the expected cell counts are more than five (the smallest is 100.47), and trials are clearly independent, so such a test is appropriate. We have $\chi^2 = 56.992$, df = 2, and P < 0.0005 (from software, P = 0.0000 to four places), so the evidence for a relationship among treatments (placebo, bupropion, or Chantix®) and cessation of smoking is overwhelming. Examining the output, we see that more Chantix® users than expected and fewer placebo users than expected were successful. CONCLUDE: The treatments are not equally successful at helping people to quit smoking. The Chantix® group, in particular, has a higher success rate. We also conclude that the placebo and bupropion success rates are not the same.

```
         Chantix  Bupropion  Placebo  Total
 No      155      97         61       313
         107.49   100.47     105.05
 Yes     197      232        283      712
         244.51   228.53     238.95

Chi-Sq = 56.992, DF = 2, P-Value = 0.000
```

(c) This is a test of homogeneity, because the subjects were randomly assigned to three different treatments (they can be considered three separate samples).

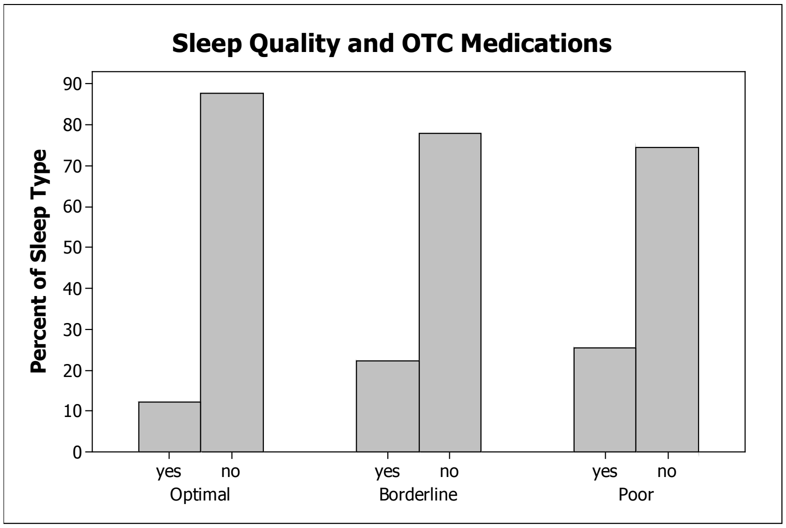
25.32 (a) These were separate random samples, so this is a test of homogeneity. **(b)** STATE: We want to determine if the distribution of age for those with a landline differs from the distribution of those with only a cell phone. PLAN: We test H_0: the distribution of age groups is the same for landline and cell-only individuals versus H_a: the distributions are different. SOLVE: All expected cell counts are more than five, so the guidelines for the chi-square test are satisfied. We have $\chi^2 = 1032.892$, df = 3, and $P < 0.0005$. CONCLUDE: There is strong evidence of an association between age group and the type of telephone. In fact, the younger age groups were much more likely than expected to have only cell phones.

	Cell Only	Landline	All
Age18-29	374	335	709
	52.75	47.25	100.00
	115	594	709
	587.04	113.20	*
Age30-49	347	1242	1589
	21.84	78.16	100.00
	257	1332	1589
	31.63	6.10	*
Age50-64	146	1625	1771
	8.24	91.76	100.00
	286	1485	1771
	68.75	13.26	*
Age65up	36	1481	1517
	2.37	97.63	100.00
	245	1272	1517
	178.51	34.42	*
All	903	4683	5586
	16.17	83.83	100.00
	903	4683	5586
	*	*	*

```
Cell Contents:   Count
                 % of Row
                 Expected count
                 Contribution to Chi-square


Pearson Chi-Square = 1032.892, DF = 3, P-Value = 0.000
```

25.33 (a) A total of 174 of the 871 students used OTC stimulants; $\hat{p} = 0.1998$. The sample counts are large, so we can use $\hat{p} \pm 1.96\ SE$ for the confidence interval, which becomes $0.1998 \pm 1.96\sqrt{\frac{0.1998(1-0.1998)}{871}}$, or 0.1732 to 0.2264. **(b)** A graph comparing the distributions is given. The conditional distributions are given in the second row for each cell in the Minitab output shown. It appears that those who use OTC medications are less likely to have optimal sleep and more likely to have poor sleep than those who do not use these medications.



	Borderline	Optimal	Poor
No	186	266	245
	77.82	87.79	74.47
	191.3	242.5	263.3
	0.144	2.284	1.269
Yes	53	37	84
	22.18	12.21	25.53
	47.7	60.5	65.7
	0.578	9.147	5.082

```
Cell Contents:   Count
                 % of Row
                 Expected count
                 Contribution to Chi-square

Pearson Chi-Square = 18.504, DF = 2, P-Value = 0.000
```

(c) We test H_0: there is no association between taking OTC medications to stay awake and sleep quality versus H_a: there is an association between OTC medications to stay awake and sleep quality. From the Minitab output in part (b), we have $\chi^2 = 18.504$, and $P < 0.0005$. We conclude that there is an association between taking over-the-counter medications to stay awake and sleep quality.

Treatment	Count / Expected / Cell Chi^2	Score 2	Score 3	Score 4	Total
	No-surgery	5	21	24	50
		6	21	23	
		0.1667	0.0000	0.0435	
	Surgery	7	21	22	50
		6	21	23	
		0.1667	0.0000	0.0435	
	Total	12	42	46	100

25.34 Test the hypotheses H_0: the distribution of Kellner-Lawrence score is the same for the nonsurgical and surgical treatment groups versus H_a: the distribution of scores is not the same for the two groups. JMP output for the chi-square test of homogeneity is provided. The expected counts are all greater than 5, so it is safe to perform a chi-square test. The test statistic is $\chi^2 = 0.1667 + 0 + \cdots + 0.0435 = 0.420$. With df $= 2$, the *P*-value is large (using technology, we get *P* = 0.81). There is essentially no evidence to suggest the distribution of scores is not the same for the two groups.

25.35 (a) We test $H_0 : p_{\text{G}} = p_{\text{NG}}$ versus $H_a : p_{\text{G}} \neq p_{\text{NG}}$. With $\hat{p}_{\text{G}} = \frac{36}{91} = 0.395604$ and $\hat{p}_{\text{NG}} = \frac{578}{2014} = 0.286991$, the pooled proportion is $\hat{p} = 0.291686$. The standard error is *SE* = 0.048713. Hence, $z = \frac{\hat{p}_{\text{G}} - \hat{p}_{\text{NG}}}{SE} = 2.23$, and *P* = 0.0258. **(b)** By inspection of the output given, we see that all expected cell counts exceed 5, and use of a chi-square test is appropriate. We find $\chi^2 = 4.971$ with df = 1. From software, *P* = 0.026.

```
              Fight  No Fight       All
No              578      1436      2014
              587.5    1426.5    2014.0
             0.1522    0.0627         *
Yes              36        55        91
               26.5      64.5      91.0
             3.3690    1.3874         *

Cell Contents:   Count
                 Expected count
                 Contribution to Chi-square

Pearson Chi-Square = 4.971, DF = 1, P-Value = 0.026
```

(c) $z^2 = (2.22965)^2 = 4.971$, which is equal to χ^2. Obviously, *P*-values also agree. **(d)** We would use a one-sided *z* test. The chi-square test is inherently two-sided, because it tests for general association instead of for a particular direction of association.

25.36 (a) The two-way table is provided. We test $H_0 : p_1 = p_2$ versus $H_a : p_1 < p_2$.

	Tumor	No Tumor
Group 1	11	19
Group 2	22	8

(b) The z test must be used, because the chi-square procedure measures evidence in support of any association and is implicitly two-sided. We have $\hat{p}_1 = 0.3667$ and $\hat{p}_2 = 0.7333$. The pooled proportion is $\hat{p} = (11 + 22)/(30 + 30) = 0.55$ and the standard error is SE = 0.12845, so z = −2.85 and P = 0.0022. We have strong evidence that rats that can stop the shock (and, therefore, presumably have better attitudes) develop tumors less often than rats that cannot (and, therefore, are presumably depressed).

25.37 (a) This was a single sample and individuals are tabled by two categorical variables, so this is a test of independence. **(b)** STATE: Is there a difference between how men and women assess their chances of being rich by age 30? PLAN: We test H_0: there is no relationship between sex and self-assessment of chances of being rich versus H_a: there is some relationship between these factors. SOLVE: Examining the Minitab output in Figure 25.8, we see that conditions for use of the chi-square test are satisfied because all expected cell counts exceed 5 (the smallest is 95.2). We have $\chi^2 = 43.946$ with df = 4, leading to P < 0.0005 (using Table D). CONCLUDE: Overall, men give themselves a better chance of being rich. This difference shows up most noticeably in the second and fifth rows of the table: Women were more likely to say "some chance, but probably not," whereas men more often responded "almost certain." There was virtually no difference between men and women in the "almost no chance" and "a 50–50 chance" responses, and little difference in the "a good chance" response.

25.38 STATE: Does sexual content of ads differ in magazines aimed at different audiences? PLAN: We test H_0: there is no relationship between sexual content of ads and magazine audience versus H_a: there is some relationship between sexual content of ads and magazine audience. SOLVE: Examining the Minitab output in Figure 25.9, we see that conditions for use of the chi-square test are satisfied because all expected cell counts exceed 5 (the smallest is 82.4). We have $\chi^2 =$ 80.874 with df = 2, leading to P < 0.0005. CONCLUDE: Magazines aimed at women are much more likely to have sexual depictions of models than the other two types of magazines. Specifically, about 39% of ads in women's magazines show sexual depictions of models, compared with 21% and 17% of ads in general-audience and men's magazines, respectively. The two women's chi-squared terms account for over half of the total chi-square value.

25.39 (a) We compare the percent of dogs in each condition type (I, II, III) that make a specified number of errors. The table summarizes the data. We see that under Type I condition (social-communicative), dogs tend to make more errors, whereas under Type III condition (nonsocial), dogs tend to make fewer errors.

	0	1	2	3
Type I	0.0%	25.0%	25.0%	50%
Type II	41.7%	25.0%	8.3%	25%
Type III	66.7%	16.7%	16.7%	0%

(b) We have many cells with expected cell counts lower than 5. We also have two zeroes in the table. Also, each dog had up to three trials, so observations are not independent. **(c)** Software should warn users against using the chi-square test. The software package Minitab does provide such a warning.

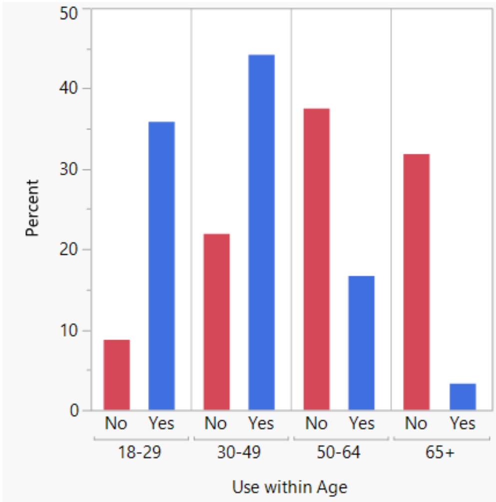
25.40 We need cell counts, not just percents. If we had been given the number of travelers in each group—leisure and business—we could have estimated this.

25.41 Presumably, many of the individuals were included in more than one category of alcohol and drug use. The chi-square test should only be used when individuals are classified to a specific category.

25.42 In order to do a chi-square test, each subject can only be counted once. In this experiment, each individual is represented for both treatments (carob and chocolate).

25.43 **(a)** Let p denote the proportion of Americans over the age of 18 who use a social networking site on their phone. The sample proportion is $\hat{p} = \frac{636}{1918} = 0.3316$. We can construct a large-sample confidence interval, since the number of successes and failures are large. The 99% confidence interval is $0.3316 \pm 2.576\sqrt{\frac{0.3316(1-0.3316)}{1918}} = 0.3039$ to 0.3593, or 30.39% to 35.93%. **(b)** The conditional distributions are shown in the second row of each cell in the provided JMP output. A bar graph is also provided. The table and graph show younger individuals (under age 50) are much more likely to use social networking on their phone than older individuals (age 50 and over).

Age	Count / Col % / Expected / Cell Chi^2	Use: No	Use: Yes	Total
18-29	Count	112	228	340
	Col %	8.74	35.85	
	Expected	227.258	112.742	
	Cell Chi^2	58.4548	117.829	
30-49	Count	281	281	562
	Col %	21.92	44.18	
	Expected	375.643	186.357	
	Cell Chi^2	23.8454	48.0657	
50-64	Count	481	106	587
	Col %	37.52	16.67	
	Expected	392.353	194.647	
	Cell Chi^2	20.0284	40.3717	
65+	Count	408	21	429
	Col %	31.83	3.30	
	Expected	286.746	142.254	
	Cell Chi^2	51.2742	103.355	
Total		1282	636	1918



(c) Test the hypotheses H_0: there is no relationship between age and phone use for social networking versus H_a: there is some relationship between age and phone use for social networking. We can do a chi-square test because all cells have at least five expected counts. The value of the test statistic is $\chi^2 = 58.45 + 117.83 + \cdots + 103.35 = 463.224$. The degrees of freedom are $\text{df} = (r-1)(c-1) = (4-1)(2-1) = 3$. If the null hypothesis were true, the mean of the test statistic would be 3. Our observed test statistic is much larger than 3. The *P*-value is approximately zero. There is strong evidence of a relationship between age and phone use for social networking. **(d)** The cells that contribute most to the chi-squared statistic are the "18-29 and Yes" cell and the "65+ and Yes" cell. For the 18-29 group, we observe a much higher number of Yes responses than expected if the null were true. For the 65+ group, we observe a much smaller number of Yes responses than expected if the null were true. It appears younger individuals are much more likely to use social networking on their phone than older individuals.

25.44 (a) We test H_0: there is no relationship between degree held and service attendance versus H_a: there is some relationship between degree held and service attendance. Examining the Minitab output shown, $\chi^2 = 14.19$ with df = 3, and *P*-value = 0.003. There is strong evidence of an association between degree held and service attendance.

```
              HS   JColl   Bachelor   Graduate     All
 No          880     101        232        105    1318
           842.7   107.3      248.9      119.2  1318.0
 Yes         400      62        146         76     684
           437.3    55.7      129.1       61.8   684.0

 Cell Contents:   Count
                  Expected count

Pearson Chi-Square = 14.190, DF = 3, P-Value = 0.003
```

(b) The new table is shown below. We find $\chi^2 = 0.73$ on df = 2, and $P = 0.694$. In this table, we find no evidence of association between religious service attendance and degree held.

	HS	Bachelor	Graduate	All
No	101	232	105	438
	98.9	229.3	109.8	438.0
Yes	62	146	76	284
	64.1	148.7	71.2	284.0

```
Cell Contents:  Count
                Expected count

Pearson Chi-Square = 0.729, DF = 2, P-Value = 0.694
```

(c) The new table is shown. We have $\chi^2 = 13.4$ with df = 1. $P = 0.000$ to three decimal places (it's actually 0.0002). There is overwhelming evidence of association between level of education (high school versus beyond high school) and religious service attendance.

	BeyondHS	HSchool	All
Attend	284	400	684
	246.7	437.3	684.0
No Attend	438	880	1318
	475.3	842.7	1318.0

```
Cell Contents:  Count
                Expected count

Pearson Chi-Square = 13.416, DF = 1, P-Value = 0.000
```

(d) In general, we find that people with degrees beyond high school attend services more often than expected; people with high school degrees attend services less often than expected. Of those with high school degrees, 31.3% attended services, and the percents are 38%, 38.6%, and 42%, respectively, for people with junior college, bachelor, and graduate degrees.

25.45 The conditional distributions and necessary output for the chi-square test are included in the following JMP output. The frequencies are very similar for males and females. There are slightly more females who have experienced either occasional or frequent bullying, whereas males have a slightly higher percent of never being bullied. Test the hypotheses H_0: there is no relationship between sex and frequency of bullying versus H_a: there is some relationship between sex and frequency of bullying. We can do a chi-square test because the expected counts are all much larger than 5. The degrees of freedom are $\text{df} = (2 - 1)(3 - 1) = 2$. The test statistic is $\chi^2 = 1.842 + 1.1941 + \cdots + 0.4059 = 7.007$. Using technology, the P-value is $P = 0.0301$. There is some evidence that the distribution of frequency of being bullied is not the same for males and females.

Sex	Count / Row % / Expected / Cell Chi^2	Bullied: Never	Occasional	Frequent	Total
female	Count	1526	1281	629	3436
	Row %	44.41	37.28	18.31	
	Expected	1579.95	1242.48	613.571	
	Cell Chi^2	1.8420	1.1941	0.3880	
male	Count	1564	1149	571	3284
	Row %	47.62	34.99	17.39	
	Expected	1510.05	1187.52	586.429	
	Cell Chi^2	1.9272	1.2493	0.4059	
Total		3090	2430	1200	6720

25.46 STATE: Is there a relationship between race and parental opinion of schools? PLAN: We use a chi-square test to test H_0: there is no relationship between race and opinion about schools versus H_a: there is some relationship between race and opinion about schools. SOLVE: All expected cell counts exceed 5 (the smallest is 21.26), so use of a chi-square test is appropriate. We find that $\chi^2 = 22.426$ with df = 8, and P = 0.004. CONCLUDE: We have strong evidence of a relationship between race and opinion of schools. Specifically, according to the sample (as illustrated in the table given), blacks are less likely and Hispanics are more likely to consider schools to be excellent, while Hispanics and whites differ in the percent considering schools good. Also, a higher percent of blacks rated schools as fair.

Schools	Count / Expected / Cell Chi^2	Race: Black	Hispanic	White	Total
Excellent	Count	12	34	22	68
	Expected	22.7041	22.7041	22.5917	
	Cell Chi^2	5.0466	5.6200	0.0155	
Good	Count	69	55	81	205
	Expected	68.4463	68.4463	68.1074	
	Cell Chi^2	0.0045	2.6415	2.4405	
Fair	Count	75	61	60	196
	Expected	65.4413	65.4413	65.1174	
	Cell Chi^2	1.3962	0.3014	0.4022	
Poor	Count	24	24	24	72
	Expected	24.0397	24.0397	23.9207	
	Cell Chi^2	0.0001	0.0001	0.0003	
Don't know	Count	22	28	14	64
	Expected	21.3686	21.3686	21.2628	
	Cell Chi^2	0.0187	2.0580	2.4808	
Total		202	202	201	605

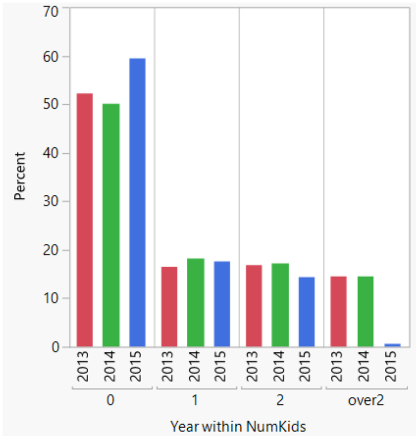
25.47 (a) A chi-square test of independence should be used, because we have an SRS and each individual was classified according to the eGFR and hearing loss. **(b)** The JMP output for the chi-square test is provided. The hypotheses are H_0: there is no relationship between hearing loss and eGFR versus H_a: there is some relationship between hearing loss and eGFR. We can do a chi-square test because all expected counts are at least 5. The test statistic is $\chi^2 = 555.16$ with df $= 4$. Using technology, the P-value is $P < 0.0001$ (using software, with Table D, we get $P <$

0.0005). There is strong evidence of a relationship between hearing loss and eGFR. Based on the observed conditional percents, it appears individuals with hearing loss are more likely to have eGFR at least 90 than those without hearing loss. Those without hearing loss are more likely to have eGFR less than 45 than those with hearing loss.

Hearing loss	Count / Row % / Expected / Cell Chi^2	eGFR <45	45-60	60-75	75-90	>90	Total
no	Count	717	27	458	207	295	1704
	Row %	42.08	1.58	26.88	12.15	17.31	
	Expected	527.017	115.638	439.292	184.755	437.298	
	Cell Chi^2	68.4863	67.9422	0.7967	2.6783	46.3042	
yes	Count	76	147	203	71	363	860
	Row %	8.84	17.09	23.60	8.26	42.21	
	Expected	265.983	58.3619	221.708	93.2449	220.702	
	Cell Chi^2	135.699	134.620	1.5786	5.3069	91.7468	
Total		793	174	661	278	658	2564

(c) Age is a quantitative variable. It could also be a lurking variable because it may be related to both eGFR and hearing loss. **(d)** If the relationship is the same for each age group, then age is not contributing to the relationship between eGFR and hearing loss.

25.48 PLAN: We compare how the number of children per group has changed from 2013 through 2015 at the Monterey Bay Aquarium. We will create a bar graph and do a chi-square test of homogeneity (each year is a separate sample). SOLVE: To examine any possible change in the number of children per group, we first look at a bar graph of the data (given). The graph indicates that the number of children has been pretty steady across the three years, with the exception that far fewer had 3 or more children in 2015 than in the earlier years. JMP output for the chi-square test is provided. Note that all expected counts are above 5, so it is safe to use the chi-square test. The test statistic is $\chi^2 = 106.319$, the degrees of freedom are df $= (4-1)(3-1) = 6$, and the P-value is $P < 0.0001$. CONCLUDE: There is strong evidence that the distribution of the number of children has changed over the three-year period. Looking at the bar graph and conditional percents, it appears that, in 2015, people were more likely to go to the aquarium with no kids and less likely to go with more than two kids than in previous years.



NumKids	Count / Col % / Expected / Cell Chi^2	Year 2013	2014	2015	Total
0	Count	1855	1751	1998	5604
	Col %	52.19	50.11	59.45	
	Expected	1913.4	1881.1	1809.5	
	Cell Chi^2	1.7827	8.9980	19.6374	
1	Count	585	636	591	1812
	Col %	16.46	18.20	17.58	
	Expected	618.681	608.236	585.083	
	Cell Chi^2	1.8336	1.2673	0.0598	
2	Count	599	601	483	1683
	Col %	16.85	17.20	14.37	
	Expected	574.636	564.934	543.43	
	Cell Chi^2	1.0330	2.3024	6.7199	
over2	Count	515	506	289	1310
	Col %	14.49	14.48	8.60	
	Expected	447.28	439.729	422.991	
	Cell Chi^2	10.2530	9.9876	42.4442	
Total		3554	3494	3361	10409

25.49 There are 2512 adults in the sample. Of these, 502 are Independent. Therefore, $\hat{p} = 502/2512 = 0.1998$, and $SE = \sqrt{\frac{0.1998(1-0.1998)}{2512}} = 0.008$. A 95% confidence interval for the proportion of Independent adults is $0.1998 \pm 1.96(0.008) = 0.1841$ to 0.2155, or 18.41% to 21.55%.

25.50 STATE: How do the conditional distributions of political leaning, given education, compare? PLAN: We compare the percents leaning toward each party within each education group. SOLVE: A table of output from JMP is provided. At each education level, we compute the percent leaning toward each party. CONCLUDE: At every education level, people leaning Democrat outweigh people leaning Republican. The difference is greatest at the "none" level of education, followed by the "graduate" level of education.

Politics	Count Col %	None	Highschool	JrCollege	Bachelor	Graduate	Total
		Education					
Democrat		145	565	80	217	155	1162
		70.39	58.01	55.56	56.22	65.13	
Republican		61	409	64	169	83	786
		29.61	41.99	44.44	43.78	34.87	
Total		206	974	144	386	238	1948

25.51 PLAN: We will find conditional distributions for political leaning at each level of education and perform a chi-square test on the full table, testing the null hypothesis of no relationship between education level and political preference. SOLVE: The conditional distributions are tabulated. Only 1 of 40 cell counts is less than five, so it is safe to use a chi-square test. Using software, the chi-square statistic is $\chi^2 = 106.239$ with df $= 28$. The P-value is $P < 0.0001$. CONCLUDE: Student observations about the full table will vary. Notice that among the "none" education group, there is a much larger proportion of Independents, and in the "graduate" education group, there is a very high percent of Democrats and people who lean Democrat. The general result of this exercise, where we find that the differences among party affiliations vary across levels of education, is consistent with our conclusion in Exercise 25.49.

Politics	Count Col % Expected Cell Chi^2	None	Highschool	JrCollege	Bachelor	Graduate	Total
		Education					
Independent		118	251	36	67	30	502
		35.87	19.98	19.67	14.32	10.87	
		65.7476	251	36.5709	93.5255	55.1561	
		41.5272	0.0000	0.0089	7.5231	11.4734	
NearDemocrat		40	163	26	66	42	337
		12.16	12.98	14.21	14.10	15.22	
		44.1373	168.5	24.5506	62.785	37.0271	
		0.3878	0.1795	0.0856	0.1646	0.6679	
NearRepublican		24	136	19	45	25	249
		7.29	10.83	10.38	9.62	9.06	
		32.6119	124.5	18.1397	46.3901	27.3583	
		2.2741	1.0622	0.0408	0.0417	0.2033	
Otherparty		5	31	3	15	8	62
		1.52	2.47	1.64	3.21	2.90	
		8.12022	31	4.51672	11.551	6.8121	
		1.1990	0.0000	0.5093	1.0299	0.2071	
StrongDemocrat		53	198	23	81	64	419
		16.11	15.76	12.57	17.31	23.19	
		54.877	209.5	30.5243	78.0621	46.0366	
		0.0642	0.6313	1.8547	0.1106	7.0093	
StrongRepublican		18	131	15	53	28	245
		5.47	10.43	8.20	11.32	10.14	
		32.088	122.5	17.8483	45.6449	26.9188	
		6.1852	0.5898	0.4546	1.1852	0.0434	
WeakDemocrat		52	204	31	70	49	406
		15.81	16.24	16.94	14.96	17.75	
		53.1744	203	29.5772	75.6401	44.6083	
		0.0259	0.0049	0.0684	0.4206	0.4324	
WeakRepublican		19	142	30	71	30	292
		5.78	11.31	16.39	15.17	10.87	
		38.2436	146	21.2723	54.4013	32.0828	
		9.6831	0.1096	3.5808	5.0645	0.1352	
Total		329	1256	183	468	276	2512

25.52 and **25.53** are Web-based exercises.