Section 11.1: Independence and Dependence (Association)

11.1 Gender gap in politics?

a) The response variable is political party identification, and the explanatory variable is gender.

b)

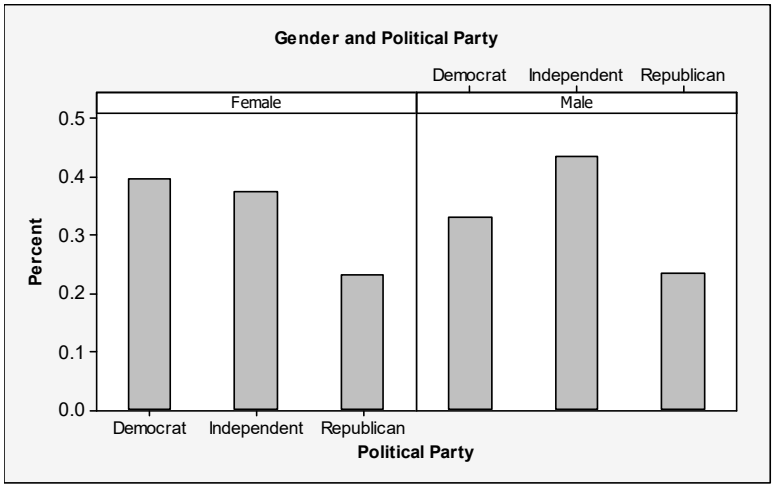
	Political Party Identification				
Gender	**Democrat**	**Independent**	**Republican**	**Total**	***n***
Female	39.6%	37.4%	23.0%	100%	1063
Male	33.0%	43.5%	23.5%	100%	843

Women are more likely than are men to be Democrats, whereas men are more likely than are women to be Independents. The likelihood of being Republican is about the same for males and females.

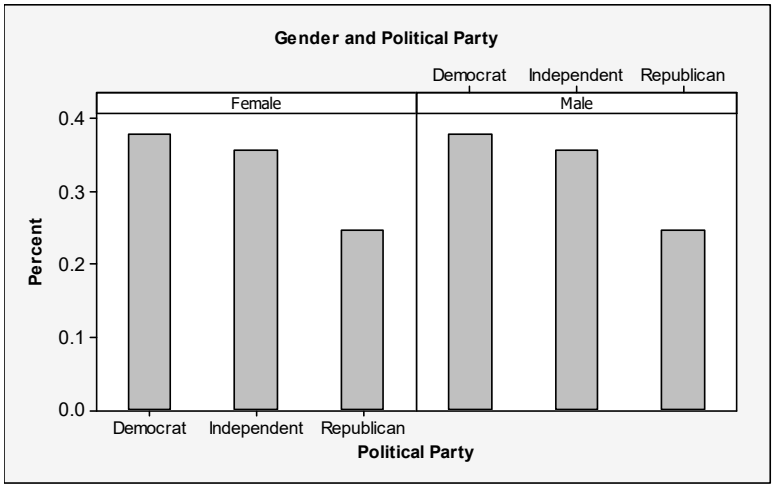
c) There are many possible hypothetical conditional distributions for which these variables would be independent. Distributions should show percentages in the party categories that are the same for men and women. Here's one such example:

	Political Party Identification				
Gender	**Democrat**	**Independent**	**Republican**	**Total**	***n***
Female	39.6%	35.7%	26.4%	100%	1063
Male	39.6%	35.7%	26.4%	100%	843

d) Graph for (b):



Graph for (c):



11.2 Beliefs of new employees

a) If the results for population of new employees were similar to these, gender and feelings of being overwhelmed would be independent.

b) There are many hypothetical populations for which these variables would be independent. Distributions should show percentages that are same for men and women. For example, if percentages for men and women both were 43%, these variables would be independent.

11.3 Williams College admission

a) These distributions refer to those of Y (admitted or not) at given categories of X (gender).

	Admitted			
Gender	**Yes**	**No**	**Total**	***n***
Male	18.2%	81.8%	100%	3195
Female	16.9%	83.1%	100%	3568

b) X and Y are dependent because the probability of a student being admitted differs by gender.

11.4 Happiness and gender

a) The response variable is happiness, and the explanatory variable is gender.

b)

	Happiness				
Gender	**Not Too Happy**	**Pretty Happy**	**Very Happy**	**Total**	***n***
Female	14.2%	54.7%	31.1%	100%	1082
Male	13.9%	56.9%	29.1%	100%	882

The conditional distributions of happiness look nearly identical between females and males, with about 14% being not too happy, about 56% being pretty happy, and 30% being very happy.

c) The following is an example of a population conditional distribution that is both consistent with this sample and for which happiness and gender are independent.

	Happiness				
Gender	**Not Too Happy**	**Pretty Happy**	**Very Happy**	**Total**	***n***
Female	14.0%	56.0%	30.0%	100%	1082
Male	14.0%	56.0%	30.0%	100%	882

11.5 Marital happiness and income

a)

	Happiness of Marriage			
Income	**Not Too Happy**	**Pretty Happy**	**Very Happy**	**Total**
Below	6	62	139	207
Average	7	125	283	415
Above	6	69	115	190

b)

	Happiness of Marriage				
Income	**Not Too Happy**	**Pretty Happy**	**Very Happy**	**Total**	***n***
Below	3%	30%	67%	100%	207
Average	2%	30%	68%	100%	415
Above	3%	36%	61%	100%	190

The conditional distributions of marital happiness for above-average and average income are nearly identical; however, among people of below-average income, the percentage of very happy people is lower, and the percentage of pretty happy people is larger.

c) Across all three income categories, the percentage of very happy people is much larger for marital happiness (always over 60%) than for general happiness (always below 40%).

11.6 What is independent of happiness?

Answers will vary. One possible answer is The region of the country in which you live, because the probability of being happy could well be the same in different regions, whereas we would expect happiness to be higher for those who believe in an afterlife, those with higher family income, those with better health, and those who are happier with their job.

11.7 Sample evidence about independence

Answers will vary depending on the column variable selected. For example, in 2008, the percentage in each of the 9 regions was about 30% for very happy, 55% for pretty happy, and 15% for not too happy. Independence seems plausible.

Section 11.2: Testing Categorical Variables for Independence

11.8 Lung cancer and smoking

a) Observed values:

	Lung Cancer		
Smoking	**Present**	**Absent**	**Total**
Smoker	605	185	790
Non-smoker	122	312	434

b) Expected values:

	Lung Cancer		
Smoking	**Present**	**Absent**	**Total**
Smoker	469	321	790
Non-smoker	258	176	434

There are fewer non-smokers but more smokers who are diagnosed with lung cancer than what would be expected under independence.

c) $$X^2 = \sum \frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$$

$$= \frac{(605-469)^2}{469} + \frac{(185-321)^2}{321} + \frac{(122-258)^2}{258} + \frac{(312-176)^2}{176} = 272.89$$

11.9 Happiness and gender

a) H_0: Gender and happiness are independent.
H_a: Gender and happiness are dependent.

b) If the null hypothesis of independence is true, it is not unusual to observe a chi-square value of 1.04 or larger because the probability is 59% of this occurring. Hence, there is no evidence of an association between gender and happiness.

11.10 What gives P-value = 0.01?

a) 1 : 6.63
b) 2 : 9.21
c) 4 : 13.28
d) 5 : 32.00
e) 5 : 32.00

11.11 Marital happiness and income

a) H_0: Marital happiness and family income are independent.
H_a: Marital happiness and family income are dependent.

b) $df = (3 - 1)(3 - 1) = 4$

c) (i) The expected value of the chi-squared statistic is $df = 4$.

(ii) The standard deviation is $\sqrt{2df} = \sqrt{2(4)} = \sqrt{8} = 2.8$. 4.58 is $(4.58 - 4)/2.8 = 0.21$, or about 0.2 standard deviations above the expected value under independence.

d) We would need a chi-squared value of 9.49 to get a P-value of exactly 0.05.

11.11 (continued)

e) The P-value for a chi-squared statistic of 4.58 is greater than 0.25 using Table C and 0.33 using technology. This is a very large P-value. Do not reject the null hypothesis, there is no evidence of an association between marital happiness and income.

11.12 First and second free throw independent?

a)

	Made Second		
Made First	**Yes**	**No**	**Total**
No	48	5	53
Yes	251	34	285

b) It does not seem that his success on the second shot depends on whether he made the first. The chi-squared statistic is small and the P-value is large. It would not be unusual for a random sample to have a chi-squared statistic of this size.

11.13 Cigarettes and marijuana

a)

	Marijuana			
Cigarettes	**Yes**	**No**	**Total**	***n***
Yes	61.1%	38.9%	100%	1495
Above	5.9%	94.1%	100%	781

This conditional distribution suggests that marijuana use is much more common for those who have smoked cigarettes than for those who have not.

b) 1) The assumptions are that there are two categorical variables (cigarette use and marijuana use in this case), that randomization was used to obtain the data, and that the expected count was at least five in all cells.

2) H_0: Cigarette use and marijuana use are independent.
H_a: Cigarette use and marijuana use are dependent.

3) $X^2 = 642.0$

4) The P-value is approximately 0.

5) If the null hypothesis were true, the probability would be close to 0 of getting a test statistic at least as extreme as the value observed. This P-value is quite low. We have extremely strong evidence that marijuana use and cigarette use are associated.

11.14 Smoking and alcohol

a) False

b) The z statistic would be the square root of the chi-squared statistic, so $z = \sqrt{X^2} = \sqrt{451.404} = 21.25$. The P-value would remain 0.000.

11.15 Help the environment

a) H_0: Willingness to accept cuts to help the environment and being in school or retired are independent.
H_a: Willingness to accept cuts to help the environment and being in school or retired are dependent.

b) $r = 2$ and $c = 5$; thus, $df = (r - 1)(c - 1) = (2 - 1)(5 - 1) = 4$

c) For a P-value of 0.05, $X^2 = 9.49$ and for a P-value of 0.025, $X^2 = 11.14$. For this data, $X^2 = 9.56$, so:

(i) The P-value is less than 0.05.

(ii) The P-value is greater than 0.025.

d) (i) With a significance level of 0.05, there is evidence for an association between helping the environment and whether someone is in school or retired.

(ii) With a significance level of 0.025, there is not enough evidence for an association between helping the environment and whether someone is in school or retired.

11.16 Primary food choice of alligators

a)

	Primary Food					
Lake	**Fish**	**Invertebrates**	**Birds & Reptiles**	**Other**	**Total**	***n***
Hancock	54.5%	7.3%	14.6%	23.6%	100%	55
Trafford	24.5%	34.0%	22.6%	18.9%	100%	53

b) H_0: The distribution of primary food choice is the same for alligators caught in lakes Hancock and Trafford (homogeneity).
H_a: The distributions differ for the two lakes.

c) $df = (2 - 1)(4 - 1) = 3$, so we expect the chi-squared statistic to be 3 with a standard deviation of $\sqrt{2df} = \sqrt{2(3)} = \sqrt{6} = 2.5$. Since 16.79 is $(16.79 - 3)/2.5 = 5.5$, or 5.5 standard deviations about the expected value of 3, it is considered extreme.

d) Since the P-value is less than 0.001, there is strong evidence that the distribution of primary food choice of alligators differs in the two lakes.

11.17 Cognitive behavioural therapy and anxiety

a)

	Anxiety Level		
Treatment	**Low**	**High**	**Total**
Placebo	684	45	729
CBT	738	29	767

b) 1) The assumptions are that there are two categorical variables (treatment and anxiety level) that randomization was used to obtain the data, and the expected count was at least five in all cells.
2) H_0: Treatment and anxiety level are independent.
H_a: Treatment and anxiety level are dependent.
3) $X^2 = 4.5$
4) The P-value is 0.03.
5) If the null hypothesis were true, the probability would be 0.03 of getting a test statistic at least as extreme as the value observed. This is strong evidence against the null hypothesis. It is plausible that the alternative hypothesis is correct; and treatment and anxiety are dependent.

11.18 *z*-test for anxiety study

a) The population proportion, p_1, is the proportion of patients who went through placebo for whom anxiety levels were increased, and p_2, is the proportion of patients who went through CBT. The hypotheses would be:
H_0: Treatment and anxiety level are independent $(H_0: p_1 = p_2)$.
H_a: Treatment and anxiety level are dependent $(H_a: p_1 \neq p_2)$.

b) The results in the two tests are identical. The P-value is the same, and the chi-squared statistic is the square of the *z*-test statistic.

11.19 Severity of fever after flu shot

a) H_0: The distribution of severity of fever is the same in the active and placebo group (homogeneity).
H_a: The distributions differ.

b) Since $df = (2 - 1)(3 - 1) = 2$, the expected chi-squared statistic would be around 2, with a standard deviation of $\sqrt{2df} = \sqrt{2(2)} = \sqrt{4} = 2$. 2.49 is $(2.49 - 2)/2 = 0.245$, or about one-fourth of a standard deviation above the expected value, which makes it not extreme.

c) Since a P-value of 0.287 is larger than any reasonable significance level, there is no evidence that the distribution of the severity of fewer differs between the active and placebo group.

11.20 What is independent of happiness?

The results for this exercise will be different based on the variables selected by each student.

11.21 Testing a genetic theory

a) H_0: $p = 0.75$ (The probability of a green seedling is 0.75.)

H_a: $p \neq 0.75$ (The probability of a green seedling is not 0.75.)

b) Under the null hypothesis, we would expect 1103(0.75) = 827.25 green seedlings and 1103(0.25) = 275.75 yellow seedlings. The chi-squared goodness-of-fit statistic is then

$$X^2 = \frac{(854 - 827.25)^2}{827.25} + \frac{(249 - 275.75)^2}{275.75} = 3.46 \text{ with } df = 2 - 1 = 1.$$

c) The P-value is 0.06. The probability of obtaining a test statistic as extreme as that observed, assuming the null hypothesis is true, is 0.06. There is evidence against the null, but not very strong.

11.22 Footfall by quarters

a) H_0: The probabilities of arrival of customers in a given quarter are same ($p = 1/4$).
H_a: The probabilities are not the same.

b) Under the null hypothesis, we would expect 0.25 × 1098 = 275.5 customers each quarter. The chi-squared goodness-of-fit statistic is:

$$X^2 = \frac{(198 - 275.5)^2}{275.5} + \frac{(340 - 275.5)^2}{275.5} + \frac{(318 - 275.5)^2}{275.5} + \frac{(242 - 275.5)^2}{275.5} = 48.31$$

c) $df = 4 - 1 = 3$

d) The P-value is approximately 0. Since this is less than the significant level of 5%, there is an evidence that probabilities of arrival of customers in a given quarter are not the same.

11.23 Checking a roulette wheel

a) P(Each Pocket) = 1/37

b) 3700(1/37) = 100

c) $$X^2 = \sum \frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}} = \frac{(110 - 100)^2}{100} = 1$$

d) $df = c - 1 = 37 - 1 = 36$; the P-value is 0.54; Since the P-value is quite large, there is not strong evidence that the roulette wheel is not balanced.

Section 11.3: Determining the Strength of the Association

11.24 Democrat, race and gender

a) 212/300 = 0.707 of blacks and 422/1468 = 0.287 of whites identify as Democrat. The difference between blacks and whites who identify as Democrat is 0.707 – 0.287 = 0.42, so the proportion of blacks identifying as Democrat is 42 percentage points higher than for whites. 421/1081 = 0.389 of females and 278/879 = 0.317 of males identify as Democrat. The difference between females and males who identify as Democrat is 0.389 – 0.317 = 0.072, so the proportion of females identifying as Democrat is 7 percentage points higher compared to males. Race has a stronger association with whether someone identifies as Democrat because the difference of proportion is much larger. The difference between blacks and whites is 0.40. Thus, race has a stronger association with whether one identifies as a Democrat.

b) The proportion identifying as Democrat is 0.707/0.287 = 2.46 times higher for blacks than for whites. (Or, blacks are 2.46 times more likely to identify as Democrat than whites.) The proportion of females identifying as Democrat is 0.389/0.316 = 1.23 times higher (or 23% higher) higher than the proportion of males. Race has a stronger association with whether someone identifies as Democrat because the ratio of proportion is larger.

c) The odds for blacks are 0.707/(1 – 0.707) = 2.4, which can be written as 2.4:1 = 1:0.42 or 100:42. For blacks, for every 100 identifying as Democrat, there are 42 not identifying as Democrat. The odds for whites are 0.287/(1 – 0.287) = 0.4, which is 0.4:1 = 4:10 or 100:250. For whites, for every 100 identifying as Democrat, there are 250 not identifying as Democrat. The odds ratio is 2.4/0.4 = 6. The odds for identifying as Democrat are 6 times higher for blacks than for whites.

11.25 Death penalty associations

a) False, a larger chi-squared value might be due to a larger sample size rather than a stronger association.
b) Yes. The confidence interval for race effect covers larger values than confidence interval for gender effect, so race has the stronger association with opinion on death penalty.

11.26 Smoking and alcohol

a) 64% of those who had not smoked cigarettes also used alcohol, whereas 97% of those who had smoked cigarettes had used alcohol. 0.97 – 0.64 = 0.33, so the proportion who had used alcohol is 0.33 higher for cigarette users than for non-cigarette users.
b) The proportion who had used alcohol is 0.97/0.64 = 1.5 times higher for those who used cigarettes compared to those who didn't.
c) The odds of having used alcohol for cigarette users are 0.97/0.03 = 32, or 32:1. For cigarette users, for every 32 using alcohol, one is not using alcohol. For nonusers of cigarettes, the odds of having used alcohol are 0.64/0.36 = 1.8, or 1.8:1. For nonusers of cigarettes, for every 1.8 using alcohol, one is not using it. The odds ratio is 32/1.8 = 17.8. The odds of having used alcohol are about 18 times higher for students who also used cigarettes compared to those who didn't.

11.27 Gender and dominant hand usage

There are several possible answers. The proportion of female individuals who are right handed = 0.92. The proportion of male individuals who are right handed = 0.83. Ratio = 0.92/0.83 = 1.1. Female individuals are 1.1 times more likely to be right handed as compared to the male individuals. (That is, relatively more females are right handed as compared to the males.)

11.28 Smelling and mortality

No, the odds ratio being greater than 3 does not imply that the relative risk is greater than 3. With the given information, we do not know whether the proportion dying is more than 3 times larger in the anosmic group. We only know that the odds of dying are more than three times larger. We cannot interpret an odds ratio as a relative risk.

11.29 Vioxx

a) The proportion with myocardial infarctions in the naproxen group was 0.001 – 0.004 = –0.003, or 0.3 percentage points lower than the proportion in the rofecoxib group.
b) The proportion with myocardial infarctions in the naproxen group is 0.2 times (or 80%) lower than the proportion in the rofecoxib group.
c) Myocardial infarctions were 1/0.2 = 5 times more likely.

11.30 Egg and cell derived vaccine

a) 26/3900 = 0.0067 of subjects that received the cell-derived flu vaccine developed the flu, whereas 24/3900 = 0.0062 of subjects that received the egg-derived flu vaccine developed the flu. The relative risk of developing the flu is $(26/3900)/(24/3900) = 1.1$. Subjects that received the cell-derived flu vaccine were 1.1 times as likely as subjects that received the egg-derived flu vaccine to develop the flu.
b) The odds ratio is $\dfrac{(26/3900)/(1-26/3900)}{(24/3900)/(1-24/3900)} = 1.1$. The odds of developing the flu are 10% larger in the cell-derived vaccine group.
c) Yes, both relative risk and odds ratio are close to 1, which is the value that occurs when the probabilities are roughly the same in the two groups.

11.31 Risk of dying for teenagers

a) The difference of proportions is 0.00135 – 0.00046 = 0.00089. The proportion of male teenagers who die is 0.00089 higher than the proportion of female teenagers who die.
b) The relative risk is 0.00135/0.00046 = 2.9. Male teenagers are 2.9 times more likely to die than are female teenagers.
c) The relative risk seems more useful because it shows there is a substantial gender effect, which the difference does not show when both proportions are close to 0.

11.32 Recreation and happiness

(a) Since the P-value is less than 0.001, there is strong evidence for an association between recreation and happiness.
(b) No, not generally. Large X^2 values can occur even for weak (but still significant) associations.
(c) The percentage of being not too happy is 37/57 – 9/187 = 0.60099, or about 60 percentage points higher for those who often indulge into recreation compared to those seldom indulge into recreation.
(d) Those who seldom indulge into recreation are about (37/57)/(9/187) = 13.52, or about 13 times as likely to be not too happy compared to those who often indulge into recreation.

11.33 Party ID and gender

a) The proportion of females who identify as Republican is 244/1063 – 198/843 = –0.005, or 0.005 smaller than the proportion for males.
b) The proportion of females who identify as Democrat is 421/1063 – 278/843 = 0.066, or 0.066 higher than the proportion for males.
c) The proportion of females who identify as Republican is $(244/1063)/(198/843) = 0.98$, or 2% lower than the proportion for males.
d) The proportion of females who identify as Democrat is $(421/1063)/(278/843) = 1.2$, or 20% higher than the proportion for males.
e) There is a rather weak association between gender and whether identifying as Republican, as seen from (a) and (c). There is a stronger association between gender and whether identifying as Democrat, as seen in (b) and (d).

11.34 Chi-squared versus measuring association

The analysis in parts (c) and (d) describes the association observed in the data, whereas the chi-squared test is an inferential procedure about the association in the population (here: the association between recreation and happiness for all residents of the city).

Section 11.4: Using Residuals to Reveal the Pattern of Association

11.35 Standardized residuals for happiness and income

a) The standardized residual indicates the number of standard errors that the observed count falls from the expected count. In this case, the observed count falls 2.49 standard errors below the expected count.
b) The standardized residuals highlighted in green designate conditions in which the observed counts are much higher (more than 3 standard deviations) than the expected counts, relative to what we'd expect due to sampling variability. For people with above-average income, many more were very happy than what independence between income and happiness would predict. For people with below-average income, many more were not happy than what independence would imply.
c) The standardized residuals highlighted in red designate conditions in which the observed counts are much lower (more than 3 standard deviations) than the expected counts, relative to what we'd expect due to sampling variability. For people with below-average income, many fewer were very happy than what independence would imply. For people with average income, many fewer were not happy than what independence would predict.

11.36 Happiness and religious attendance

a) The large chi-squared statistic and small P-value indicate that we have strong evidence that there is an association between happiness and religious attendance.
b) The cell for attendance at most several times a year and "not too happy," and that for attendance every week or more and "very happy" have strong evidence that in the population there are more people than if the variables were independent.
c) The cell for attendance at most several times a year and "very happy" gives strong evidence that in the population there are fewer people than if the variables were independent.

11.37 Recreation and happiness

a) The relatively small standardized residual of –0.9 indicates that the observed count for this cell is only 0.5 standard deviations below expected count. This is not unusual under the null hypothesis of independence, so it is not strong evidence that there is a true effect in that cell.

b) These are less than 1% chance that a standardized residual would exceed 3 in absolute value, if the variables were independent. Based on this criterion, the following cells would lead us to infer that the population has many more cases than would occur if recreation and happiness were independent. Consider the cell (seldom, not too happy), a high absolute value of standardized residual as 8.2 implies that people who do not indulge into recreation are more likely to be not too happy than would be expected if the variables were independent. Consider the cell (often, very happy), a high absolute value of standardized residual as 5.1 implies that the people who often indulge into recreation are more likely to be very happy than would be expected if the variables were independent.

11.38 Happiness and marital status

(i) There are more people who say they are "very happy" in the married category than we would expect if the variables were independent.

(ii) There are fewer people who say they are "very happy" in the divorced and never married categories than we would expect if the variables were independent.

11.39 Gender gap?

There are more women who identify as Democrats, and fewer men who identify as Democrats than would be expected if there were no association between political party and gender. It does seem that political party and gender are not independent – that there is an association.

11.40 Ideology and political party

a) There is a very large chi-squared statistic: $X^2 = 723$, $df = 42$, and the P-value is approximately 0. There is extremely strong evidence for an association between political ideology and party identification.

b) The standardized residuals help us to understand the natural of the association. It appears that people who think of themselves as (extremely) liberal tend to identify strongly with the Democratic Party, whereas people who think of themselves as (extremely) conservative tend to identify with the Republican party.

Section 11.5: Fisher's Exact and Permutation Tests

11.41 Keeping old dogs mentally sharp

a)

	Could Solve Task		
Care And Diet	**Yes**	**No**	**Total**
Standard	2	6	8
Extra	12	0	12

b) 1) There are two binary categorical variables, and randomization was used.

2) H_0: Care and ability to solve the task are independent. ($H_0: p_1 = p_2$)

H_a: Care and ability to solve the task are dependent. ($H_a: p_1 \neq p_2$)

3) Test statistic: 2

4) The P-value is 0.001.

5) If the null hypothesis were true, the probability would be 0.001 of getting a test statistic at least as extreme as the value observed. The P-value is quite low (lower than a significance level of 0.05, for example); we can reject the null hypothesis. We have strong evidence that a dog's care and diet are associated with its ability to solve a task.

c) It is improper to conduct the chi-squared test for these data because the expected cell counts are less than 5 for at least some cells.

11.42 Tea-tasting results

	Prediction		
Actual	**Milk**	**Tea**	**Total**
Milk	4	0	4
Tea	0	4	4
Total	**4**	**4**	**8**

The P-value for a Fisher's Exact Test of Independence is 0.014. There is strong evidence of an association between the actual tea preparation and the prediction with respect to the tea preparation.

11.43 Claritin and nervousness

a) The P-value for the small-sample test is 0.24. It is plausible that the null hypothesis is true and that nervousness and treatment are independent.

b) It is not appropriate to conduct the chi-squared test for these data because two cells have an expected count of less than five, (2.5 and 3.5, respectively).

11.44 AIDS and condom use:

a) p_1 is the population proportion of those who always used condoms who became infected, and p_2 is the population proportion of those who did not always use condoms who became infected.
H_0: Condom use and infection are independent.
H_a: Condom use and infection are dependent.

b) A chi-squared test would not be appropriate because the expected count of the cell for those who did not use condoms and who did become infected is less than five. The Fisher's exact test gives a P-value of 0.0007, which rounds to 0.001. This is very strong evidence against the null hypothesis. There seems to be an association between condom use and HIV infection.

11.45 Fitness workshop worthwhile?

a)

	Shifted to Proper Meal Timings?		
Gender	**Yes**	**No**	**Total**
Female	5	0	5
Male	6	6	12

b) 462 out of 6188, or 0.075, of the tables have first cell count as large or larger than 5.

c) The (exact or permutation) P-value of 0.075 is greater than the significance level of 0.05, which indicates insufficient evidence to conclude that the probability of shifting to proper meal timings is larger in female participants.

11.46 Proper meal timings enhance fitness?

a) H_0: The conditional distribution of the responses on fitness status after 2 months is same for both who shifted to proper meal timings and who did not, i.e. shifting to proper meal timings and fitness status are independent.
H_a: The conditional distributions differ, i.e. shifting to proper meal timings and fitness are dependent.

b) X^2

c) The statistic would follow a chi-squared distribution with $df = 1$.

d) The permutation P-value is 475/10000 = 0.0475. There is strong evidence that the distribution of responses on fitness status after 2 months differs for both who shifted to proper meal timings and who did not.

Chapter Problems: Practicing the Basics

11.47 Female participation in defense services?

a)

	Opinion	
Gender	**Yes**	**No**
Female	91%	9%
Male	91%	9%

11.47 (continued)

b) If results for the entire population are similar, it does seem possible that gender and opinion about active participation of women in defense services are independent. The percentages of men and women who would have had opinion of active participation of women in defence services may be the same.

11.48 Down syndrome diagnostic test

a) $P(\text{Positive} \mid D) = 48/54 = 0.8889$, or about 89% and $P(\text{Positive} \mid D^c) = 1307/5228 = 0.25$, or 25%.

b) For the Down cases, 89% were correctly diagnosed. For the unaffected cases, 25% get a negative result. The test seems fairly good, but there are a good number of false positives and false negatives.

c) $P(\text{Positive} \mid D^c) = 48/(48+1307+6) = 48/1361 = 0.035$; Of the positive cases, only 0.035 truly have Down syndrome. This result is not surprising because there are so few cases overall. The fairly large number of false positives will overwhelm the much smaller number of actual cases.

11.49 Down and chi-squared

1) The assumptions are that there are two categorical variables (Down syndrome status and blood test result), that randomization was used to obtain the data, and that the expected count was at least five in all cells.
2) H_0: Down syndrome status and blood test result are independent.
   H_a: Down syndrome status and blood test result are dependent.
3) $X^2 = 114.4;\ df = 1$
4) P-value: 0.000
5) If the null hypothesis were true, the probability would be almost 0 of getting a test statistic at least as extreme as the value observed. There is very strong evidence of an association between test result and actual status.

11.50 Herbs and the common cold

a) The response variable is whether or not the individual's cold symptoms improved, and the explanatory variable is treatment (placebo versus Immumax).

b) I would explain that if improved cold symptoms did not depend on whether one took Immumax or placebo, then it would be quite unusual to observe the results actually obtained. This provides relatively strong evidence of improved cold symptoms for adults taking Immumax.

11.51 Study hours and grades

a) If the two variables were independent, it would mean that the chance of any particular effect on grades category (such as positive) would be identical for each category of study hours per week.

b) An expected cell count is the number of cases we'd expect in a given cell if the two variables were not associated (i.e., were independent). For the first cell, the expected count = (90 × 53)/200 = 23.9.

c) (i) Those with least study hours per week tend to have a positive effect on their grades.
   (ii) Those with highest study hours per week tend to have a negative effect on their grades.

11.52 Gender gap?

1) The assumptions are that there are two categorical variables (party identification and gender), that randomization was used to obtain the data, and that the expected count was at least five in all cells.
2) H_0: Party identification and gender are independent.
   H_a: Party identification and gender are dependent.
3) $X^2 = 10.04;\ df = 2$
4) The P-value is 0.0066.
5) If the null hypothesis were true, the probability would be 0.0066 (less than any reasonable significance level) of getting a test statistic at least as extreme as the value observed. We have very strong evidence that party identification depends on gender.

11.53 Gender gap in employment?

a) $df = 1$; the sampling distribution is the chi-squared probability distribution for $df = 1$.

b) The small P-value indicates that it is not plausible that the null hypothesis is correct and that income and job satisfaction are not independent.

c) With a 0.05 significance level, we would reject the null hypothesis. No, we cannot accept H_0 and conclude that employment is independent of gender?

11.54 Aspirin and heart attacks for women

a) (i) The assumptions are that there are two categorical variables (group and cardiovascular event), that randomization was used to obtain the data, and that the expected count was at least five in all cells.

(ii) H_0: Group and cardiovascular event are independent.
H_a: Group and cardiovascular event are dependent.

(iii) $X^2 = 10.66;\ df = 2$

(iv) The P-value is 0.005.

(v) The P-value is very small. If the null hypothesis were true, the probability would be 0.005 of getting a test statistic at least as extreme as the value observed. We have very strong evidence that there is an association between cardiovascular event and group for women.

b) The proportion of women on placebo who had a stroke was 0.013. The proportion of those on aspirin who had a stroke is 0.011. Thus, the relative risk is 0.013/0.011 = 1.2. Women on placebo are 1.2 times as likely as those on aspirin to have a stroke.

11.55 Crossing Peas

a)

Pea Type				
RY	**RG**	**WY**	**WG**	**Total**
315	108	101	32	556

b) P(RY) = 9/16, P(RG) = 3/16, P(WY) = 3/16, and P(WG) = 1/16

c)

Expected Values				
RY	**RG**	**WY**	**WG**	**Total**
$556\left(\frac{9}{16}\right)$	$556\left(\frac{3}{16}\right)$	$556\left(\frac{3}{16}\right)$	$556\left(\frac{1}{16}\right)$	
$= 312.75$	$= 104.25$	$= 104.25$	$= 34.75$	556

d) $df = 4 - 1 = 3$

e) The P-value will be large because the chi-square statistic of 0.47 is well below the expected value of 3, so area above 0.47 under chi-squared curve with $df = 3$ is large.

11.56 Women's role

a) The difference of proportions based on gender is 0.148 – 0.159 = –0.011.

b) The difference of proportions based on education is 0.390 – 0.117 = 0.273.

c) Educational level seems to have the stronger association with opinion. The difference between the proportions of the two educational groups is much larger than the difference between the proportions of the two genders. Education level makes a larger difference than gender.

11.57 Seat belt helps?

a) The proportion of those who were injured given that they did not wear a seat belt is 0.125. The proportion of those who were injured given that they wore a seat belt is 0.064. The difference between proportions is 0.125 – 0.064 = 0.061. The proportion who were injured is 0.06 higher for those who did not wear a seat belt than for those who did wear a seat belt.

b) The relative risk is 0.125/0.064 = 1.95. People were 1.95 times as likely to be injured if they were not wearing a seat belt than if they were wearing a seat belt.

11.58 Serious side effects

a)

	Serious Side Effect		
Treatment	**Yes**	**No**	**Total**
Zelnorm	13	11,601	11,614
Placebo	1	7030	7031

b) The relative risk is $(13/11,614)/(1/7031) = 7.87$. Patients receiving the drug were 7.87 times more likely to experience a serious side effect than patients receiving placebo.

c) The odds ratio is $\dfrac{(13/11,614)/(1-13/11,614)}{(1/7031)/(1-1/7031)} = \dfrac{13/11,601}{1/7031} = 7.88$. The odds of experiencing a serious side effect were 7.88 times larger for patients receiving the drug rather than placebo.

d) There is sufficient evidence to reject the null hypothesis and conclude the probability of a side effect differs between the drug and placebo group.

11.59 Pesticides

a) Relative risk is $(29/127)/(19,485/26,571) = 0.31$. The proportion of organic food samples with pesticide residues present was 69% (or 0.31 times) lower than the proportion of conventional food samples with pesticide residues present.

b) The proportion of conventional food samples with pesticide residues present was 3.2 (from 1/0.31) times larger than the proportion for organic food samples with pesticide residues.

c) The odds ratio is $\dfrac{(29/127)/(1-29/127)}{(19,485/26,571)/(1-19,485/26,571)} = \dfrac{29/89}{19,485/7086} = 0.11$. The odds of finding pesticide residues on organic food samples were 89% (or 0.11 times) lower than on conventional food samples.

d) The odds of finding pesticide residues on conventional food samples were 1/0.11 = 9.1 times higher than on organic food samples.

e) The *proportion* is only about 3.2 times larger (the relative risk), but the *odds* are more than 9 times larger. This statement confuses relative risk with the odds ratio.

11.60 Race and party ID

a) The expected count for the first cell is $(275 \times 651)/1791 = 100.0$. It is what we would expect if there were no association.

b) The standardized residual of 12.5 for the first cell indicates that the observed count falls 12.5 standard errors from the expected count. It is very likely that the population proportion in this cell is higher than what would be expected if the two variables were independent.

c) The four corner cells indicate that there are many more blacks who are Democrats and many more whites who are Republicans than we would expect if political party and race were independent. Similarly, there are far fewer blacks who are Republicans and far fewer whites who are Democrats than we would expect if these two variables were independent.

11.61 Happiness and sex

It seems as though those with no partners or two or more partners are less likely to be very happy than would be expected if these variables were not associated, and those with one partner are more likely to be very happy than would be expected if these variables were independent.

11.62 Education and religious beliefs

For each of these cells (bachelor or graduate and fundamentalist, bachelor or graduate and liberal and less than high school and fundamentalist), the observed value is much higher than what we would expect if the variables were independent, relative to what we'd expect due to sampling variability.

11.63 TV and aggression

a) H_0: Amount of TV watching and aggression are independent.
H_a: Amount of TV watching and aggression are dependent.
If p_1 is the population proportion of those who are aggressive in the group that watches less than one hour of TV per day and p_2 is the population proportion of those who are aggressive in the group that watches more than one hour of TV per day, then the hypotheses can be expressed as H_0: $p_1 = p_2$ and H_a: $p_1 \neq p_2$.

b) The P-value is 0.0001. This is a very small P-value. If the null hypothesis were true, the probability would be close to 0 of getting a test statistic at least as extreme as the value observed. We have very strong evidence that TV watching and aggression are associated.

11.64 Botox side effects

1) The assumptions are that there are two binary categorical variables, and randomization was used.
2) H_0: Treatment and pain status are independent.
H_a: Treatment and pain status are dependent.
3) Test statistic: 9
4) The P-value is 0.12.
5) If the null hypothesis were true, the probability would be 0.12 of getting a test statistic at least as extreme as the value observed. It is plausible that the null hypothesis is correct and that treatment and pain status are independent.

11.65 Clarity of diamonds

a) Technology will confirm that $X^2 = 0.267$.

b) No, many cell counts are very small, leading to expected cell counts that are less than 5, so the sampling distribution of X^2 may not be approximately chi-squared.

c) Approximate permutation P-value is 9908/10,000 = 0.9908. This P-value is large and almost 1. There is no evidence that the clarity depends on whether the diamond's cut is good or fair.

11.66 Benford's Law

a)

Leading Digit	Expected Value
1	0.301(130) = 39.13
2	0.176(130) = 22.88
3	0.125(130) = 16.25
4	0.097(130) = 12.61
5	0.079(130) = 10.27
6	0.067(130) = 8.71
7	0.058(130) = 7.54
8	0.051(130) = 6.63
9	0.046(130) = 5.98

b) Chi-squared goodness-of-fit test with $df = 9 - 1 = 8$. From technology: $X^2 = 7.2$, the P-value is 0.5147. Because P-value is not small, there is no evidence that the distribution deviates significantly from Benford's Law.

Chapter Problems: Concepts and Investigations

11.67 Student data

Each student's short report will be different, but could include the following findings.

From MINITAB:

```
Rows: religiosity   Columns: life_after_death
            n       u       y     All
0           8       3       4      15
        3.250   4.000   7.750  15.000
1           5      11      13      29
        6.283   7.733  14.983  29.000
2           0       2       5       7
        1.517   1.867   3.617   7.000
3           0       0       9       9
        1.950   2.400   4.650   9.000
All        13      16      31      60
       13.000  16.000  31.000  60.000
Cell Contents:      Count
                    Expected count
Pearson Chi-Square = 21.386, DF = 6, P-Value = 0.002
```

11.68 Marital happiness decreasing?

The conditional distributions show a slight trend for the response of "very happy" such that the percentages drop over time. For the most part, percentages are in the high 60's in the beginning, dropping to the low 60's over time. The standardized residuals show a similar pattern. They're higher in the first few years with one as high as 3.92. Through the years, the standardized residuals tend to get closer to zero, and then become negative in a number of the most recent years.

11.69 Another predictor of happiness?

The one-page report will be different depending on the variable that each student finds to be associated with happiness.

11.70 Market price associated with factor cost?

We would expect these two variables to be associated. As factor cost, i.e. the price of mangoes rises, the market price of mango juice will also rise.

11.71 Babies and gray hair

a)

	Has Young Children	
Gray Hair	**Yes**	**No**
Yes	0	4
No	5	0

b)

	Has Young Children	
Gray Hair	**Yes**	**No**
Yes	0%	100%
No	100%	0%

There does seem to be an association. All women in the sample who have gray hair do not have young children, whereas all women in the sample who do not have gray hair do have young children.

c) There often are third factors that influence an association. Gray hair is associated with age (older women being more likely to be gray), and age is associated with having or not having young children (older women being less likely to have young children). Just because two things are associated, doesn't mean that one causes the other.

11.72 When is chi-squared not valid?

The examples of contingency tables can differ for each student; however, a student could choose to include numbers for which the expected scores would be less than five.

11.73 Gun homicide in United States and Britain

a) The proportion in the U.S. is 0.000047. The proportion in Britain is 0.00001. The difference of proportions with the U.S. as group 1 is 0.000037, and with Britain as group 1 is –0.000037. The only thing that changes is the sign.

b) The relative risk with the U.S. as group 1 is 0.000047/0.00001 = 4.7. The relative risk with Britain as group 1 is 0.00001/0.000047 = 0.213. One value is the reciprocal of the other.

c) When both proportions are so small, the relative risk is more useful for describing the strength of association. The difference between proportions might be very small, even when one is many times larger than the other.

11.74 Colon cancer and race

The relative risk of developing colorectal cancer was $(46.5/100{,}000)/(57.3/100{,}000) = 0.81$ times lower for white residents of North Carolina than for African American residents of North Carolina during 2002-2006. Thus, African American residents of North Carolina were 19% more likely to have been diagnosed with colorectal cancer (1 – 0.81 = 0.19) than white residents.

11.75 True or false: $X^2 = 0$

False

11.76 True or false: Group 1 becomes Group 2

True

11.77 True or false: Relative risk

False

11.78 True or false: Relative risk versus odds ratio

False

11.79 True or false: Statistical but not practical significance

True

11.80 Statistical versus practical significance

Given enough participants in a study, we might find a very weak association to be statistically significant. It is important to examine the size of an association in addition to its statistical significance. Otherwise, the significant association might not be practically important.

11.81 Normal and chi-squared with $df = 1$

a) The chi squared value for a right-tail probability of 0.05 and $df = 1$ is 3.84, which is the z value for a two-tail probability of 0.05 squared: (1.96)(1.96) = 3.84.

b) The chi-squared value for P-value of 0.01 and $df = 1$ is 6.64 which is (apart from rounding) the z value for a two-tail probability of 0.05 squared: (2.58)(2.58) = 6.63.

♦♦11.82 Multiple response variables

a) Because participants were able to give more than one response, these are dependent samples, and so it is not valid to do a chi-squared test. We would need different participants in each cell.

b) If we only look at men versus women for *one* of the factors, the data are independent and we can use a chi-squared test. For example, here is the contingency table for factor A.

	Income Gap Responsible	
Gender	**Yes**	**No**
Men	60	40
Women	75	25

♦♦11.83 Standardized residuals for 2×2 tables

The two observed values in a given row (or column) must add up to the same total as the two expected values in that same row (or column). Thus, we know that if one of these two expected values is above its related observed count, then the other expected value in that same row (or column) must be below its related observed count. For example, let's say that we have a row with observed count of 50 in one cell and an observed count of 50 in the other cell. If the expected count for the first cell is 60, then the expected count for the other must be 40. Both pairs must add up to the same number – in this case, 100.

11.84 Degrees of freedom explained

a) The order of the calculations is given in the table.

	Vote for Female President		
Political Views	**Yes**	**No**	**Total**
Extremely Liberal	56	1st: 58 – 56 = 2	58
Moderate	490	2nd: 509 – 490 = 19	509
Extremely Conservative	4th: 61 – 3 = 58	3rd: 24 – 2 – 19 = 3	61
Total	**604**	**24**	**628**

b) The order of the calculations is given in the table.

	Vote for Female President		
Political Views	**Yes**	**No**	**Total**
Extremely Liberal	2nd: 58 – 2 = 56	1st: 24 – 3 – 19 = 2	58
Moderate	3rd: 509 – 19 = 490	19	509
Extremely Conservative	4th: 61 – 3 = 58	3	61
Total	**604**	**24**	**628**

11.85 What is *df*?

For the top row the last number must be 33 because the numbers must add up to the row total of 100. For the bottom row, the numbers must add to the column totals of 40 for each column.

	A	**B**	**C**	**D**	**E**	**Total**
	24	21	12	10	100 – 67 = 33	100
	40 – 24 = 16	40 – 21 = 19	40 – 12 = 28	40 – 10 = 30	40 – 33 = 7	100
Total	**40**	**40**	**40**	**40**	**40**	**200**

♦♦11.86 Variability of chi-squared

a) When *df* is large enough, the chi-squared distribution is fairly bell-shaped, so about 95% of cases fall within two standard deviations of the mean. For the chi-squared distribution (which has *df* as its mean), the mean plus and minus two standard deviations would be $\mu \pm 2\sigma$, or $df \pm 2\sqrt{2(df)}$.

b) $df \pm 2\sqrt{2(df)} = 8 \pm 2\sqrt{2(8)} = 8 \pm 2\sqrt{16} = 8 \pm 2(4) = 8 \pm 8$, or (0, 16).

The chi-squared distribution always has 0 as its lower limit. The chi-squared table tells us that for 8 *df* and a P-value of 0.05, the chi-squared value is 15.5. This would mark the upper 95% of the curve.

♦♦11.87 Explaining Fisher's exact test

a) The twenty distinct possible samples that could have been selected are as follows:

F1, F2, F3	F1, F2, M1	F1, F2, M2	F1, F2, M3
F1, F3, M1	F1, F3, M2	F1, F3, M3	F1, M1, M2
F1, M1, M3	F1, M2, M3	F2, F3, M1	F2, F3, M2
F2, F3, M3	F2, M1, M2	F2, M1, M3	F2, M2, M3
F3, M1, M2	F3, M1, M3	F3, M2, M3	M1, M2, M3

This contingency table shows that two males were chosen (M1 and M3) and one was not (M2). It also shows that one female was chosen (F2) and two were not (F1 and F3).

11.87 (continued)

b) 10 of the 20 tables have a first cell count of 2 or 3. Under a null hypothesis of no gender bias versus an alternative hypothesis of a preference for males, observing tables with 2 or 3 males selected (i.e., a 2 or 3 in the first cell) has a probability of 0.5, which is the P-value from Fisher's exact test.

These samples are as follows:

F1, M1, M2	F1, M1, M3	F1, M2, M3	F2, M1, M2
F2, M1, M3	F2, M2, M3	F3, M1, M2	F3, M1, M3
F3, M2, M3	M1, M2, M3		

♦♦11.88 Likelihood-ratio chi-squared

a) If the observed count equals the expected count, then the all the ratios of observed count over expected count in the equation for G^2 are equal to 1. The log of 1 is 0. Any observed count multiplied by 0 is 0. Thus, we are summing several 0's and multiplying them by 2, which, of course, gives 0.

b) In practice, we would not expect to get *exactly* $G^2 = X^2 = 0$, even if the variables are truly independent. A given random sample is not likely to have exactly the same breakdown as the population, because of sampling variability.

💻11.89 Voting with 16

a) Technology will confirm that $X^2 = 14.1$.

b) No, there are 6 cells with expected cell counts below 5. The chi-squared distribution may not approximate well the actual sampling distribution of X^2.

c) Answers will vary. One permutation yielded a value of 3.1, smaller than the observed 14.1.

d) Answers will vary. In one sample, none of the 10 permutations gave a value as large or larger than 14.1.

e) Answers may vary slightly. In one sample, 233 of the 10,000 permutations resulted in a chi-squared statistic as large or larger than the observed one, giving a P-value of 0.0233. This means that if there is no association between grade level and opinion, observing a test statistic as large or larger than the one we have observed is unlikely (probability of 0.02). With such a small P-value, we reject the null hypothesis and conclude that there is an association between the grade level and the opinion about voting with 16.

Chapter Problems: Student Activities

11.90 Conduct a research study using the GSS

Responses to this exercise will depend on the categorical response variable assigned by the instructor and on the explanatory variables chosen by each student.