Chapter 22 – Comparing Counts

Section 22.1

1. Human births.

a) If there were no "seasonal effect" we would expect 25% of births to occur in each season. The expected number of births is 0.25(120) = 30 births per season.

b) $\chi^2 = \frac{(25-30)^2}{30} + \frac{(35-30)^2}{30} + \frac{(32-30)^2}{30} + \frac{(28-30)^2}{30} \approx 1.933$

c) There are 4 seasons, so there are 4 – 1 = 3 degrees of freedom.

2. Bank cards.

a) If the historical percentages hold, we would expect 0.60(200) = 120, 0.30(200) = 60, and 0.10(200) = 20 customers to apply for Silver, Gold, and Platinum cards, respectively.

b) $\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(110-120)^2}{120} + \frac{(55-60)^2}{60} + \frac{(35-20)^2}{20} \approx 12.5$

c) There are 3 types of cards, so there are 3 – 1 = 2 degrees of freedom.

3. Human births, again.

a) The mean of the χ^2 distribution is the number of degrees of freedom, so we would expect the χ^2 statistic to be 3 if there were no seasonal effect.

b) Since 1.933 is less than the mean of 3, it does not seem large in comparison.

c) We should fail to reject the null hypothesis. These data do not provide evidence of a seasonal effect on human births.

d) The critical value of χ^2 with 3 degrees of freedom and $\alpha = 0.05$ is 7.815.

e) Since $\chi^2 = 1.933$ is less than the critical value, 7.815, we fail to reject the null hypothesis. There is no evidence of a seasonal effect on human births.

4. Bank cards, again.

a) The mean of the χ^2 distribution is the number of degrees of freedom, so we would expect the χ^2 statistic to be 2 if customers applied for bank cards according to the historical proportions.

b) The χ^2 statistic calculated in the Exercise 2, 12.5, in much greater than the mean of 2.

c) We should reject the null hypothesis. We are unlikely to see a value of χ^2 this high if customers applied for bank cards according to the historical proportions.

d) The critical value of χ^2 with 2 degrees of freedom and $\alpha = 0.05$ is 5.991.

e) Since $\chi^2 = 12.5$ is greater than the critical value, 5.991, we reject the null hypothesis. There is strong evidence that the customers are not applying for bank cards according to the historical proportions.

Section 22.2

5. Customer ages.

a) The null hypothesis is that the age distributions of the customers are the same at the two branches.

b) There are three age groups and one variable, location of the branch. This is a χ^2 test of homogeneity.

5. (continued)

c)

Expected Counts		Age		
	Less than 30	30-55	56 or older	Total
In-Town Branch	25	45	30	100
Mall Branch	25	45	30	100
Total	50	90	60	200

d) $$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(20-25)^2}{25} + \frac{(40-45)^2}{45} + \frac{(40-30)^2}{30} + \frac{(30-25)^2}{25} + \frac{(50-45)^2}{45} + \frac{(20-30)^2}{30}$$
$$\approx 9.778$$

e) There are 2 rows and 3 columns, so there are (2 – 1)(3 – 1) = 2 degrees of freedom.

f) The probability of having a χ^2 value over 9.778 with df = 2 is 0.0075.

g) Since the P-value is so low, reject the null hypothesis. There is strong evidence that the distribution of ages is not the same at the two branches. The mall branch had more customers under 30, and fewer customers 56 or older, than expected. The in-town branch had more customers 56 and older, and fewer customers under 30, than expected.

6. Bank cards, once more.

a) The null hypothesis is that the proportion of customers applying for the three types of cards is the same for each of the three mailings.

b) There are three mailings and one variable, the type of card. This is a χ^2 test of homogeneity.

c)

Expected Counts		Type of Card		
	Silver	Gold	Platinum	Total
Mailing 1	133.33	51.67	35	200
Mailing 2	133.33	51.67	35	200
Mailing 3	133.33	51.67	35	200
Total	340	155	105	600

d) $$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(120-113.33)^2}{113.33} + \frac{(50-51.67)^2}{51.67} + \frac{(30-35)^2}{35} + \frac{(115-113.33)^2}{113.33}$$
$$+ \frac{(50-51.67)^2}{51.67} + \frac{(35-35)^2}{35} + \frac{(105-113.33)^2}{113.33} + \frac{(55-51.67)^2}{51.67} + \frac{(40-35)^2}{35} \approx 2.7806$$

e) There are 3 rows and 3 columns, so there are (3 – 1)(3 – 1) = 4 degrees of freedom.

f) The probability of having a χ^2 value over 2.7806 with df = 4 is 0.5952.

g) Since the P-value is high, fail to reject the null hypothesis. There is no evidence that the proportions of card type differ for the three mailings.

Section 22.3

7. Human births, last time.

a) The standardized residuals are given below.

$$c_1 = \frac{(25-30)}{\sqrt{30}} \approx -0.913 \qquad c_2 = \frac{(35-30)}{\sqrt{30}} \approx 0.913 \qquad c_3 = \frac{(32-30)}{\sqrt{30}} \approx 0.365 \qquad c_4 = \frac{(28-30)}{\sqrt{30}} \approx -0.365$$

b) None of the standardized residuals are large. Since they are *z*-scores, they are actually quite small.

c) We did not reject the null hypothesis, so we should expect the standardized residuals to be relatively small.

8. Bank cards, last time.

a) The standardized residuals are given below.

$$c_1 = \frac{(110-120)}{\sqrt{120}} \approx -0.913 \qquad c_2 = \frac{(55-60)}{\sqrt{60}} \approx -0.645 \qquad c_3 = \frac{(35-20)}{\sqrt{20}} \approx 3.354$$

b) The standardized residual for the Platinum card is quite large. A *z*-score of 3.354 is large.

c) This large standardized residual suggests that the main difference in the proportion for this group is the increase in the number applying for the Platinum card.

Section 22.4

9. *Iliad* injuries 800 BCE

a) The null hypothesis is that the lethality of an injury is independent of the injury site.

Expected Counts		Lethal?		
		Fatal	Not fatal	Total
Injury Site	body	54.15	12.85	67
	head/neck	36.37	8.63	45
	limb	27.48	6.52	34
	Total	118	28	1467

b)
$$\chi^2 = \sum_{all\ cells} \frac{(Obs-Exp)^2}{Exp} = \frac{(61-54.15)^2}{54.15} + \frac{(6-12.85)^2}{12.85} + \frac{(44-36.37)^2}{36.37} + \frac{(1-8.63)^2}{8.63} + \frac{(13-27.48)^2}{27.48}$$
$$+ \frac{(21-6.52)^2}{6.52} \approx 52.65$$

c) There are 3 rows and 2 columns, so there are (3 – 1)(2 – 1) = 2 degrees of freedom.

d) The probability of having a χ^2 value over 52.65 with df = 2 is less than 0.0001.

e) Since the P-value is low, reject the null hypothesis. There is evidence of an association between the site of the injury and whether or not the injury was lethal. Injuries to the head/neck and body are more likely to be fatal, while injuries to the limbs are less likely to be fatal.

10. *Iliad* weapons.

a) The null hypothesis is that the injury site is independent of the type of weapon.

Expected Counts		Injury Site			
		Body	Head/Neck	Limb	Total
Weapon	Arrow	5.54483	3.72414	2.73103	12
	Ground/Rock	5.08276	3.41379	2.50345	11
	Sword	56.3724	37.8621	27.7655	122
	Total	67	45	33	145

10. (continued)

b) $$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(5-5.54483)^2}{5.54483} + \frac{(2-3.72414)^2}{3.72414} + \frac{(5-2.73103)^2}{2.73103} + \frac{(1-5.08276)^2}{5.08276}$$
$$+ \frac{(5-3.41379)^2}{3.41379} + \frac{(5-2.50345)^2}{2.50345} + \frac{(61-56.3724)^2}{56.3714} + \frac{(38-37.8621)^2}{37.8621} + \frac{(23-27.7655)^2}{27.7655} \approx 10.44$$

c) There are 3 rows and 3 columns, so there are (3 – 1)(3 – 1) = 4 degrees of freedom.

d) The probability of having a χ^2 value over 10.44 with df = 4 is 0.0336.

e) Since the P-value is low, reject the null hypothesis. There is evidence of an association between type of weapon and injury site. Injuries to the limb are more likely to be caused by an arrow or by the ground/rock, while injuries to the body are more likely to be caused by the ground/rock. However, we should be very cautious about making too much of these results, since 4 of the expected counts were less than 5. This data set is too small to make distinctions between injuries caused by arrow and ground/rock.

One way to work with these data is to combine arrows and ground/rock into one category, so that we only have 2 categories for weapon: sword and other. Now all expected counts are greater than 5.

	Observed (Expected)	Injury Site			
		Body	Head/Neck	Limb	Total
Weapon	Arrow/Ground/Rock	6 (10.62759)	7 (7.13793)	10 (5.23448)	12
Weapon	Sword	61 (56.3724)	38 (37.8621)	23 27.7655	122
	Total	67	45	33	145

There are now only 2 rows and 3 columns, so there are (2 – 1)(3 – 1) = 2 degrees of freedom. The value of χ^2 is now 7.5545.

The probability of having a χ^2 value over 7.5545 with df = 2 is 0.0229.

Since the P-value is low, reject the null hypothesis. There is evidence of an association between type of weapon and injury site. Injuries to the limb are more likely to be caused by an arrow or by the ground/rock, while injuries to the body are more likely to be caused by a sword.

Chapter Exercises.

11. Which test?

a) Chi-square test of Independence. We have one sample and two variables. We want to see if the variable *account type* is independent of the variable *trade type.*

b) Some other statistics test. The variable *account size* is quantitative, not counts.

c) Chi-square test of Homogeneity. We have two samples (residential and non-residential students), and one variable, *courses.* We want to see if the distribution of *courses* is the same for the two groups.

12. Which test, again?

a) Chi-square goodness-of-fit test. We want to see if the distribution of defects is uniform over the variable *day*.

b) Some other statistical test. *Cholesterol level* is a quantitative variable, not counts.

c) Chi-square test of Independence. We have data on two variables, *political leaning* and *major*, for one group of students.

13. Dice.

a) If the die were fair, you'd expect each face to show 10 times.

b) Use a chi-square test for goodness-of-fit. We are comparing the distribution of a single variable (outcome of a die roll) to an expected distribution.

c) H_0: The die is fair. (All faces have the same probability of coming up.)
H_A: The die is not fair. (Some faces are more or less likely to come up than others.)

d) **Counted data condition:** We are counting the number of times each face comes up.
Randomization condition: Die rolls are random and independent of each other.
Expected cell frequency condition: We expect each face to come up 10 times, and 10 is greater than 5.

e) Under these conditions, the sampling distribution of the test statistic is χ^2 on 6 – 1 = 5 degrees of freedom. We will use a chi-square goodness-of-fit test.

f)

Face	Observed	Expected	Residual $(Obs - Exp)$	Residual² $(Obs - Exp)^2$	Component $(Obs - Exp)^2 / Exp$
1	11	10	1	1	0.1
2	7	10	–3	9	0.9
3	9	10	–1	1	0.1
4	15	10	5	25	2.5
5	12	10	2	4	0.4
6	6	10	–4	16	1.6

g) $$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(11-10)^2}{10} + \frac{(7-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(12-10)^2}{10} + \frac{(6-10)^2}{10} = 5.6;\ \text{df} = 5$$

Since the P-value ≈ 0.3471 is high, we fail to reject the null hypothesis. There is no evidence that the die is unfair.

14. M&M's

a) There are 29 + 23 + 12 + 14 + 8 + 20 = 106 M&M's in the bag. The expected number of M&M's of each color is: 106(0.20) = 21.2 red, 106(0.20) = 21.2 yellow, 106(0.10) = 10.6 orange, 106(0.10) = 10.6 blue, 106(0.10) = 10.6 green, and 106(0.30) = 31.8 brown.

b) Use a chi-square test for goodness-of-fit. We are comparing the distribution of a single variable (color) to an expected distribution.

c) H_0: The distribution of colors of M&M's is as specified by the company.
H_A: The distribution of colors of M&M's is different than specified by the company.

d) **Counted data condition:** The author counted the M&M's in the bag.
Randomization condition: These M&M's are mixed thoroughly at the factory.
Expected cell frequency condition: The expected counts (calculated in part (b)) are all greater than 5.

e) Since there are 6 different colors, there are 6 – 1 = 5 degrees of freedom.

14. (continued)

f) Under these conditions, the sampling distribution of the test statistic is χ^2 on 6 – 1 = 5 degrees of freedom. We will use a chi-square goodness-of-fit test.

Color	Observed	Expected	Residual $(Obs - Exp)$	Residual² $(Obs - Exp)^2$	Component $(Obs - Exp)^2 / Exp$
yellow	29	21.2	7.8	60.84	2.8698
red	23	21.2	1.8	3.24	0.1528
orange	12	10.6	1.4	1.96	0.1849
blue	14	10.6	3.4	11.56	1.0906
green	8	10.6	–2.6	6.76	0.6377
brown	20	31.8	–11.8	139.24	4.3786

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(29-21.2)^2}{21.2} + \frac{(23-21.2)^2}{21.2} + \frac{(12-10.6)^2}{10.6} + \frac{(14-10.6)^2}{10.6} + \frac{(8-10.6)^2}{10.6}$$

$$+ \frac{(20-31.8)^2}{31.8} \approx 9.3145.6;\ \text{df} = 5$$

Since the P-value ≈ 0.0972 is high, we fail to reject the null hypothesis.

g) There is no evidence that the distribution of colors is anything other than the distribution specified by the company.

15. Nuts.

a) The weights of the nuts are quantitative. Chi-square goodness-of-fit requires counts.

b) In order to use a chi-square test, you could count the number of each type of nut. However, it's not clear whether the company's claim was a percentage by number or a percentage by weight.

16. Mileage.

The average number of miles traveled is quantitative date, not categorical. Chi-square is for comparing counts.

17. NYPD and race.

H_0: The distribution of ethnicities in the police department represents the distribution of ethnicities of the youth of New York City.
H_A: The distribution of ethnicities in the police department does not represent the distribution of ethnicities of the youth of New York City.

Counted data condition: The percentages reported must be converted to counts.
Randomization condition: Assume that the current NYPD is representative of recent departments with respect to ethnicity.
Expected cell frequency condition: The expected counts are all much greater than 5.

(Note: The observed counts should be whole numbers. They are actual policemen. The expected counts may be decimals, since they behave like averages.)

Ethnicity	Observed	Expected
White	16,965	7644.852
Black	3796	7383.042
Latino	5001	8247.015
Asian	367	2382.471
Other	52	523.620

17. (continued)

Under these conditions, the sampling distribution of the test statistic is χ^2 on 5 – 1 = 4 degrees of freedom. We will use a chi-square goodness-of-fit test.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(16{,}965 - 7644.852)^2}{7644.852} + \frac{(3796 - 7383.042)^2}{7383.042} + \frac{(5001 - 8247.015)^2}{8247.015}$$
$$+ \frac{(367 - 2382.471)^2}{2382.471} + \frac{(52 - 523.620)^2}{523.620} \approx 16{,}512.7$$

With χ^2 of over 16,500, on 4 degrees of freedom, the P-value is essentially 0, so we reject the null hypothesis. There is strong evidence that the distribution of ethnicities of NYPD officers does not represent the distribution of ethnicities of the youth of New York City. Specifically, the proportion of white officers is much higher than the proportion of white youth in the community. As one might expect, there are also lower proportions of officers who are black, Latino, Asian, and other ethnicities than we see in the youth in the community.

18. Violence against women.

H_0: The weapon use rates in murders of women have the same distribution as the weapon use rates of all murders.
H_A: The weapon use rates in murders of women have a different distribution than the weapon use rates of all murders.

Counted data condition: The percentages reported must be converted to counts.
Randomization condition: Assume that the weapon use rates from 2009 are representative of the weapon use rates for all recent years.
Expected cell frequency condition: The expected counts are all much greater than 5.

Weapon	Observed	Expected
guns	861	1048.636
knives	364	216.674
other	214	277.872
personal	215	110.818

(Note: The observed counts should be whole numbers. They are actual murders. The expected counts may be decimals, since they behave like averages.)

Under these conditions, the sampling distribution of the test statistic is χ^2 on 4 – 1 = 3 degrees of freedom. We will use a chi-square goodness-of-fit test.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(861 - 1048.636)^2}{1048.636} + \frac{(364 - 216.674)^2}{216.674} + \frac{(214 - 277.872)^2}{277.872} + \frac{(215 - 110.818)^2}{110.818}$$
$$\approx 246.4;\ df = 3$$

The P-value is essentially 0, so we reject the null hypothesis. There is strong evidence that the distribution of weapon use rates is different for murders of women than for all murders. Women are much more likely to be killed by personal attacks and knives, and less likely to be killed with guns or other weapons.

19. Fruit flies.

a) H_0: The ratio of traits in this type of fruit fly is 9:3:3:1, as genetic theory predicts.
H_A: The ratio of traits in this type of fruit fly is not 9:3:3:1.

Counted data condition: The data are counts.
Randomization condition: Assume that these flies are representative of all fruit flies of this type.
Expected cell frequency condition: The expected counts are all greater than 5.

Trait	Observed	Expected
YN	59	56.25
YS	20	18.75
EN	11	18.75
ES	10	6.25

Under these conditions, the sampling distribution of the test statistic is χ^2 on 4 – 1 = 3 degrees of freedom. We will use a chi-square goodness-of-fit test.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(59-56.25)^2}{56.25} + \frac{(20-18.75)^2}{18.75} + \frac{(11-18.75)^2}{18.75} + \frac{(10-6.25)^2}{6.25} \approx 5.671;\ \text{df} = 3$$

The P-value ≈ 0.1288 is high, so we fail to reject the null hypothesis. There is no evidence that the ratio of traits is different than the theoretical ratio predicted by the genetic model. The observed results are consistent with the genetic model.

b) With $\chi^2 \approx 11.342$, on 3 degrees of freedom, the P-value = 0.0100 is low, so we reject the null hypothesis. There is strong evidence that the ratio of traits is different than the theoretical ratio predicted by the genetic model. Specifically, there is evidence that the normal wing length occurs less frequently than expected and the short wing length occurs more frequently than expected.

Trait	Observed	Expected
YN	118	112.5
YS	40	37.5
EN	22	37.5
ES	20	12.5

c) At first, this seems like a contradiction. We have two samples with exactly the same ratio of traits. The smaller of the two provides no evidence of a difference, yet the larger one provides strong evidence of a difference. This is explained by the sample size. In general, large samples decrease the proportion of variation from the true ratio. Because of the relatively small sample in the first test, we are unwilling to say that there is a difference. There just isn't enough evidence. But the larger sample allows us to be more certain about the difference.

20. Pi.

H_0: Digits of π are uniformly distributed (all occur with frequency 1/10).
H_A: Digits of π are not uniformly distributed.

Counted data condition: The data are counts.
Randomization condition: Assume that the first million digits of π are representative of all digits.
Expected cell frequency condition: The expected counts are all greater than 5.

20. (continued)

Digit	Observed	Expected
0	99,959	100,000
1	99,758	100,000
2	100,026	100,000
3	100,229	100,000
4	100,230	100,000
5	100,359	100,000
6	99,548	100,000
7	99,800	100,000
8	99,985	100,000
9	100,106	100,000

Under these conditions, the sampling distribution of the test statistic is χ^2 on 10 – 1 = 9 degrees of freedom. We will use a chi-square goodness-of-fit test.

With $\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 5.509$, on 9 degrees of freedom, the P-value = 0.7879 is high, so we fail to reject the null hypothesis. There is no evidence that the digits of π are not uniformly distributed. These data are consistent with the null hypothesis.

21. Hurricane frequencies.

a) We would expect 96/16 = 6 hurricanes per time period.

b) We are comparing the distribution of the number of hurricanes, a single variable, to a theoretical distribution. A Chi-square test for goodness-of-fit is appropriate.

c) H_0: The number of large hurricanes remains constant over decades.
H_A: The number of large hurricanes has changed.

d) There are 16 time periods, so there are 16 – 1 = 15 degrees of freedom.

e) $P(\chi^2_{df=15} > 12.67) \approx 0.63$

f) The very high P-value means that these data offer no evidence that the number of hurricanes large hurricanes has changed.

22. Lottery numbers.

a) We are comparing the distribution of the number of times each lottery number has occurred, either as a regular number, or as the Powerball. We will use a Chi-square test for goodness-of-fit.

b) We expect the Powerballs to be distributed uniformly over the 49 numbers. We expect each number to be the bonus ball 655/49 = 13.367 times.

c) H_0: All numbers are equally likely to be the bonus ball.
H_A: Some numbers are more likely than others to be the bonus ball.

d) There are 49 numbers, so there are 49 – 1 = 48 degrees of freedom.

e) $P(\chi^2_{df=48} > 34.5) \approx 0.93$

f) The very high P-value means that these data offer no evidence that some lottery numbers are more likely than others to be the bonus ball.

23. Childbirth, part 1.

a) There are two variables, breastfeeding and having an epidural, from a single group of births. We will perform a Chi-square test for Independence.

b) H_0: Breastfeeding success is independent of having an epidural.
H_A: There is an association between breastfeeding success and having an epidural.

24. Does your doctor know?

a) There is one variable, whether or not statistics is used, over three time periods in which the articles were published. We will perform a Chi Square test of Homogeneity.

b) H_0: The same proportion of articles used statistics in the three time periods.
H_A: The proportion of articles that used statistics was different in the three time periods.

25. Childbirth, part 2.

a) The table has 2 rows and 2 columns, so there are $(2-1)\times(2-1)=1$ degree of freedom.

b) We expect $\frac{474}{1178} \approx 40.2\%$ of all babies to not be breastfeeding after 6 months, so we expect that 40.2% of the 396 epidural babies, or 159.34, to not be breastfeeding after 6 months.

c) Breastfeeding behavior should be independent for these babies. They are fewer than 10% of all babies, and we assume they are representative of all babies. We have counts, and all the expected cells are at least 5.

26. Does your doctor know? (part 2).

a) The table has 2 rows and 3 columns, so there are $(2-1)\times(3-1)=2$ degrees of freedom.

b) We expect $\frac{144}{758} \approx 19\%$ of all articles to contain no statistics, so we expect 19% of the 115 articles from 1989, or 21.85, to contain no statistics.

c) These are counted data. One article shouldn't affect another article (except perhaps for a follow-up article to another article included in the study). We can regard the selected years as representative of other years, and the authors seem to want to regard these articles as representative of those appearing in similar-quality medical journals, so they are fewer than 10% of all articles. All expected counts are greater than 5.

27. Childbirth, part 3.

a) $\frac{(Obs-Exp)^2}{Exp} = \frac{(190-159.34)^2}{159.34} = 5.90$

b) $P(\chi^2_{df=1} > 14.87) < 0.005$

c) The P-value is very low, so reject the null hypothesis. There's strong evidence of an association between having an epidural and subsequent success in breastfeeding.

28. Does your doctor know? (part 3).

a) $\frac{(Obs-Exp)^2}{Exp} = \frac{(14-21.85)^2}{21.85} = 2.82$

b) $P(\chi^2_{df=2} > 25.28) < 0.001$

c) The P-value is very low, so reject the null hypothesis. There's strong evidence that the proportion of medical journal articles that contain statistics is different for the three time periods.

29. Childbirth, part 4.

a) $c = \dfrac{Obs - Exp}{\sqrt{Exp}} = \dfrac{190 - 159.34}{\sqrt{159.34}} = 2.43$

b) It appears that babies whose mothers had epidurals during childbirth are much more likely to be breastfeeding 6 months later.

30. Does your doctor know? (part 4).

a) $c = \dfrac{Obs - Exp}{\sqrt{Exp}} = \dfrac{14 - 21.85}{\sqrt{21.85}} = -1.68$

b) The residuals for No stats are decreasing and those for Stats are increasing over time, indicating that, over time, a smaller proportion of articles are appearing without statistics.

31. Childbirth, part 5.

These factors would not have been mutually exclusive. There would be yes or no responses for every baby for each.

32. Does your doctor know? (part 5).

These methods would not have been mutually exclusive. Articles might use more than one statistical method.

33. *Titanic*.

a) $P(\text{crew}) = \dfrac{889}{2208} \approx 0.4026$

b) $P(\text{third and alive}) = \dfrac{180}{2208} \approx 0.0815$

c) $P(\text{alive} \mid \text{first}) = \dfrac{P(\text{alive and first})}{\text{P(first)}} = \dfrac{201/2208}{324/2208} = \dfrac{201}{324} \approx 0.6204$

d) The overall chance of survival is $\dfrac{712}{2208} \approx 0.3225$**,** so we would expect about 32.3% of the crew, or about 286.67 members of the crew, to survive.

e) H_0: Survival was independent of status on the ship.
H_A: Survival depended on status on the ship.

f) The table has 2 rows and 4 columns, so there are $(2-1)\times(4-1) = 3$ degrees of freedom.

g) With $\chi^2 \approx 187.56$, on 3 degrees of freedom, the P-value is essentially 0, so we reject the null hypothesis. There is strong evidence of an association between survival and status. First-class passengers were more likely to survive than any other class or crew.

34. NYPD.

a) $P(\text{female}) = \dfrac{5613}{37{,}379} \approx 0.150$

b) $P(\text{detective}) = \dfrac{4864}{37{,}379} \approx 0.130$

c) The overall percentage of females is 15%**,** so we would expect about 15% of the detectives, or about 730 detectives, to be female.

d) We have one group, categorized according to two variables, rank and gender, so we will perform a chi-square test for independence.

e) H_0: Rank is independent of gender in the NYPD.
H_A: Rank is associated with gender in the NYPD.

34. (continued)

f) **Counted data condition:** The data are counts.
Randomization condition: These data are not a random sample, but all NYPD officers. Assume that these officers are representative with respect to the recent distribution of sex and rank in the NYPD.
Expected cell frequency condition: The expected counts are all greater than 5.

Rank	Male	Female
Officer	22249.5	3931.5
Detective	4133.6	730.4
Sergeant	3665.3	647.7
Lieutenant	1208.5	213.5
Captain	315.3	55.7
Higher ranks	193.8	34.2

g) The table has 6 rows and 2 columns, so there are $(6-1)\times(2-1)=5$ degrees of freedom.

h) With $\chi^2 = 343.9$, the P-value is very low. We reject the null hypothesis. There is strong evidence of an association between the sex and rank of NYPD officers.

35. *Titanic*, again.

First class passengers were most likely to survive, while third class passengers and crew were under-represented among the survivors.

36. NYPD again.

Women are over-represented at the lower ranks and under-represented at every rank from sergeant up.

37. Cranberry juice.

a) This is an experiment. Volunteers were assigned to drink a different beverage.

b) We are concerned with the proportion of urinary tract infections among three different groups. We will use a chi-square test for homogeneity.

c) H_0: The proportion of urinary tract infection is the same for each group.
H_A: The proportion of urinary tract infection is different among the groups.

d) **Counted data condition:** The data are counts.
Randomization condition: Although not specifically stated, we will assume that the women were randomly assigned to treatments.
Expected cell frequency condition: The expected counts are all greater than 5.

	Cranberry (*Obs / Exp*)	Lactobacillus (*Obs / Exp*)	Control (*Obs / Exp*)
Infection	8 / 15.333	20 / 15.333	18 / 15.333
No infection	42 / 34.667	30 / 34.667	32 / 34.667

e) The table has 2 rows and 3 columns, so there are $(2-1)\times(3-1)=2$ degrees of freedom.

f) $$\chi^2 = \sum_{all\ cells} \frac{(Obs-Exp)^2}{Exp} = \frac{(8-15.333)^2}{15.333} + \frac{(20-15.333)^2}{15.333} + \frac{(18-15.333)^2}{15.333} + \frac{(42-34.667)^2}{34.667}$$
$$+\frac{(30-34.667)^2}{34.667} + \frac{(32-34.667)^2}{34.667} \approx 7.776,\ \text{df} = 2;\ \text{P-value} \approx 0.020$$

37. (continued)

g) Since the P-value is low, we reject the null hypothesis. There is strong evidence of difference in the proportion of urinary tract infections for cranberry juice drinkers, lactobacillus drinkers, and women that drink neither of the two beverages.

h) A table of the standardized residuals is below, calculated by using $c = \frac{Obs - Exp}{\sqrt{Exp}}$.

	Cranberry	Lactobacillus	Control
Infection	–1.87276	1.191759	0.681005
No infection	1.245505	–0.79259	–0.45291

There is evidence that women who drink cranberry juice are less likely to develop urinary tract infections, and women who drank lactobacillus are more likely to develop urinary tract infections.

38. Car origins.

a) We have two groups, staff and students (selected from different lots), and we are concerned with the distribution of one variable, origin of car. We will perform a chi-square test for homogeneity.

b) H_0: The distribution of car origin is the same for students and staff.
H_A: The distribution of car origin is different for students and staff.

c) **Counted data condition:** The data are counts.
Randomization condition: Cars were surveyed randomly.
Expected cell frequency condition: The expected counts are all greater than 5.

	Driver	
Origin	**Student (*Obs / Exp*)**	**Staff (*Obs / Exp*)**
American	107 / 115.15	105 / 96.847
European	33 / 24.443	12 / 20.557
Asian	55 / 55.404	47 / 46.596

d) Under these conditions, the sampling distribution of the test statistic is χ^2 on 2 degrees of freedom. We will use a chi-square test for homogeneity.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 7.828;\ \text{df} = 2;\ \text{P-value} \approx 0.020$$

e) Since the P-value is low, we reject the null hypothesis. There is strong evidence that the distribution of car origins at this university differs between students and staff. Students are more likely to drive European cars than staff and less likely than staff to drive American cars.

39. Montana.

a) We have one group, categorized according to two variables, political party and being male or female, so we will perform a chi-square test for independence.

b) H_0: Political party is independent of being male or female in Montana.
H_A: There is an association between political party and being male or female in Montana.

c) **Counted data condition:** The data are counts.
Randomization condition: Although not specifically stated, we will assume that the poll was conducted randomly.
Expected cell frequency condition: The expected counts are all greater than 5.

39. (continued)

	Democrat (*Obs / Exp*)	Republican (*Obs / Exp*)	Independent (*Obs / Exp*)
Male	36 / 43.663	45 / 40.545	24 / 20.792
Female	48 / 40.337	33 / 37.455	16 / 19.208

d) Under these conditions, the sampling distribution of the test statistic is χ^2 on 2 degrees of freedom. We will use a chi-square test for independence.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} = \frac{(36-46.663)^2}{43.663} + \frac{(45-40.545)^2}{40.545} + \frac{(24-20.792)^2}{20.792} + \frac{(48-40.337)^2}{40.337}$$

$$+\frac{(33-37.455)^2}{37.455} + \frac{(16-19.208)^2}{19.208} \approx 4.851,\ \text{df} = 2;\ \text{P-value} \approx 0.0884$$

e) Since the P-value is fairly high, we fail to reject the null hypothesis. There is little evidence of an association between being male or female and political party in Montana.

40. Fish diet.

a) This is an observational prospective study. Swedish men were selected, and then followed for 30 years.

b) We have one group, categorized according to two variables, fish consumption and incidence of prostate cancer, so we will perform a chi-square test for independence.

c) H_0: Prostate cancer and fish consumption are independent.
H_A: There is an association between prostate cancer and fish consumption.

Counted data condition: The data are counts.
Randomization condition: Assume that these men are representative of all men.
Expected cell frequency condition: The expected counts are all greater than 5.

Fish Consumption	Prostate Cancer (*Obs / Exp*)	No Prostate Cancer (*Obs / Exp*)
Never/Seldom	14 / 9.21	110 / 114.79
Small part	201 /194.74	2420 / 2426.3
Moderate part	209 / 221.26	2769 / 2756.7
Large part	42 / 40.79	507 / 508.21

Under these conditions, the sampling distribution of the test statistic is χ^2 on 3 degrees of freedom. We will use a chi-square test for independence.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 3.677,\ \text{df} = 3;\ \text{P-value} \approx 0.2985$$

Since the P-value is high, we fail to reject the null hypothesis. There is no evidence of an association between prostate cancer and fish consumption.

d) This does not prove that eating fish does not prevent prostate cancer. There is merely a lack of evidence of a relationship. Furthermore, association (or lack thereof) does not prove a cause-and-effect relationship. We would need to conduct a controlled experiment before anything could be proven.

41. Montana revisited.

H_0: Political party is independent of region in Montana.
H_A: There is an association between political party and region in Montana.

Counted data condition: The data are counts.
Randomization condition: Although not specifically stated, we will assume that the poll was conducted randomly.
Expected cell frequency condition: All expected counts are greater than 5.

	Democrat (*Obs / Exp*)	Republican (*Obs / Exp*)	Independent (*Obs / Exp*)
West	39 / 28.277	17 / 26.257	12 / 13.465
Northeast	15 / 23.703	30 / 22.01	12 / 11.287
Southeast	30 /32.02	31 / 29.733	16 / 15.248

Under these conditions, the sampling distribution of the test statistic is χ^2 on 4 degrees of freedom. We will use a chi-square test for independence.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 13.849,\ \text{df} = 4; \text{and the P-value} \approx 0.0078$$

Since the P-value is low, reject the null hypothesis. There is strong evidence of an association between region and political party in Montana. Residents in the West are more likely to be Democrats than Republicans, and residents in the Northeast are more likely to be Republicans than Democrats.

42. Working parents.

a) This is a survey of adults.

b) We have two groups, males and females, and we are concerned with the distribution of one variable, attitude about the child care options. We will perform a chi-square test for homogeneity.

c) H_0: The distribution of attitudes about child care is the same for men and women.
H_A: The distribution of attitudes about child care is not the same for men and women.

Counted data condition: The data are counts.
Randomization condition: Adults were surveyed randomly.
Expected cell frequency condition: The expected counts are all greater than 5.

	Male (*Obs / Exp*)	Female (*Obs / Exp*)
Both work full time	161 / 150.5	140/ 150.5
One works full time, other part time	259/ 283.5	308 / 283.5
One works, other stays at home	189 / 175	161/ 175
Both parents work part time	49 / 56	63 / 56
No opinion	42 / 35	28 / 35

Under these conditions, the sampling distribution of the test statistic is χ^2 on 4 degrees of freedom. We will use a chi-square test for homogeneity.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 12.49,\ \text{df} = 4;\ \text{P-value} \approx 0.0141$$

Since the P-value is low, we reject the null hypothesis. There is strong evidence of a difference in the distribution of attitudes about child care options between men and women. Men are more likely than woman to prefer a situation where both parents work full time, and women are more likely to prefer a situation where one parent works full time and the other works part time.

43. Grades.

a) We have two groups, students of Professor Alpha and students of Professor Beta, and we are concerned with the distribution of one variable, grade. We will perform a chi-square test for homogeneity.

b) H_0: The distribution of grades is the same for the two professors.
H_A: The distribution of grades is different for the two professors.

c) The expected counts are organized in the table below:

	Prof. Alpha	Prof. Beta
A	6.667	5.333
B	12.778	10.222
C	12.222	9.778
D	6.111	4.889
F	2.222	1.778

Since three cells have expected counts less than 5, the chi-square procedures are not appropriate. Cells would have to be combined in order to proceed. (We will do this in another exercise.)

44. Full moon.

a) We have two groups, weeks of six full moons and six other weeks, and we are concerned with the distribution of one variable, type of offense. We will perform a chi-square test for homogeneity.

b) H_0: The distribution of type of offense is the same for full moon weeks as it is for weeks in which there is not a full moon.
H_A: The distribution of type of offense is different for full moon weeks than it is for weeks in which there is not a full moon.

c) The expected counts are organized in the table below:

Offense	Full Moon	Not Full
Violent	2.558	2.442
Property	19.442	18.558
Drugs / Alcohol	23.535	22.465
Domestic Abuse	12.791	12.209
Other offenses	7.674	7.326

Since two cells have expected counts less than 5, the chi-square procedures are not appropriate. Cells would have to be combined in order to proceed. (We will do this in another exercise.)

45. Grades again.

a) **Counted data condition:** The data are counts.
Randomization condition: Assume that these students are representative of all students that have ever taken courses from the professors.
Expected cell frequency condition: The expected counts are all greater than 5.

	Prof. Alpha (*Obs / Exp*)	Prof. Beta (*Obs / Exp*)
A	3 / 6.667	9 / 5.333
B	11 / 12.778	12 / 10.222
C	14 / 12.222	8 / 9.778
Below C	12 / 8.333	3 / 6.667

45. (continued)

b) Under these conditions, the sampling distribution of the test statistic is χ^2 on 3 degrees of freedom, instead of 4 degrees of freedom before the change in the table. We will use a chi-square test for homogeneity.

c) $$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 9.306,\ \text{df} = 3;\ \text{P-value} \approx 0.0255$$

Since the P-value is low, we reject the null hypothesis. There is evidence that the grade distributions for the two professors are different. Professor Alpha gives fewer As and more grades below C than Professor Beta.

46. Full moon, next phase.

a) **Counted data condition:** The data are counts.
Randomization condition: Assume that these weeks are representative of all weeks.
Expected cell frequency condition:
It seems reasonable to combine the violent offenses and domestic abuse, since both involve some sort of violence. Combining violent crimes with the "other offenses" is okay, but that may put very minor offenses in with violent offenses, which doesn't seem right. Once the cells are combined, all expected counts are greater than 5.

Offense	**Full Moon (*Obs / Exp*)**	**Not Full (*Obs / Exp*)**
Violent / Domestic Abuse	13 / 15.349	17 / 14.651
Property	17 / 19.442	21 / 18.558
Drugs / Alcohol	27 / 23.535	19 / 22.465
Other offenses	9 / 7.674	6 / 7.326

b) Under these conditions, the sampling distribution of the test statistic is χ^2 on 3 degrees of freedom, instead of 4 degrees of freedom before the change in the table. We will use a chi-square test for homogeneity.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 2.877,\ \text{df} = 3;\ \text{P-value} \approx 0.4109$$

Since the P-value is high, we fail to reject the null hypothesis. There is no evidence that the distribution of offenses is different during the full moon than during other phases.

47. Racial steering.

H_0: There is no association between race and section of the complex in which people live.
H_A: There is an association between race and section of the complex in which people live.

Counted data condition: The data are counts.
Randomization condition: Assume that the recently rented apartments are representative of all apartments in the complex.
Expected cell frequency condition: The expected counts are all greater than 5.

	White (***Obs / Exp***)	**Black** (***Obs / Exp***)
Section A	87 / 76.179	8 / 18.821
Section B	83 / 93.821	34 / 23.179

Under these conditions, the sampling distribution of the test statistic is χ^2 on 1 degree of freedom. We will use a chi-square test for independence.

$$\chi^2 = \sum_{\text{all cells}} \frac{(Obs - Exp)^2}{Exp} = \frac{(87-79.179)^2}{79.179} + \frac{(8-18.821)^2}{18.821} + \frac{(83-93.821)^2}{93.821} + \frac{(34-23.179)^2}{23.179} \approx 14.058,\ \text{df} = 1$$

Since the P-value ≈ 0.0002 is low, we reject the null hypothesis. There is strong evidence of an association between race and the section of the apartment complex in which people live. An examination of the components shows us that whites are much more likely to rent in Section A (component = 6.2215), and blacks are much more likely to rent in Section B (component = 5.0517).

48. Survival on the *Titanic*.

H_0: Survival was independent of gender on the *Titanic*.
H_A: There is an association between survival and gender on the *Titanic*.

Counted data condition: The data are counts.
Randomization condition: We have the entire population of the *Titanic*.
Expected cell frequency condition: The expected counts are all greater than 5.

	Female (***Obs / Exp***)	**Male** (***Obs / Exp***)
Alive	359 / 157.685	353 / 554.315
Dead	130 / 331.315	1366 / 1164.685

Under these conditions, the sampling distribution of the test statistic is χ^2 on 1 degree of freedom. We will use a chi-square test for independence.

$$\chi^2 = \sum_{\text{all cells}} \frac{(Obs - Exp)^2}{Exp} = \frac{(359-157.685)^2}{157.685} + \frac{(353-554.315)^2}{554.315} + \frac{(130-331.315)^2}{331.315} + \frac{(1366-1164.685)^2}{1164.685}$$
$$\approx 487.25,\ \text{df} = 1$$

Since the P-value is so low (essentially 0), we reject the null hypothesis. There is strong evidence of an association between survival and gender on the *Titanic*. Females were much more likely to survive than males.

49. Pregnancies.

H_0: Pregnancy outcome is independent of age.
H_A: There is an association between pregnancy outcome and age.

Counted data condition: The data are counts.
Randomization condition: Assume that these women are representative of all pregnant women.
Expected cell frequency condition: The expected counts are all greater than 5.

	Early Preterm (*Obs / Exp*)	Late Preterm (*Obs / Exp*)
Under 20	129 / 116.55	270 / 282.45
20 – 29	243 / 249.75	612 / 605.25
30 – 39	165 / 172.05	424 / 416.95
40 and over	18 / 16.65	39 / 40.35

Under these conditions, the sampling distribution of the test statistic is χ^2 on 3 degrees of freedom. We will use a chi-square test for independence.

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 2.699,\ \text{df} = 3;\ \text{P-value} \approx 0.440$$

Since the P-value is high, we fail to reject the null hypothesis. There is no evidence of an association between pregnancy outcome and age.

50. Education by age.

H_0: The distribution of education level attained is the same for different age groups.
H_A: The distribution of education level attained is different for different age groups.

Counted data condition: The data are counts.
Randomization condition: Assume that the sample was taken randomly.
Expected cell frequency condition: The expected counts are all greater than 5.

	25 – 34 ((*Obs / Exp*)	35 – 44 (*Obs / Exp*)	45 – 54 (*Obs / Exp*)	55 – 64 (*Obs / Exp*)	65 + (*Obs / Exp*)
Not HS Grad	27 / 60.2	50 / 60.2	52 / 60.2	71 / 60.2	101 / 60.2
HS	82 / 66.2	19 / 66.2	88 / 66.2	83 / 66.2	59 / 66.2
1 – 3 years college	43 / 33	56 / 33	26 / 33	20 / 33	20 / 33
4+ years college	48 / 40.6	75 / 40.6	34 / 40.6	26 / 40.6	20 / 40.6

Under these conditions, the sampling distribution of the test statistic is χ^2 on 12 degrees of freedom. We will use a chi-square test for homogeneity. (There are 200 people in each age group, an indication that we are examining 5 age groups, with respect to one variable, education level attained.)

$$\chi^2 = \sum_{all\ cells} \frac{(Obs - Exp)^2}{Exp} \approx 178.453,\ \text{df} = 12$$

Since the P-value is so low (essentially 0), we reject the null hypothesis. There is strong evidence that the distribution of education level attained is different between the groups. Generally, younger people are more likely to have higher levels of education than older people, who are themselves over represented at the lower education levels. Specifically, people in the 35 – 44 age group were less likely to have only a high school diploma, and more likely to have at least four years of college.