11.10: The Chi-Square Distribution (Exercises)

These are homework exercises to accompany the Textmap created for "Introductory Statistics" by OpenStax.

11.1: Introduction

11.2: Facts about the Chi-Square Distribution

Decide whether the following statements are true or false.

Q 11.2.1

As the number of degrees of freedom increases, the graph of the chi-square distribution looks more and more symmetrical.

S 11.2.1

true

Q 11.2.2

The standard deviation of the chi-square distribution is twice the mean.

Q 11.2.3

The mean and the median of the chi-square distribution are the same if $df = 24$.

S 11.2.3

false

11:3: Goodness-of-Fit Test

For each problem, use a solution sheet to solve the hypothesis test problem. Go to [link] for the chi-square solution sheet. Round expected frequency to two decimal places.

Q 11.3.1

A six-sided die is rolled 120 times. Fill in the expected frequency column. Then, conduct a hypothesis test to determine if the die is fair. The data in Table are the result of the 120 rolls.

Face Value	Frequency	Expected Frequency
1	15	
2	29	
3	16	
4	15	
5	30	
6	15	

The marital status distribution of the U.S. male population, ages 15 and older, is as shown in Table.Q 11.3.2

Marital Status	Percent	Expected Frequency
never married	31.3	
married	56.1	
widowed	2.5	
divorced/separated	10.1	

Suppose that a random sample of 400 U.S. young adult males, 18 to 24 years old, yielded the following frequency distribution. We are interested in whether this age group of males fits the distribution of the U.S. adult population. Calculate the frequency one would expect when surveying 400 people. Fill in Table, rounding to two decimal places.

Marital Status	Frequency
never married	140
married	238
widowed	2
divorced/separated	20

S 11.3.2

Marital Status	Percent	Expected Frequency
never married	31.3	125.2
married	56.1	224.4
widowed	2.5	10
divorced/separated	10.1	40.4

a. The data fits the distribution.
b. The data does not fit the distribution.
c. 3
d. chi-square distribution with $df = 3$
e. 19.27
f. 0.0002
g. Check student's solution.
h. i. $\alpha = 0.05$
   ii. Decision: Reject null
   iii. Reason for decision: p-value $< \alpha$
   iv. Conclusion: Data does not fit the distribution.

Use the following information to answer the next two exercises: The columns in Table contain the Race/Ethnicity of U.S. Public Schools for a recent year, the percentages for the Advanced Placement Examinee Population for that class, and the Overall Student Population. Suppose the right column contains the result of a survey of 1,000 local students from that year who took an AP Exam.

Race/Ethnicity	AP Examinee Population	Overall Student Population	Survey Frequency
Asian, Asian American, or Pacific Islander	10.2%	5.4%	113
Black or African-American	8.2%	14.5%	94
Hispanic or Latino	15.5%	15.9%	136
American Indian or Alaska Native	0.6%	1.2%	10
White	59.4%	61.6%	604
Not reported/other	6.1%	1.4%	43

Q 11.3.3

Perform a goodness-of-fit test to determine whether the local results follow the distribution of the U.S. overall student population based on ethnicity.

Q 11.3.4

Perform a goodness-of-fit test to determine whether the local results follow the distribution of U.S. AP examinee population, based on ethnicity.

S 11.3.4

a. H_0: The local results follow the distribution of the U.S. AP examinee population
b. H_0: The local results do not follow the distribution of the U.S. AP examinee population
c. $df = 5$
d. chi-square distribution with $df = 5$
e. chi-square test statistic = 13.4
f. $p\text{-value} = 0.0199$
g. Check student's solution.
h. i. $\alpha = 0.05$

ii. Decision: Reject null when $a = 0.05$

iii. Reason for Decision: $p\text{-value} < \alpha$

iv. Conclusion: Local data do not fit the AP Examinee Distribution.

v. Decision: Do not reject null when $a = 0.01$

vi. Conclusion: There is insufficient evidence to conclude that local data do not follow the distribution of the U.S. AP examinee distribution.

Q 11.3.5

The City of South Lake Tahoe, CA, has an Asian population of 1,419 people, out of a total population of 23,609. Suppose that a survey of 1,419 self-reported Asians in the Manhattan, NY, area yielded the data in Table. Conduct a goodness-of-fit test to determine if the self-reported sub-groups of Asians in the Manhattan area fit that of the Lake Tahoe area.

Race	Lake Tahoe Frequency	Manhattan Frequency
Asian Indian	131	174
Chinese	118	557
Filipino	1,045	518
Japanese	80	54
Korean	12	29
Vietnamese	9	21
Other	24	66

Use the following information to answer the next two exercises: UCLA conducted a survey of more than 263,000 college freshmen from 385 colleges in fall 2005. The results of students' expected majors by gender were reported in *The Chronicle of Higher Education (2/2/2006)*. Suppose a survey of 5,000 graduating females and 5,000 graduating males was done as a follow-up last year to determine what their actual majors were. The results are shown in the tables for Exercise and Exercise. The second column in each table does not add to 100% because of rounding.

Q 11.3.6

Conduct a goodness-of-fit test to determine if the actual college majors of graduating females fit the distribution of their expected majors.

Major	Women - Expected Major	Women - Actual Major
Arts & Humanities	14.0%	670
Biological Sciences	8.4%	410
Business	13.1%	685
Education	13.0%	650
Engineering	2.6%	145
Physical Sciences	2.6%	125
Professional	18.9%	975
Social Sciences	13.0%	605
Technical	0.4%	15
Other	5.8%	300
Undecided	8.0%	420

S 11.3.6

a. H_0: The actual college majors of graduating females fit the distribution of their expected majors
b. H_a: The actual college majors of graduating females do not fit the distribution of their expected majors
c. $df = 10$
d. chi-square distribution with $df = 10$
e. $\text{test statistic} = 11.48$
f. $p\text{-value} = 0.3211$
g. Check student's solution.
h.
  i. $\alpha = 0.05$
  ii. Decision: Do not reject null when $a = 0.05$ and $a = 0.01$
  iii. Reason for decision: $p\text{-value} > \alpha$
  iv. Conclusion: There is insufficient evidence to conclude that the distribution of actual college majors of graduating females fits the distribution of their expected majors.

Q 11.3.7

Conduct a goodness-of-fit test to determine if the actual college majors of graduating males fit the distribution of their expected majors.

Major	Men - Expected Major	Men - Actual Major
Arts & Humanities	11.0%	600
Biological Sciences	6.7%	330
Business	22.7%	1130
Education	5.8%	305
Engineering	15.6%	800
Physical Sciences	3.6%	175
Professional	9.3%	460

Major	Men - Expected Major	Men - Actual Major
Social Sciences	7.6%	370
Technical	1.8%	90
Other	8.2%	400
Undecided	6.6%	340

Read the statement and decide whether it is true or false.

Q 11.3.8

In a goodness-of-fit test, the expected values are the values we would expect if the null hypothesis were true.

S 11.3.8

true

Q 11.3.9

In general, if the observed values and expected values of a goodness-of-fit test are not close together, then the test statistic can get very large and on a graph will be way out in the right tail.

Q 11.3.10

Use a goodness-of-fit test to determine if high school principals believe that students are absent equally during the week or not.

S 11.3.10

true

Q 11.3.11

The test to use to determine if a six-sided die is fair is a goodness-of-fit test.

Q 11.3.12

In a goodness-of fit test, if the p-value is 0.0113, in general, do not reject the null hypothesis.

S 11.3.12

false

Q 11.3.13

A sample of 212 commercial businesses was surveyed for recycling one commodity; a commodity here means any one type of recyclable material such as plastic or aluminum. Table shows the business categories in the survey, the sample size of each category, and the number of businesses in each category that recycle one commodity. Based on the study, on average half of the businesses were expected to be recycling one commodity. As a result, the last column shows the expected number of businesses in each category that recycle one commodity. At the 5% significance level, perform a hypothesis test to determine if the observed number of businesses that recycle one commodity follows the uniform distribution of the expected values.

Business Type	Number in class	Observed Number that recycle one commodity	Expected number that recycle one commodity
Office	35	19	17.5
Retail/Wholesale	48	27	24
Food/Restaurants	53	35	26.5
Manufacturing/Medical	52	21	26
Hotel/Mixed	24	9	12

Q 11.3.14

Table contains information from a survey among 499 participants classified according to their age groups. The second column shows the percentage of obese people per age class among the study participants. The last column comes from a different study at the national level that shows the corresponding percentages of obese people in the same age classes in the USA. Perform a hypothesis test at the 5% significance level to determine whether the survey participants are a representative sample of the USA obese population.

Age Class (Years)	Obese (Percentage)	Expected USA average (Percentage)
20–30	75.0	32.6
31–40	26.5	32.6
41–50	13.6	36.6
51–60	21.9	36.6
61–70	21.0	39.7

S 11.3.14

a. H_0: Surveyed obese fit the distribution of expected obese
b. H_a: Surveyed obese do not fit the distribution of expected obese
c. $df = 4$
d. chi-square distribution with $df = 4$
e. $\text{test statistic} = 54.01$
f. $p\text{-value} = 0$
g. Check student's solution.
h. i. $\alpha : 0.05$
   ii. Decision: Reject the null hypothesis.
   iii. Reason for decision: $p\text{-value} < \alpha$
   iv. Conclusion: At the 5% level of significance, from the data, there is sufficient evidence to conclude that the surveyed obese do not fit the distribution of expected obese.

11.4: Test of Independence

For each problem, use a solution sheet to solve the hypothesis test problem. Go to Appendix E for the chi-square solution sheet. Round expected frequency to two decimal places.

Q 11.4.1

A recent debate about where in the United States skiers believe the skiing is best prompted the following survey. Test to see if the best ski area is independent of the level of the skier.

U.S. Ski Area	Beginner	Intermediate	Advanced
Tahoe	20	30	40
Utah	10	30	60
Colorado	10	40	50

Q 11.4.2

Car manufacturers are interested in whether there is a relationship between the size of car an individual drives and the number of people in the driver's family (that is, whether car size and family size are independent). To test this, suppose that 800 car owners were randomly surveyed with the results in Table. Conduct a test of independence.

Family Size	Sub & Compact	Mid-size	Full-size	Van & Truck
1	20	35	40	35

Family Size	Sub & Compact	Mid-size	Full-size	Van & Truck
2	20	50	70	80
3–4	20	50	100	90
5+	20	30	70	70

S 11.4.2

a. H_0: Car size is independent of family size.
b. H_a: Car size is dependent on family size.
c. $df = 9$
d. chi-square distribution with $df = 9$
e. $\text{test statistic} = 15.8284$
f. $p\text{-value} = 0.0706$
g. Check student's solution.
h. i. $\alpha : 0.05$
   ii. Decision: Do not reject the null hypothesis.
   iii. Reason for decision: $p\text{-value} > \alpha$
   iv. Conclusion: At the 5% significance level, there is insufficient evidence to conclude that car size and family size are dependent.

Q 11.4.3

College students may be interested in whether or not their majors have any effect on starting salaries after graduation. Suppose that 300 recent graduates were surveyed as to their majors in college and their starting salaries after graduation. Table shows the data. Conduct a test of independence.

Major	< \$50,000	\$50,000 – \$68,999	\$69,000 +
English	5	20	5
Engineering	10	30	60
Nursing	10	15	15
Business	10	20	30
Psychology	20	30	20

Q 11.4.4

Some travel agents claim that honeymoon hot spots vary according to age of the bride. Suppose that 280 recent brides were interviewed as to where they spent their honeymoons. The information is given in Table. Conduct a test of independence.

Location	20–29	30–39	40–49	50 and over
Niagara Falls	15	25	25	20
Poconos	15	25	25	10
Europe	10	25	15	5
Virgin Islands	20	25	15	5

a. H_0: Honeymoon locations are independent of bride's age.
b. H_a: Honeymoon locations are dependent on bride's age.
c. $df = 9$
d. chi-square distribution with $df = 9$
e. $\text{test statistic} = 15.7027$

f. p-value $= 0.0734$
g. Check student's solution.
h. i. $\alpha : 0.05$
ii. Decision: Do not reject the null hypothesis.
iii. Reason for decision: p-value $> \alpha$
iv. Conclusion: At the 5% significance level, there is insufficient evidence to conclude that honeymoon location and bride age are dependent.

Q 11.4.5

A manager of a sports club keeps information concerning the main sport in which members participate and their ages. To test whether there is a relationship between the age of a member and his or her choice of sport, 643 members of the sports club are randomly selected. Conduct a test of independence.

Sport	18 - 25	26 - 30	31 - 40	41 and over
racquetball	42	58	30	46
tennis	58	76	38	65
swimming	72	60	65	33

Q 11.4.6

A major food manufacturer is concerned that the sales for its skinny french fries have been decreasing. As a part of a feasibility study, the company conducts research into the types of fries sold across the country to determine if the type of fries sold is independent of the area of the country. The results of the study are shown in Table. Conduct a test of independence.

Type of Fries	Northeast	South	Central	West
skinny fries	70	50	20	25
curly fries	100	60	15	30
steak fries	20	40	10	10

S 11.4.6

a. H_0: The types of fries sold are independent of the location.
b. H_a: The types of fries sold are dependent on the location.
c. $df = 6$
d. chi-square distribution with $df = 6$
e. $\text{test statistic} = 18.8369$
f. p-value $= 0.0044$
g. Check student's solution.
h. i. $\alpha : 0.05$
ii. Decision: Reject the null hypothesis.
iii. Reason for decision: p-value $> \alpha$
iv. Conclusion: At the 5% significance level, There is sufficient evidence that types of fries and location are dependent.

Q 11.4.7

According to Dan Lenard, an independent insurance agent in the Buffalo, N.Y. area, the following is a breakdown of the amount of life insurance purchased by males in the following age groups. He is interested in whether the age of the male and the amount of life insurance purchased are independent events. Conduct a test for independence.

Age of Males	None	< \$200,000	\$200,000–\$400,000	\$401,001–\$1,000,000	\$1,000,001+
20–29	40	15	40	0	5

Age of Males	None	< $200,000	$200,000–$400,000	$401,001–$1,000,000	$1,000,001+
30–39	35	5	20	20	10
40–49	20	0	30	0	30
50+	40	30	15	15	10

Q 11.4.8

Suppose that 600 thirty-year-olds were surveyed to determine whether or not there is a relationship between the level of education an individual has and salary. Conduct a test of independence.

Annual Salary	Not a high school graduate	High school graduate	College graduate	Masters or doctorate
< $30,000	15	25	10	5
$30,000–$40,000	20	40	70	30
$40,000–$50,000	10	20	40	55
$50,000–$60,000	5	10	20	60
$60,000+	0	5	10	150

S 11.4.8

a. H_0: Salary is independent of level of education.
b. H_a: Salary is dependent on level of education.
c. $df = 12$
d. chi-square distribution with $df = 12$
e. $\text{test statistic} = 255.7704$
f. $p\text{-value} = 0$
g. Check student's solution.
h. i. $\alpha : 0.05$
   ii. Decision: Reject the null hypothesis.
   iii. Reason for decision: $p\text{-value} > \alpha$
   iv. Conclusion: At the 5% significance level, There is sufficient evidence that types of fries and location are dependent.

Read the statement and decide whether it is true or false.

Q 11.4.9

The number of degrees of freedom for a test of independence is equal to the sample size minus one.

Q 11.4.10

The test for independence uses tables of observed and expected data values.

S 11.4.10

true

Q 11.4.11

The test to use when determining if the college or university a student chooses to attend is related to his or her socioeconomic status is a test for independence.

Q 11.4.12

In a test of independence, the expected number is equal to the row total multiplied by the column total divided by the total surveyed.

S 11.4.12

true

Q 11.4.13

An ice cream maker performs a nationwide survey about favorite flavors of ice cream in different geographic areas of the U.S. Based on Table, do the numbers suggest that geographic location is independent of favorite ice cream flavors? Test at the 5% significance level.

U.S. region/Flavor	Strawberry	Chocolate	Vanilla	Rocky Road	Mint Chocolate Chip	Pistachio	Row total
East	8	31	27	8	15	7	96
Midwest	10	32	22	11	15	6	96
West	12	21	22	19	15	8	97
South	15	28	30	8	15	6	102
Column Total	45	112	101	46	60	27	391

Q 11.4.14

Table provides a recent survey of the youngest online entrepreneurs whose net worth is estimated at one million dollars or more. Their ages range from 17 to 30. Each cell in the table illustrates the number of entrepreneurs who correspond to the specific age group and their net worth. Are the ages and net worth independent? Perform a test of independence at the 5% significance level.

Age Group\ Net Worth Value (in millions of US dollars)	1–5	6–24	≥25	Row Total
17–25	8	7	5	20
26–30	6	5	9	20
Column Total	14	12	14	40

S 11.4.14

a. H_0: Age is independent of the youngest online entrepreneurs’ net worth.
b. H_5: Age is dependent on the net worth of the youngest online entrepreneurs.
c. $df = 2$
d. chi-square distribution with $df = 2$
e. test statistic $= 1.76$
f. p-value $= 0.4144$
g. Check student’s solution.
h. i. $\alpha : 0.05$
   ii. Decision: Do not reject the null hypothesis.
   iii. Reason for decision: p-value $> \alpha$
   iv. Conclusion: At the 5% significance level, there is insufficient evidence to conclude that age and net worth for the youngest online entrepreneurs are dependent.

Q 11.4.15

A 2013 poll in California surveyed people about taxing sugar-sweetened beverages. The results are presented in Table, and are classified by ethnic group and response type. Are the poll responses independent of the participants’ ethnic group? Conduct a test of independence at the 5% significance level.

Opinion/Ethnicity	Asian-American	White/Non-Hispanic	African-American	Latino	Row Total
Against tax	48	433	41	160	628
In Favor of tax	54	234	24	147	459
No opinion	16	43	16	19	84
Column Total	118	710	71	272	1171

11.5: Test for Homogeneity

For each word problem, use a solution sheet to solve the hypothesis test problem. Go to [link] for the chi-square solution sheet. Round expected frequency to two decimal places.

Q 11.5.1

A psychologist is interested in testing whether there is a difference in the distribution of personality types for business majors and social science majors. The results of the study are shown in Table. Conduct a test of homogeneity. Test at a 5% level of significance.

	Open	Conscientious	Extrovert	Agreeable	Neurotic
Business	41	52	46	61	58
Social Science	72	75	63	80	65

S 11.5.1

a. H_0: The distribution for personality types is the same for both majors
b. H_a: The distribution for personality types is not the same for both majors
c. $df = 4$
d. chi-square with $df = 4$
e. $\text{test statistic} = 3.01$
f. $p\text{-value} = 0.5568$
g. Check student's solution.
h.
  i. $\alpha : 0.05$
  ii. Decision: Do not reject the null hypothesis.
  iii. Reason for decision: $p\text{-value} > \alpha$
  iv. Conclusion: There is insufficient evidence to conclude that the distribution of personality types is different for business and social science majors.

Q 11.5.2

Do men and women select different breakfasts? The breakfasts ordered by randomly selected men and women at a popular breakfast place is shown in Table. Conduct a test for homogeneity at a 5% level of significance.

	French Toast	Pancakes	Waffles	Omelettes
Men	47	35	28	53
Women	65	59	55	60

Q 11.5.3

A fisherman is interested in whether the distribution of fish caught in Green Valley Lake is the same as the distribution of fish caught in Echo Lake. Of the 191 randomly selected fish caught in Green Valley Lake, 105 were rainbow trout, 27 were other trout, 35 were bass, and 24 were catfish. Of the 293 randomly selected fish caught in Echo Lake, 115 were rainbow trout, 58 were other trout, 67 were bass, and 53 were catfish. Perform a test for homogeneity at a 5% level of significance.

S 11.5.3

a. H_0: The distribution for fish caught is the same in Green Valley Lake and in Echo Lake.
b. H_a: The distribution for fish caught is not the same in Green Valley Lake and in Echo Lake.
c. $df = 3$
d. chi-square with $df = 3$
e. $\text{test statistic} = 11.75$
f. $p\text{-value} = 0.0083$
g. Check student's solution.
h. i. $\alpha : 0.05$
  ii. Decision: Reject the null hypothesis.
  iii. Reason for decision: $p\text{-value} > \alpha$
  iv. Conclusion: There is evidence to conclude that the distribution of fish caught is different in Green Valley Lake and in Echo Lake

Q 11.5.4

In 2007, the United States had 1.5 million homeschooled students, according to the U.S. National Center for Education Statistics. In Table you can see that parents decide to homeschool their children for different reasons, and some reasons are ranked by parents as more important than others. According to the survey results shown in the table, is the distribution of applicable reasons the same as the distribution of the most important reason? Provide your assessment at the 5% significance level. Did you expect the result you obtained?

Reasons for Homeschooling	Applicable Reason (in thousands of respondents)	Most Important Reason (in thousands of respondents)	Row Total
Concern about the environment of other schools	1,321	309	1,630
Dissatisfaction with academic instruction at other schools	1,096	258	1,354
To provide religious or moral instruction	1,257	540	1,797
Child has special needs, other than physical or mental	315	55	370
Nontraditional approach to child's education	984	99	1,083
Other reasons (e.g., finances, travel, family time, etc.)	485	216	701
Column Total	5,458	1,477	6,935

Q 11.5.5

When looking at energy consumption, we are often interested in detecting trends over time and how they correlate among different countries. The information in Table shows the average energy use (in units of kg of oil equivalent per capita) in the USA and the joint European Union countries (EU) for the six-year period 2005 to 2010. Do the energy use values in these two areas come from the same distribution? Perform the analysis at the 5% significance level.

Year	European Union	United States	Row Total
2010	3,413	7,164	10,557
2009	3,302	7,057	10,359
2008	3,505	7,488	10,993

Year	European Union	United States	Row Total
2007	3,537	7,758	11,295
2006	3,595	7,697	11,292
2005	3,613	7,847	11,460
Column Total	45,011	20,965	65,976

S 11.5.5

a. H_0: The distribution of average energy use in the USA is the same as in Europe between 2005 and 2010.
b. H_a: The distribution of average energy use in the USA is not the same as in Europe between 2005 and 2010.
c. $df = 4$
d. chi-square with $df = 4$
e. $\text{test statistic} = 2.7434$
f. $p\text{-value} = 0.7395$
g. Check student's solution.
h. i. $\alpha : 0.05$
   ii. Decision: Do not reject the null hypothesis.
   iii. Reason for decision: $p\text{-value} > \alpha$
   iv. Conclusion: At the 5% significance level, there is insufficient evidence to conclude that the average energy use values in the US and EU are not derived from different distributions for the period from 2005 to 2010.

Q 11.5.6

The Insurance Institute for Highway Safety collects safety information about all types of cars every year, and publishes a report of Top Safety Picks among all cars, makes, and models. Table presents the number of Top Safety Picks in six car categories for the two years 2009 and 2013. Analyze the table data to conclude whether the distribution of cars that earned the Top Safety Picks safety award has remained the same between 2009 and 2013. Derive your results at the 5% significance level.

Year \ Car Type	Small	Mid-Size	Large	Small SUV	Mid-Size SUV	Large SUV	Row Total
2009	12	22	10	10	27	6	87
2013	31	30	19	11	29	4	124
Column Total	43	52	29	21	56	10	211

11.6: Comparison of the Chi-Square Tests

For each word problem, use a solution sheet to solve the hypothesis test problem. Go to [link] for the chi-square solution sheet. Round expected frequency to two decimal places.

Q 11.6.1

Is there a difference between the distribution of community college statistics students and the distribution of university statistics students in what technology they use on their homework? Of some randomly selected community college students, 43 used a computer, 102 used a calculator with built in statistics functions, and 65 used a table from the textbook. Of some randomly selected university students, 28 used a computer, 33 used a calculator with built in statistics functions, and 40 used a table from the textbook. Conduct an appropriate hypothesis test using a 0.05 level of significance.

S 11.6.1

a. H_0: The distribution for technology use is the same for community college students and university students.
b. H_a: The distribution for technology use is not the same for community college students and university students.
c. $df = 2$
d. chi-square with $df = 2$

e. test statistic $= 7.05$
f. p-value $= 0.0294$
g. Check student's solution.
h. i. $\alpha : 0.05$
 ii. Decision: Reject the null hypothesis.
 iii. Reason for decision: p-value $> \alpha$
 iv. Conclusion: There is sufficient evidence to conclude that the distribution of technology use for statistics homework is not the same for statistics students at community colleges and at universities.

Read the statement and decide whether it is true or false.

Q 11.6.2

If $df = 2$, the chi-square distribution has a shape that reminds us of the exponential.

11.7: Test of a Single Variance

Use the following information to answer the next twelve exercises: Suppose an airline claims that its flights are consistently on time with an average delay of at most 15 minutes. It claims that the average delay is so consistent that the variance is no more than 150 minutes. Doubting the consistency part of the claim, a disgruntled traveler calculates the delays for his next 25 flights. The average delay for those 25 flights is 22 minutes with a standard deviation of 15 minutes.

Q 11.7.1

Is the traveler disputing the claim about the average or about the variance?

Q 11.7.2

A sample standard deviation of 15 minutes is the same as a sample variance of __________ minutes.

S 11.7.2

225

Q 11.7.3

Is this a right-tailed, left-tailed, or two-tailed test?

Q 11.7.4

H_0: __________

S 11.7.4

$H_0 : \sigma^2 \leq 150$

Q 11.7.5

$df =$ ________

Q 11.7.6

chi-square test statistic = ________

S 11.7.6

36

Q 11.7.7

p-value $=$ ________

Q 11.7.8

Graph the situation. Label and scale the horizontal axis. Mark the mean and test statistic. Shade the p-value.

S 11.7.8

Check student's solution.

Q 11.7.9

Let $\alpha = 0.05$

Decision: ________

Conclusion (write out in a complete sentence.): ________

Q 11.7.10

How did you know to test the variance instead of the mean?

S 11.7.10

The claim is that the variance is no more than 150 minutes.

Q 11.7.11

If an additional test were done on the claim of the average delay, which distribution would you use?

Q 11.7.12

If an additional test were done on the claim of the average delay, but 45 flights were surveyed, which distribution would you use?

S 11.7.12

a Student's t- or normal distribution

For each word problem, use a solution sheet to solve the hypothesis test problem. Go to [link] for the chi-square solution sheet. Round expected frequency to two decimal places.

Q 11.7.13

A plant manager is concerned her equipment may need recalibrating. It seems that the actual weight of the 15 oz. cereal boxes it fills has been fluctuating. The standard deviation should be at most 0.5 oz. In order to determine if the machine needs to be recalibrated, 84 randomly selected boxes of cereal from the next day's production were weighed. The standard deviation of the 84 boxes was 0.54. Does the machine need to be recalibrated?

Q 11.7.14

Consumers may be interested in whether the cost of a particular calculator varies from store to store. Based on surveying 43 stores, which yielded a sample mean of $84 and a sample standard deviation of $12, test the claim that the standard deviation is greater than $15.

S 11.7.14

a. $H_0 : \sigma = 15$
b. $H_a : \sigma > 15$
c. $df = 42$
d. chi-square with $df = 42$
e. test statistic = 26.88
f. p-value $= 0.9663$
g. Check student's solution.
h. i. $alpha = 0.05$
   ii. Decision: Do not reject null hypothesis.
   iii. Reason for decision: p-value $> \alpha$
   iv. Conclusion: There is insufficient evidence to conclude that the standard deviation is greater than 15.

Q 11.7.15

Isabella, an accomplished **Bay to Breakers** runner, claims that the standard deviation for her time to run the 7.5 mile race is at most three minutes. To test her claim, Rupinder looks up five of her race times. They are 55 minutes, 61 minutes, 58 minutes, 63 minutes, and 57 minutes.

Q 11.7.16

Airline companies are interested in the consistency of the number of babies on each flight, so that they have adequate safety equipment. They are also interested in the variation of the number of babies. Suppose that an airline executive believes the average number of babies on flights is six with a variance of nine at most. The airline conducts a survey. The results of the 18 flights surveyed give a sample average of 6.4 with a sample standard deviation of 3.9. Conduct a hypothesis test of the airline executive's belief.

S 11.7.16

a. $H_0 : \sigma \leq 3$
b. $H_a : \sigma > 3$
c. $df = 17$
d. chi-square distribution with $df = 17$
e. test statistic = 28.73
f. $p\text{-value} = 0.0371$
g. Check student's solution.
h. i. $\alpha : 0.05$
   ii. Decision: Reject the null hypothesis.
   iii. Reason for decision: $p\text{-value} < \alpha$
   iv. Conclusion: There is sufficient evidence to conclude that the standard deviation is greater than three.

Q 11.7.17

The number of births per woman in China is 1.6 down from 5.91 in 1966. This fertility rate has been attributed to the law passed in 1979 restricting births to one per woman. Suppose that a group of students studied whether or not the standard deviation of births per woman was greater than 0.75. They asked 50 women across China the number of births they had had. The results are shown in Table. Does the students' survey indicate that the standard deviation is greater than 0.75?

# of births	Frequency
0	5
1	30
2	10
3	5

Q 11.7.18

According to an avid aquarist, the average number of fish in a 20-gallon tank is 10, with a standard deviation of two. His friend, also an aquarist, does not believe that the standard deviation is two. She counts the number of fish in 15 other 20-gallon tanks. Based on the results that follow, do you think that the standard deviation is different from two? Data: 11; 10; 9; 10; 10; 11; 11; 10; 12; 9; 7; 9; 11; 10; 11

S 11.7.18

a. $H_0 : \sigma = 2$
b. $H_a : \sigma \neq 2$
c. $df = 14$
d. chi-square distribution with $df = 14$
e. chi-square test statistic = 5.2094
f. $p\text{-value} = 0.0346$
g. Check student's solution.
h. i. $\alpha : 0.05$
   ii. Decision: Reject the null hypothesis
   iii. Reason for decision: $p\text{-value} < \alpha$

iv. Conclusion: There is sufficient evidence to conclude that the standard deviation is different than 2.

Q 11.7.19

The manager of "Frenchies" is concerned that patrons are not consistently receiving the same amount of French fries with each order. The chef claims that the standard deviation for a ten-ounce order of fries is at most 1.5 oz., but the manager thinks that it may be higher. He randomly weighs 49 orders of fries, which yields a mean of 11 oz. and a standard deviation of two oz.

Q 11.7.20

You want to buy a specific computer. A sales representative of the manufacturer claims that retail stores sell this computer at an average price of $1,249 with a very narrow standard deviation of $25. You find a website that has a price comparison for the same computer at a series of stores as follows: $1,299; $1,229.99; $1,193.08; $1,279; $1,224.95; $1,229.99; $1,269.95; $1,249. Can you argue that pricing has a larger standard deviation than claimed by the manufacturer? Use the 5% significance level. As a potential buyer, what would be the practical conclusion from your analysis?

S 11.7.20

a. $H_0 : \sigma = 25^2$
b. $H_a : \sigma > 25^2$
c. $df = n - 1 = 7$
d. test statistic: $\chi^2 = \chi^2_7 = \frac{(n-1)s^2}{25^2} = \frac{(8-1)(34.29)^2}{25^2} = 13.169$
e. p-value : $P(\chi^2_7 > 13.169) = 1 - P(\chi^2_7 \leq 13.169) = 0.0681$
f. i. $\alpha : 0.05$
 ii. Decision: Do not reject the null hypothesis
 iii. Reason for decision: p-value $< \alpha$
 iv. Conclusion: At the 5% level, there is insufficient evidence to conclude that the variance is more than 625.

Q 11.7.21

A company packages apples by weight. One of the weight grades is Class A apples. Class A apples have a mean weight of 150 g, and there is a maximum allowed weight tolerance of 5% above or below the mean for apples in the same consumer package. A batch of apples is selected to be included in a Class A apple package. Given the following apple weights of the batch, does the fruit comply with the Class A grade weight tolerance requirements. Conduct an appropriate hypothesis test.

a. at the 5% significance level
b. at the 1% significance level

Weights in selected apple batch (in grams): 158; 167; 149; 169; 164; 139; 154; 150; 157; 171; 152; 161; 141; 166; 172;

11.8: Lab 1: Chi-Square Goodness-of-Fit

11.9: Lab 2: Chi-Square Test of Independence
