Chapter 11
Study Guide

Chi-Square

THE ANALYSIS OF CATEGORICAL DATA AND GOODNESS-OF-FIT TESTS

This section will explore inference for categorical data using chi-square test procedures. These procedures are used with univariate as well as bivariate data sets if the variables are categorical.

OBJECTIVES

- Carry out a chi-square goodness-of-fit test.
- Carry out chi-square tests for homogeneity of proportions and for independence.

CHI-SQUARE TESTS FOR UNIVARIATE DATA

(*Introduction to Statistics & Data Analysis* 3rd ed. pages 647–656/4th ed. pages 700–708)

In this section, we extend techniques for analyzing univariate categorical data sets. Here we can consider questions about variables that involve two or more categories. A chi-square test (χ^2) for univariate data allows us to test hypotheses about the proportions falling into the different categories for a categorical variable. We do this by looking at the frequencies observed in each of the categories and comparing them to what would be expected if a null hypothesis were true. The hypothesized proportions can be equal or they can be different for each category.

EXAMPLE A new donut shop plans to sell plain, strawberry, blueberry, and cinnamon donuts. They wonder if there is a preference for one of these types of donuts or if each type is preferred by the same proportion of customers. If we let

p_p = proportion of customers preferring plain donuts
p_s = proportion of customers preferring strawberry donuts
p_b = proportion of customers preferring blueberry donuts
p_c = proportion of customers preferring cinnamon donuts

We are interested in knowing if $p_p = p_s = p_b = p_c = 0.25$ or if there is evidence that these proportions are not all the same. To answer this question, a random sample of 60 customers is surveyed, and each person in the sample is asked which of the four donut types they prefer.

Instead of running a separate two-sample proportions test for each of the possible pairs of proportions, a chi-square test will allow us to decide if the proportions we have observed in our sample are significantly different from the hypothesized proportions. The chi-square goodness-of-fit test will allow us to test the following hypotheses:

$H_0 : p_p = p_s = p_b = p_c$
H_a : not all of the proportions are equal to 0.25

(For chi-square tests, it is acceptable to state hypotheses in words on the AP Exam.)

To make the decision to reject or fail to reject the null hypothesis, the chi-square test will compare the number of responses observed in each category to what we would expect to see in each category if the null hypothesis is true. If the difference is too large, we reject the null hypothesis. Otherwise, we fail to reject, the null hypothesis.

Data on a single categorical variable is usually summarized using a one-way frequency table. Returning to the previous example, suppose that the sample of 60 customers resulted in the data summarized in the table below. The table entries are observed frequencies or counts.

Types of Donut Preferred

	Plain	Strawberry	Blueberry	Cinnamon
Observed Count	13	12	16	19

These counts represent the number of times a person in the sample selected each particular donut type. *Expected* counts are calculated using the hypothesized proportions from the null hypothesis. In this example, all four hypothesized proportions are equal to 0.25, so if the null hypothesis is true, we would expect to see the same number preferring each type. We calculate expected counts by multiplying the sample size by each hypothesized proportion:

$np_p = 60(0.25) = 15$
$np_s = 60(0.25) = 15$
$np_b = 60(0.25) = 15$
$np_c = 60(0.25) = 15$

The expected counts can now be entered into the table as shown here.

Types of Donuts Sold

	Plain	Strawberry	Blueberry	Cinnamon
Observed Count	13	12	16	19
Expected Count	60(0.25) = 15	60(0.25) =15	60(0.25) = 15	60(0.25) = 15

Now we are ready to calculate the value of the test statistic. This test statistic is called the *goodness-of-fit* statistic because it considers how well the observed values fit to what we expected to see. Calculating the value of a chi-square (χ^2) test statistic is relatively easy.

$$\chi^2 = \Sigma \frac{(\text{observed cell count} - \text{expected cell count})^2}{\text{expected cell count}}$$

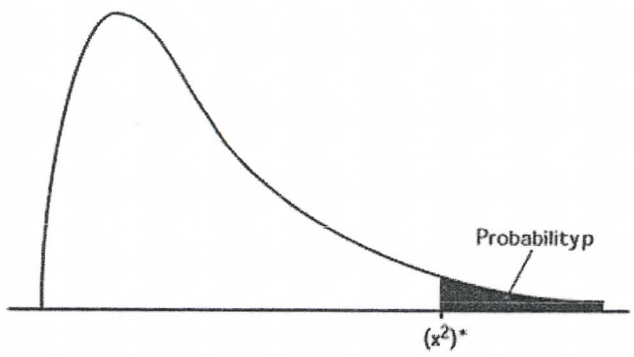
which is often abbreviated

$$\chi^2 = \Sigma \frac{(O - E)^2}{E}$$

For the example problem, the value of this test statistic is

$$\begin{aligned}\chi^2 &= \frac{(13-15)^2}{15} + \frac{(12-15)^2}{15} + \frac{(16-15)^2}{15} + \frac{(19\text{-}15)^2}{15} \\ &= \frac{(^-2)^2}{15} + \frac{(^-3)^2}{15} + \frac{(1)^2}{15} + \frac{(4)^2}{15} \\ &= \frac{4}{15} + \frac{9}{15} + \frac{1}{15} + \frac{16}{15} \\ &= \frac{30}{15} \\ &= 2\end{aligned}$$

To find the associated P-value, we use a chi-square distribution with df = k - 1, where k is the number of categories of a categorical variable. The P-value is the area to the right of the computed test statistic value (the chi-square goodness-of-fit test is an upper tail test). This area is found by referring to the chi-square table or by using a graphing calculator or other technology.



The chi-square table is similar to the t table. You locate the df in the leftmost column then read across the row to find the value of the χ^2 test statistic. If the value is between two columns, then the P-value is between the two corresponding tail probabilities.

Chi-Square Tail Probability (χ^2)						
df	0.25	0.20	0.15	0.10	0.05	0.025
1	1.32	1.64	2.07	2.71	3.84	5.02
2	2.77	3.22	3.79	4.61	5.99	7.38
3	4.11	4.64	5.32	6.25	7.81	9.35
4	5.39	5.99	6.74	7.78	9.49	11.14
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Since our $\chi^2 = 2.00$ and this value is smaller than 4.11, we know the tail probability is greater than 0.25. Using any reasonable α level will result in a failure to reject the null hypothesis.

The only things that we have not yet discussed are the overall assumptions required for the chi-square test. The first assumption is that we have a random sample of observations of a categorical variable. In addition, the sample size must be large enough that the following conditions are met:

1. No expected counts are < 1,
2. All of our expected counts should be ≥ 5 and if they are not,
3. No more than 20% of the expected counts are < 5.

Returning to the donut example:

HYPOTHESIS

$H_0 : p_p = 0.25,\ p_s = 0.25,\ p_b = 0.25,\ p_c = 0.25$

$H_a : H_0$ is not true; in other words, not all of the proportions are equal

We will use a significance level of 0.05 for this test.

TEST AND ASSUMPTIONS

Test: Chi-square goodness-of-fit test

Assumptions:

1. The sample was a random sample of customers
2. All expected counts are $\geq$ 5, so the sample size is large enough.

TEST STATISTIC As seen earlier, the calculated test statistic for this test would be

$$\begin{aligned}\chi^2 &= \frac{(13-15)^2}{15}+\frac{(12-15)^2}{15}+\frac{(16-15)^2}{15}+\frac{(19-15)^2}{15}\\ &= \frac{(-2)^2}{15}+\frac{(-3)^2}{15}+\frac{(1)^2}{15}+\frac{(4)^2}{15}\\ &= \frac{4}{15}+\frac{9}{15}+\frac{1}{15}+\frac{16}{15}\\ &= \frac{30}{15}\\ &= 2\end{aligned}$$

with $df = 3$

Based on $df = 3$ and $\chi^2 = 2$, the P-value is greater than 0.25.

CONCLUSION Since P-value > α, there is insufficient evidence to reject the null hypothesis. In other words, there is not convincing evidence that the four types of donuts are not equally preferred.

AP Tip

For all chi-square tests, make sure to provide the expected counts and verify that they are large enough.

SAMPLE PROBLEM 1 A music producer is interested in marketing a new artist via ads in movie theaters. The target age group is teenagers from 14–18 years of age. The company developing the advertisement has offered to run the ad in conjunction with the following types of movies: 20% of the time with comedies, 50% with dramas, and 30% with action films. However, the movie producer is not sure these percentages reflect the types of movies that teens attend. To investigate, each person in a random sample of teens was asked what type of movie they had seen most recently, resulting in the following data:

Movie Most Recently Watched By 100 Teens

	Comedy	Drama	Action
Observed Count	41	35	24

Do these data provide convincing evidence that the proportions of teens watching the different types of movies are different than the proportions proposed by the company developing the ad? Use a significance level of 0.05.

Solution to Problem 1

Hypothesis

p_c = proportion of teens who watched a comedy
p_d = proportion of teens who watched a drama
p_a = proportion of teens who watched an action films

$H_0 : p_c = 0.2;\quad p_d = 0.5;\quad p_a = 0.3$
H_a : the null hypothesis is not true (at least one of the proportions of teens watching the various film types is not equal to the hypothesized proportion).
$\alpha = 0.05$

Assumptions The problem states the teens were a random sample.
All the expected counts are greater than or equal to 5 (see table below).

Movie Most Recently Watched By 100 Teens

	Comedy	Drama	Action
Expected Counts	100(0.2) = 20	100(0.5) = 50	100(0.3) = 30

Test Statistic

χ^2 Goodness-of-fit test:

$$\begin{aligned}\chi^2 &= \frac{(41-20)^2}{20} + \frac{(35-50)^2}{50} + \frac{(24-30)^2}{30} \\ &= \frac{(21)^2}{20} + \frac{(-15)^2}{50} + \frac{(-6)^2}{30} \\ &= 22.05 + 4.5 + 1.2 \\ &= 27.75 \qquad \text{with } df = 2\end{aligned}$$

Conclusion The P-value associated with $\chi^2 = 27.72$ and $df = 2$ is approximately 0.000. Since 0.000 < 0.05, there is strong evidence to reject the null hypothesis. In other words, based on this sample there is strong evidence that at least one of the proportions is not equal to the proportion stated by the ad company.

Just as with other hypothesis tests, it is important to address all parts of the test. Clearly conveying your understanding of the procedure will include

- explaining the notation used,
- addressing all assumptions that are required for the test,
- demonstrating correct mechanics,
- and finally writing a conclusion in context.

As with other tests, the mechanics can be done using a graphing calculator, but it is strongly recommended that you still show the initial set-up. This lets the AP Reader know you understand how to compute the value of the test statistic.

TESTS FOR HOMOGENEITY AND INDEPENDENCE IN A TWO-WAY TABLE

(*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722)

A chi-square test procedure can also be used with bivariate categorical data, which is usually displayed in a two-way frequency table. There are two different types of investigations that arise from this type of data. One type involves inferences about association between two different categorical variables being observed on a single sample. The other type involves comparing two or more populations or treatments when a single categorical variable is observed. The calculation procedure is the same for both types of investigation, but the primary question of interest is different.

TEST FOR HOMOGENEITY

In a chi-square test for homogeneity, we are interested in whether the proportions falling into each of the possible categories of a categorical variable are the same for all of the treatments or populations studied. In this case, the null hypothesis is that the distribution of the categorical variable is the same for each population or treatment.

EXAMPLE A snack manufacturer produces three types of chips in two different locations (Location A and Location B). Sometimes bags of chips are damaged in the packaging process. Each bag in a random sample of 45 bags of chips packaged at Location A and in a random sample of 30 bags of chips packaged at Location B was classified into one of three categories: no damage, minimal damage, and severe damage. The resulting data is summarized in the table below. The manufacturer was interested in determining if there was sufficient evidence to conclude that the proportions falling into each of the three damage categories is not the same for the two locations.

	No Damage	Minimal Damage	Severe Damage	Row Totals
Location A	15	18	12	45
Location B	8	12	10	30
Column Totals	23	30	22	75

To answer the question of interest, the hypotheses would be

H_0: There is no difference in the proportions falling into each damage category for the two locations

H_a: The proportions falling into each damage category are not the same for the two locations

The assumptions that need to be are (1) the sample must be a random sample and (2) all expected counts must be at least 5.

The expected counts for a chi-square test of homogeneity are calculated using the following formula:

$$\frac{\text{(row total)(column total)}}{\text{table total}}$$

For example, the expected count for the cell of no damage and Location A is

$$\frac{(53)(23)}{76} = 16.04$$

A way to record these expected counts is either in a separate two-way table of just the expected counts or in parenthesis beside each observed count as shown in the table below. All of the expected counts are greater than or equal to 5.

	No Damage	Minimal Damage	Severe Damage	Row Totals
Location A	19 (16.04)	20 (18.13)	14 (18.83)	53
Location B	4 (6.96)	6 (7.87)	13 (8.17)	23
Column Totals	23	26	27	76

Degrees of freedom for the chi-square test of homogeneity is computed as follows:

$df = (r-1)(c-1)$,
where r = number of rows and c = number of columns

Note that the total column and row are not counted in computing df.

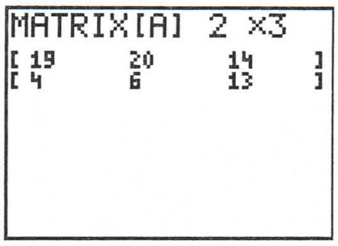
The chi-square test statistic that was used in the goodness-of-fit test is also used here.

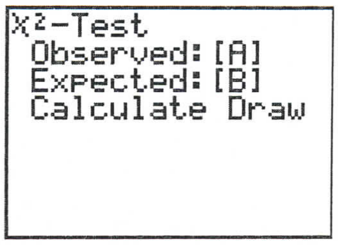
χ^2 test for homogeneity

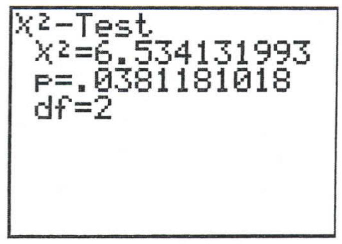
$$\chi^2 = \frac{(19-16.04)^2}{16.04} + \frac{(20-18.13)^2}{18.13} + \frac{(14-18.83)^2}{18.83} + \frac{(4-6.96)^2}{6.96} + \frac{(6-7.87)^2}{7.87} + \frac{(13-8.17)^2}{8.17}$$
$$= 6.5$$

with $df = (2-1)(3-1) = 2$

Using the graphing calculator technology, you first enter a matrix of the observed counts and then run a chi-square test from the test menu. The calculator will calculate the expected counts and store them in matrix[B].







Finally, since $P < \alpha$ $(0.038 < 0.05)$, we reject the null hypothesis. There is strong evidence that the proportions falling into each of the three damage categories is not the same for both locations.

TEST FOR INDEPENDENCE

The chi-square test for independence, also known as the chi-square test for association, is used to investigate if there is an association between two categorical variables. The calculations will proceed in the same manner as the test for homogeneity; however, we are actually looking to see if knowing the value of one variable provides information about the value of the other variable.

EXAMPLE A car manufacturer has two production lines building three types of cars. An engineer is wondering if there is an association between the type of car and the production line that made the vehicle for cars that are found to have major defects. Each car in a random sample of 75 cars selected from all cars found to have major defects was classified according to the type of car and the production line that produced the car. The resulting data is given in the table below.

	Sedan	Wagon	Truck
Line A	13	9	12
Line B	18	12	11

The question of interest is whether there is an association between car type and production line for cars with major defects. To answer this question, we use the chi-square test for independence.

HYPOTHESIS

H_0 : Production line and car type are independent.

H_a : Production line and car type are not independent.

$\alpha = 0.05$

ASSUMPTIONS The sample was a random sample of cars with major defects.

The expected counts are all greater than or equal to 5, so we can proceed (see table below).

Expected Counts for Auto Errors

	Sedan	Wagon	Truck
Line A	14.1	9.5	10.4
Line B	16.9	11.5	12.6

TEST STATISTIC Since this is a χ^2 test for independence, the calculations would be

$$\chi^2 = \frac{(13-14.1)^2}{14.1} + \frac{(9-9.5)^2}{9.5} + \frac{(12-10.4)^2}{10.4} + \frac{(18-16.9)^2}{16.9} + \frac{(12-11.5)^2}{11.5} + \frac{(11-12.6)^2}{12.6} = 0.63$$

$df = 2$

$P = 0.73$

CONCLUSION Since 0.73 is not smaller than 0.05, we fail to reject the null hypothesis. Therefore, we do not have sufficient evidence of an association between car type and production line.

Notice, the steps in the chi-square hypothesis tests are the same as for all other tests: State hypotheses, identify the test by name or formula and check assumptions required for the test, calculate the value of the test statistic and the associated P-value, and write a conclusion in context that is linked to the P-value.

INTERPRETING AND COMMUNICATING THE RESULTS OF STATISTICAL ANALYSES

(*Introduction to Statistics & Data Analysis* 3rd ed. pages 677–680/4th ed. pages 727–730)

This particular set of tests involved writing our conclusions in terms that match the setting of the individual test. Each chi-square test is performed in a different setting depending on the data we have to work with. It is important to state which of the three chi-square tests you are using when carrying out one of these tests. The goodness-of-fit, homogeneity, and independence tests all have different hypotheses. Be clear which test you are using.

CATEGORICAL DATA AND GOODNESS-OF-FIT TESTS: STUDENT OBJECTIVES FOR THE AP EXAM

- You will be able to determine degrees of freedom for chi-square tests of goodness-of-fit, homogeneity, and independence.
- You will be able to calculate expected counts for chi-square tests of goodness-of-fit, homogeneity, and independence.
- You will be able to carry out chi-square tests of goodness-of-fit, homogeneity, and independence.

- You will identify the correct hypothesis for each of the three chi-square scenarios.
- You will be able to interpret conclusions in contest for chi-square tests of goodness-of-fit, homogeneity, and independence.

MULTIPLE-CHOICE QUESTIONS

1. Each car in a random sample of 200 cars sold in 2010 at a large car dealership was classified by color. Sixty of the cars were white, 80 were blue, 20 were silver and 10 were red. What is the value of the chi-square statistic in a test of $H_0 : p_w = p_b = p_s = p_r = 0.25$?
(A) 18
(B) 32
(C) 50
(D) 70
(E) 110

2. Which of the following is not true of the χ^2 probability distribution?
(A) For small degrees of freedom, the distribution is right skewed.
(B) As the degrees of freedom increase, the χ^2 distribution approaches a normal distribution.
(C) All of the area under a chi-square curve is associated with positive values.
(D) The total area under the χ^2 curve is equal to one.
(E) The mean of a chi-square distribution is 0.

3. There are four surgical methods currently being used to place medical implants in patients. After surgery, patients are monitored and their pain level is recorded as severe, moderate, or mild. What are the degrees of freedom for a test to determine if there is an association between surgical method and pain level?
(A) 2
(B) 3
(C) 4
(D) 6
(E) 12

Questions 4–7 refer to the following set of data:

A group of AP Statistics students wanted to see if plain, peanut, and almond M&Ms have the same color distribution. To test this, students took a random sample of each type of M&M and classified the candies in the sample by color. They plan to carry out a chi-square test to decide if there is evidence that the color distributions are not the same for the three types of M&Ms.

	Red	Blue	Yellow	Green	Orange	Brown
Plain	20	18	15	10	14	12
Peanut	8	6	8	25	5	7
Almond	7	11	10	12	10	9

4. What is the expected count for green peanut M&Ms?
   (A) $\frac{25}{47}$
   (B) $\frac{25\times 47}{47}$
   (C) $\frac{59\times 47}{59}$
   (D) $\frac{59\times 47}{207}$
   (E) $\frac{25\times 47}{207}$

5. What are the correct degrees of freedom for the appropriate test?
   (A) 206
   (B) 18
   (C) 10
   (D) 7
   (E) 4

6. What is an appropriate set of hypotheses?
   (A) H_0: There is no association between the type of M&M and color.
   H_a: There is an association between the type of M&M and color.
   (B) H_0: There is an association between the type of M&M and color.
   H_a: There is no association between the type of M&M and color.
   (C) H_0: The color proportions are the same for all 3 types of M&Ms.
   H_a: The color proportions are not the same for all 3 types of M&Ms.
   (D) H_0: The proportions of red M&Ms are the same for all 3 types of M&Ms.
   H_a: The proportions of red M&Ms are not the same for all 3 types of M&Ms.
   (E) H_0: There is no difference in the type of M&M for green M&Ms.
   H_a: There is a difference in the type of M&M for green M&Ms.

7. The value of the chi-square test statistic is 22.9. What conclusion would be reached in a test of the hypotheses of interest using a significance level of 0.05?
   (A) Fail to reject H_0 and conclude there is strong evidence of an association.
   (B) Fail to reject H_0 and conclude there is strong evidence that the color distributions are the same.
   (C) Fail to reject H_0 and conclude there is not strong evidence that the color distributions are the same.
   (D) Reject H_0 and conclude there is strong evidence of an association.
   (E) Reject H_0 and conclude there is strong evidence that the color distributions are not the same.

8. A cupcake store manager believes that half of the cupcakes sold are chocolate and that the other half is equally divided between vanilla and cherry. If the manager's belief is correct and 300 cupcakes are sold, how many would you expect to be cherry cupcakes?
   (A) 300
   (B) 200
   (C) 150
   (D) 100
   (E) 75

9. Each person in a random sample of patrons of a local mall was surveyed regarding a public smoking area outside one of the mall entrances. Each person was asked if they approved of the idea of a public smoking area in the mall. The resulting data is summarized in the table below. The mall management would like to know if there is a relationship between gender and approval of the smoking area. What would be an appropriate set of hypotheses?

Public Smokers

	Approve	Do Not Approve
Males	28	57
Females	39	31

   (A) H_0: Gender and approval are not independent.
   H_a: Gender and approval are independent.
   (B) H_0: Gender and approval are independent.
   H_a: Gender and approval are not independent.
   (C) H_0: Knowing a person does not approve of a public smoking area indicates their gender.
   H_a: Knowing a person does not approve of a public smoking area does not indicate their gender.
   (D) H_0: There is no difference in gender distribution based on approval.
   H_a: There is a difference in gender distribution based on approval.
   (E) H_0: There is an association between gender and approval.
   H_a: There is no association between gender and approval.

10. A local bagel shop makes six types of bagels and eight types of cream cheese toppings. Suppose that all customers order one type of bagel and one type of cream cheese. If customers are classified according to the type of bagel and type of cream cheese, how many degrees of freedom would be appropriate for a test to determine if there is an association between type of bagel and type of cream cheese purchased?
(A) 7
(B) 12
(C) 14
(D) 35
(E) 48

11. In a recent opinion poll, likely voters were asked if they would continue to support health care reform if the cost per taxpaying citizen was increased by $800 per year. Political preference was also recorded. The *P*-value for a chi-square test of independence is 0.001. Which of the following is a correct interpretation of this result?

	Democrat	Republican	Independent	Libertarian
Support	125	87	99	48
Don't Support	75	113	101	52

(A) Since $P = 0.001$, there is strong evidence that political preference and support are independent of each other. In other words, knowing their political preference gives no insight into their support of the increase.
(B) Since $P = 0.001$, there is strong evidence that political preference and support are not independent of each other. In other words, knowing their political preference gives insight into their support of the increase.
(C) Since $P = 0.001$, there is strong evidence that political preference and support are independent of each other. Since they are independent, there is no association between the two variables.
(D) Since $P = 0.001$, there is insufficient evidence that political preference and support are independent of each other. Since they are not independent, there is no association between the two variables.
(E) Since $P = 0.001$, there is insufficient evidence that political preference and support are not independent of each other. In this case, knowing their political preference would not help in knowing their support.

12. Two different cereal companies each make a bran cereal, a corn cereal, and a rice cereal. Each person in a random sample of 300 potential customers tasted all six cereals and selected their favorite. The choice was then classified according to type of cereal and the company that made the cereal. The data is shown in the table below. What type of test would be appropriate?

	Cereal 1	Cereal 2	Cereal 3
Company A	51	48	47
Company B	43	65	46

(A) Chi-square goodness-of-fit test
(B) Chi-square test of independence
(C) Chi-square test of homogeneity
(D) Either a chi-square test for independence or homogeneity is appropriate
(E) A chi-square test is not appropriate in this situation

13. Which is not true for a chi-square test of independence?
I. The test is based on bivariate categorical data.
II. *df* is calculated using $k - 1$, where k = the number of categories.
III. The test is an upper-tailed test.
(A) I, II, and III
(B) I and II
(C) II and III
(D) I only
(E) II only

Questions 14–15 refer to the following set of data:

A movie theater recorded the type of snack item purchased and the type of movie the customer was attending for each person in a random sample of theater customers who made a snack bar purchase. The resulting data is given in the table below.

	Soda	Popcorn	Candy
Children's Movie	70	83	47
Action Movie	61	49	20

14. If the manager wanted to determine if there is an association between the type of movie and type of snack sold, which type test would be appropriate?
(A) Chi-square goodness-of-fit
(B) Chi-square test for homogeneity
(C) Chi-square test for independence
(D) Two-sample proportions test
(E) Multiple sample proportions test

15. Using the data in the table above, the value of the chi-square statistic is 5.66. Which of the following is a correct statement about the P-value for this test?
(A) P-value > 0.10
(B) $0.05 < P$-value < 0.10
(C) $0.01 < P$-value < 0.05
(D) $0.001 < P$-value < 0.01
(E) P-value < 0.001

FREE-RESPONSE PROBLEMS

1. Regina is worried that the color of her new cardigan will attract the attention of killer bees in southern California where she is going to hike. To settle her nerves she looks at the American Killer Bee Association website. It shows that these bees are highly agitated by various colors. They have found that 75% of bees are agitated by green, 9% by blue, 6% by purple or pink, and the remaining 10% by other colors.
(a) In a random sample of 200 killer bees, how many would you expect to be agitated by each color?
(b) A recent study of 120 randomly selected people stung by killer bees last year found the individuals were wearing the colors shown in the table below. Do these data provide convincing evidence that the color distribution of colors worn by people stung by killer bees is different from the percentages given on the website?

Color Worn By Individual Stung By Killer Bees

Green	Blue	Purple/Pink	Other
86	21	6	7

2. A restaurant offers both dine-in and take-out service. Customers can pay for their meal in cash, by credit card or by debit card. The restaurant owner wonders if there is an association between the method of payment and the type of service. To investigate, a random sample is selected from the orders placed during the last year and the method of payment and the type of service is recorded for each of these orders. The data is summarized in the table below.

	Cash	**Credit**	**Debit**
Dine-in	34	122	32
Take-out	70	95	47

(a) Should the restaurant owner carry out a test of homogeneity or a test of independence to answer his question?

(b) Carry out a test to answer the question of interest to the owner. Use a significance level of 0.05 for your test.

Answers

Multiple-Choice Questions

1. **D.** The calculation would be found as follows:
$$\chi^2 = \frac{(60-50)^2}{50} + \frac{(80-50)^2}{50} + \frac{(30-50)^2}{50} + \frac{(10-50)^2}{50} = 2+18+18+32=70$$
(*Introduction to Statistics & Data Analysis* 3rd ed. pages 647–656/4th ed. pages 700–708).

2. **E.** The χ^2 distribution is right skewed with all of its area associated with positive values. The mean of a chi-square distribution can't be 0 (*Introduction to Statistics & Data Analysis* 3rd ed. pages 647–656/4th ed. pages 700–708).

3. **D.** There are four surgical methods and three pain levels. $df = (r-1)(c-1) = (4-1)(3-1) = 6$ (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

4. **D.** The expected counts in a two-way table are found using $\frac{(\text{row total} \times \text{column total})}{\text{table total}}$ (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

5. **C.** In a two-way table, degrees of freedom are calculated by $(r-1)(c-1) = (3-1)(6-1) = 10$ (*Introduction to Statistics & Data Analysis* 3rd ed. pages 667–671/4th ed. pages 711–722).

6. **C.** This is a test of homogeneity. Three populations are being compared on the basis of a categorical variable (color) (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

7. **E.** $\chi^2 = 22.9$ and $df = 10$, gives a P-value = 0.01 which is less than 0.05. This means there is strong evidence to conclude that the color distributions are not all the same (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

8. **E.** The expected count is calculated is 300(0.25) = 75 (*Introduction to Statistics & Data Analysis* 3rd ed. pages 647–656/4th ed. pages 700–708).

9. **B.** A test of independence between the variables would be used to answer the question of interest (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

10. **D.** (6 – 1)(8 – 1) = 5(7) = 35 (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

11. **B.** The *P*-value is smaller than the significance level, so the null hypothesis of independence would be rejected. This means there is an association and knowing people's political party provides information about their support of this issue (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

12. **B.** Because a single sample is classified on the basis of two categorical variables, the appropriate test is a test for independence (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

13. **E.** The degrees of freedom for a goodness-of-fit test is calculated using $k - 1$. However, for a chi-square test of independence, degrees of freedom is calculated using $(r - 1)(c - 1)$ (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

14. **C.** This is a test of independence since we are trying to see if knowing the type of movie provides any information about type of snack purchased (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

15. **B.** With $df = 2$, the *P*-value is 0.059 (*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).

FREE-RESPONSE PROBLEMS

1. (a) The expected counts by color for 200 bees would be:
   Green: $200(0.75) = 150$,
   Blue: $200(0.09) = 18$,
   Purple/Pink: $200(0.06) = 12$,
   Other: $200(0.10) = 20$.

   (b) Hypotheses
   H_0 : The proportions of bee stings by color is as specified on the web site ($p_{green} = 0.75$, $p_{blue} = 0.09$, $p_{purple/pink} = 0.06$, $p_{other} = 0.10$)
   H_a : At least one of the color proportions is different from what is specified inthe null hypothesis
   $\alpha = 0.05$
   Test: Chi-square goodness-of-fit test
   Assumptions
   1. The problem states that the individuals in the sample were randomly selected.
   2. The expected counts are shown in the table below (note that the sample size is 150, not 200 as in part (a). All expected counts are greater than 5, so the sample size is large enough.

Expected Counts for 150 Stings

Green	Blue	Purple/Pink	Other
(90.0)	(10.8)	(7.2)	(12.0)

Test Statistic

$\chi^2 = 12.09, \quad df = 3, \quad P = 0.007$

Conclusion

Since 0.007 < 0.05, there is strong evidence that to reject the null hypothesis. Therefore, there is reason to say the proportions of bee stings by killer bees are not all the same as reported on the website

(*Introduction to Statistics & Data Analysis* 3rd ed. pages 647–656/4th ed. pages 700–708).

2. (a) This is a test of independence. There was only one sample and each individual in the sample was classified according to two categorical variables.

(b)

Hypotheses

H_0 : method of payment and order type are independent.

H_a : method of payment and order type are not independent.

$\alpha = 0.05$

Assumptions

(1) The sample was a random sample of orders.

(2) The expected counts are given in the table below. All expected counts are greater than 5, so the sample size is large enough.

Expected Counts

	Cash	**Credit**	**Debit**
Dine-in	\$48.88	\$101.99	\$37.13
Take-out	\$55.12	\$115.01	\$41.87

Test Statistic

$$\chi^2 = \frac{(34-48.88)^2}{48.88} + \frac{(122-101.99)^2}{101.99} + \ldots = 17.291, \quad \text{with } df = 2$$

Conclusion

With a test statistic so extreme, our P-value = 0.000. This is smaller than any reasonable significance level, so we reject the null hypothesis. There is strong evidence that there is an association between method of payment and order type.

(*Introduction to Statistics & Data Analysis* 3rd ed. pages 660–671/4th ed. pages 711–722).