Chi-Square Tests

Math Objectives

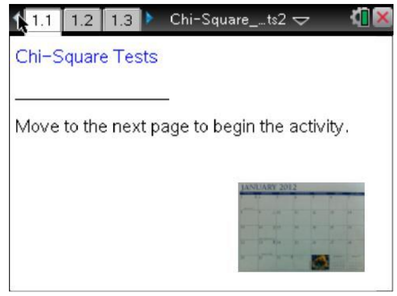
- Students will recognize that chi-squared tests are for counts of categorical data.
- Students will identify the appropriate chi-squared test to use for a given situation: X^2 Goodness of Fit Test, X^2 Test of Independence, or X^2 Test of Homogeneity.
- Students will learn how to calculate the degrees of freedom for each type of chi-squared test.
- Students will interpret the results of a chi-square test.
- Students will reason abstractly and quantitatively (CCSS Mathematical Practices).

Vocabulary

- alpha
- categorical data
- chi-squared (X^2)distribution
- degrees of freedom
- expected counts
- goodness-of-fit test
- observed counts
- *p*-value
- test of homogeneity
- test of independence

About the Lesson

- This lesson involves investigating chi-squared tests and distributions.
- As a result, students will:
  - Compare different scenarios and determine which chi-square test is appropriate.
  - Write the appropriate null and alternative hypotheses for the given scenario.
  - Determine the degrees of freedom for the chi-square test.
  - Look at chi-square test results and make the correct decision to reject or fail to reject the null hypothesis and write their conclusions in context.



TI-Nspire™ Technology Skills:

- Download a TI-Nspire document
- Open a document
- Move between pages

Tech Tips:

- Make sure the font size on your TI-Nspire handhelds is set to Medium.

Lesson Files:
Student Activity
chi_square_student.pdf
chi_square_student.doc
chi_square_Create.doc

TI-Nspire document
chi_square.tns

Visit www.mathnspired.com for lesson updates and tech tip videos.

TI-Nspire™ Navigator™ System

- Use Screen Capture to monitor student progress in creating the .tns file.
- Use Quick Poll to check for student understanding.

Prerequisite Knowledge

- Students should be familiar with hypothesis testing: writing null and alternative hypotheses, finding a test statistic and *p*-value, and making a decision based on the results.

Related Activities

- Statistics Nspired activity: Contingency Tables and Chi-square

Discussion Points and Possible Answers

> **Teacher Tip:** Students can create the *tns* file following the steps in the *Chi Square Create* document, or they can use the premade file C*hi_ Square.tns*.

Three different chi-squared tests will be discussed in this activity:

- χ^2 **Goodness-of-Fit (1)**: Compares sample counts (sometimes given as proportions) to expected counts based on a given population distribution.
- χ^2 **Two-way tables (2 & 3)**: There are two chi-squared tests using two-way tables—independence and homogeneity. The two tests differ in their hypotheses and conclusions but are mechanically identical. Determining which to use depends on how the data were collected.
  - **Test of Independence**: Compares two categorical variables in a *single population* to determine whether there is a significant association between the two variables.
  - **Test of homogeneity**: Compares categorical variables from *two or more different populations* to determine whether proportions are the same across different populations.

1. a. Suppose that in a typical week the number of absences from a large high school was 805. About how many would you expect per day? Explain your reasoning.

   Sample Answer: Some students might suggest that about the same number might be absent each day, so 805 divided by 5; others might think more students would be absent on Monday and Friday than on other days of the week.

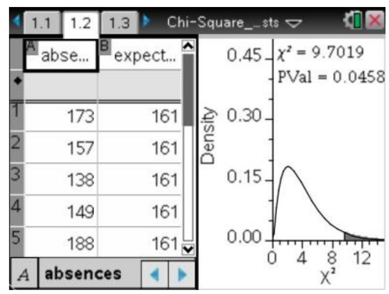
b. The school wants to see whether student absences are the same on different days of a randomly selected week of school. What type of hypothesis test should be used? Explain your answer.

Sample Answer: This is a chi-square goodness of fit test because the data are categorical—counts of absences on each day of the school week. We want to see if the sample absences fit a population pattern— is the number of students absent about the same each day?

c. Write the null and alternative hypotheses for this test.

Sample Answer: H_o: the proportions of absences are the same for each day of the week ($p_{monday}=p_{tuesday}=p_{wednesday}=p_{thrusday}=p_{friday} = 0.2$) H_a: at least one proportion is different

Move to page 1.2.



2. The left side of Page 1.2 shows the average number of observed absences per day of the week in Column A. Column B is the expected number of absences if the null hypothesis were true.
   a. How do the observed number of absences compare to your conjecture in question 1a?

Sample Answer: I was right; most of the absences seem to be on Friday, and then Monday. The fewest absences were on Wednesday.

b. How were the expected number of absences calculated, and what do they represent? Fill in the table with the values you found.

Sample Answers: The expected counts were found by taking the total of the observed absences (805) and dividing that by the typical number of days in the school week (5). The resulting answer (161) represents the number of absences that would be expected each day if the absences were the same for every day.

Day of week	Observed absences	Expected absences
Monday	173	161
Tuesday	157	161
Wednesday	138	161
Thursday	149	161
Friday	188	161
Total:	805	805

c. What are the conditions for this test? Are the conditions met?

Sample Answers: The conditions for a X^2 Goodness-of-Fit Test are that the sampling was random, less than 20% of the expected values are less than 5, and all of the expected values are greater than 1. Yes, the conditions are met.

d. The chi-square statistic is dependent on the degrees of freedom. The number of degrees of freedom for a X^2 Goodness of Fit test is found using the number of categories minus one. What degrees of freedom should be used in this situation?

Sample Answer: There are five categories –each day of the school week. So, 5 – 1 = 4 degrees of freedom.

3. The chi-square test statistic and the associated *p*-value appear on the right side of the page with the graph of the chi-square distribution.

Tech Tip: The test statistic and *p*-value are also available on the left side of the screen if the students scroll over to columns C and D.

Tech Tip: If students chose to display the graph on the same page with the data and Chi-square test results, they can later separate the page into two distinct pages using **Doc > Page Layout > Ungroup**.

a. Describe the graph.

Sample Answer: The curve is skewed to the right with the area above the χ^2 value of about 9.7 shaded.

b. Why is chi-squared always a positive value?

Sample Answer: The chi-squared statistic is found by summing the $\frac{(observed-expected)^2}{expected}$ values. Because the differences (observed – expected) are squared, the answer will always be positive.

c. What is the area of the shaded region? Explain your answer in the context of the problem.

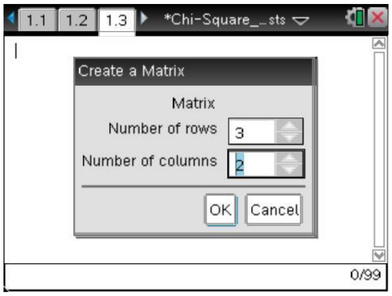
Sample Answer: The shaded region has an area of about 0.046. It represents the *p*-value—the probability of getting a chi-squared test statistic this extreme or greater if the absences are the same each day.

d. Make a decision to reject or fail to reject your null hypothesis using an alpha value of 0.01. Write your conclusion in context.

Sample Answer: I would fail to reject the null hypothesis because the *p*-value of 0.0457 is not less than the alpha value of 0.01. Based on the average weekly data, the evidence at the 0.01 level is not sufficient to suggest that student absences are different on different days of the week.

Teacher Tip: Point out to the students that at the .05 level, the null hypothesis would be rejected.

Move to page 1.3.



4. An advertiser who buys time on television suspects male and female viewers have different television viewing preferences. The company commissioned a survey of 100 males and 120 females asking their preferences among crime, reality, and comedy formats.
   a. Why would the advertiser care about such a difference?

Sample Answer: If a difference between viewing habits of shows and gender existed, the advertiser could tailor advertisements towards that specific audience.

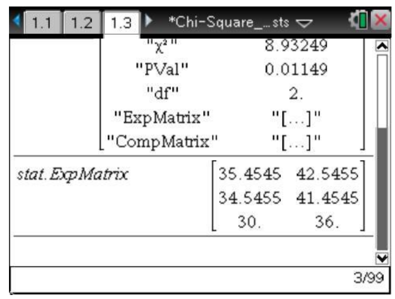
b. What type of hypothesis test should the advertiser use to analyze the results? Explain your reasoning.

Sample Answer: This example compares categorical variables from male/female populations to determine whether proportions of TV format preferences are the same across the two genders, so it uses a Chi-squared Test of Homogeneity.

c. Write the null and alternative hypotheses for this test.

Sample Answer: H_o: the three television formats have the same proportions of male and female viewers (p_{males}=$p_{females}$ for each format); H_a: the proportions are not the same.

5. a. The table below shows the survey results. Scroll down on Page 3.1 to find the expected counts calculated by the TI-Nspire. Fill the values and the totals into the table below.



Sample Answers:

Program Format	Males from survey	Males expected	Females from survey	Females expected	Totals
Crime	29	**35.4545**	49	**42.5455**	78
Reality	31	**34.5455**	45	**41.4545**	76
Comedy	40	**30**	26	**36**	66
Totals	**100**		**120**		**220**

b. How do you think the expected count for Males—Crime was calculated? Why does this make sense?

Sample Answer: 100 x (78/220) = 35.4545 – to get the answer on the handheld. If the null hypothesis is true, the expected count for males would have to give the same proportion for males as the total proportion for crime, which is 78/220 and there are 100 males so it would be 78/220 times the 100.

c. Explain what is meant by the expected count for the cell Males—Crime.

Sample Answer: The predicted number of men watching crime would be 35.4545 or about 35 if the proportions of men/women watching crime were the same.

d. What are the conditions for this hypothesis test and are the conditions met? Explain your answer.

Sample Answer: The conditions are that the sampling was random, each expected cell count should be greater than 1 and no more than 20% of them should be less than 5. Yes, all conditions are met.

e. Does it appear from the results of the survey that there is a difference in the viewing preferences of men and women? Explain your reasoning.

Sample Answer: It kind of does look like there is a difference, especially for crime and comedy shows. Fewer males liked crime than expected, and more females liked it. More males liked comedy than expected, and fewer women liked it.

f. The degrees of freedom for a X^2 two-way table is found using (# rows – 1)(# columns –1). What is the number of degrees of freedom for this test?

Sample Answer: (3 – 1)(2 – 1) = 2 degrees of freedom.

6. a. Interpret the results given on Page 1.3 for the X2 test.

Sample Answer: The X^2 test statistic of 8.93249 has a *p*-value of 0.01149, which means that a X^2 as large as or larger than 8.9 would occur by chance in about 1.1% of the samples.

b. Make a decision to reject or fail to reject your null hypothesis using an alpha value of 0.05. Explain your reasoning.

Sample Answer: Using an alpha value of 0.05, I would reject the null hypothesis because my decision is based on having a chi-square occur at least as great as 8.9 in less than 0.05 of the sample outcomes. The *p*-value for the survey results of 0.01 is less than 0.05, so having a chi-square of 8.9 is unlikely.

c. Write your conclusion in the context of the problem.

Sample Answer: There is strong evidence at the 0.05 level that males and females had different television viewing preferences.

Extension

You might want to have students explore the following two examples:

- A Goodness-of-Fit test in which the expected counts are not all the same to ensure they understand that the GOF test can be used in such situations.
- A Chi-Square test of independence, which would reinforce with students that both tests for two-way tables (Independence and Homogeneity) are conducted the same but the hypotheses and conclusions are different.

Wrap Up

Upon completion of the lesson, the teacher should ensure that students are able to understand:

- recognize that chi-squared tests are for counts or categorical data.
- identify the appropriate chi-squared test to use for a given situation: X^2 Goodness of Fit Test, X^2 Test of Independence, and X^2 Test of Homogeneity.
- can calculate the degrees of freedom for each type of chi-squared test.
- make a decision to reject or fail to reject the null hypothesis based on the *p*-value and write the conclusion in context.

Assessment

The following questions can be used as part of the lesson as a self-check for students or can be used as an assessment to determine how well students understand the concepts.

1. Choose the appropriate chi-squared test for each situation and explain your choice:

   a. A school wants to compare how its students did on the AP Statistics exam this year compared to the national scores.

   Answer: X^2 Goodness of Fit—the school's results are being compared to the total population (national scores)

   b. A restaurant samples customers to determine if there is a relationship between customer age and satisfaction with the restaurant's service.

   Answer: X^2 Test of Independence—one group of customers is surveyed and asked both their age and satisfaction with the restaurant's service.

c. A consumer safety organization wants to see if there is a difference in seat belt use in Los Angeles, California; Miami, Florida; and Dallas, Texas.

Answer: X^2 Test of Homogeneity—three samples are chosen, one from each city, and the sample proportions are compared.

d. A survey asked men and women how confident they were, on a scale from 1 to 5, that they could change a flat tire.

Answer: X^2 Test of Homogeneity because two samples are chosen, one of men and one of women, and the sample proportions in each category of the confidence scale are compared.

e. The proportion of each color of M&M's in a bag are compared to the color distribution that the manufacturer claims to make.

Answer: X^2 Goodness of Fit—the color distribution of M&M's in the bag is compared to the manufacturer's claim.

2. Decide whether the following statements are always, sometimes or never true. Explain your reasoning in each case.
   a. The X^2 curve is right-tailed.

   Sample Answer: Always true—the (observed – expected) values are squared to get the X^2 test statistic so it is always positive.

   b. A *p*-value is the probability of making a correct decision.

   Sample Answer: Never true—a *p*-value is the probability of getting a test statistic as extreme as the one we got if the null hypothesis is true.

   c. The number of degrees of freedom is n – 1 for X^2 tests, where n is the sample size.

   Sample Answer: Never true—the number of degrees of freedom is (# categories – 1) for the X^2 Goodness-of fit test and (# rows – 1)(# columns – 1) for X^2 two-way tables.

TI-Nspire Navigator

Note 1

Name of Feature: Live Presenter

Live Presenter might be used to have a student demonstrate how to build the .tns file.