Goodness-Of-Fit Tests
ID: 10286

Activity Overview
*In this activity, students test claims of whether given distributions "fit" theoretical distributions. The activity begins with an introduction to the goodness-of-fit test and a review of the chi-square distribution. Students then work through two problems, one in which the theoretical proportions of each category are the same and one in which they are not. Students will use spreadsheets to calculate test statistics and the **Integral** tool on the Graphs & Geometry application to calculate P-values. They check their calculations by using the chi square goodness-of-fit test provided in the Calculator application.*

Topic: Hypothesis Testing
- *Use a χ^2 test to test the hypothesis that an observed frequency distribution fits an expected frequency distribution.*

Teacher Preparation and Notes
- *This activity is designed to be used with students studying Statistics and Probability.*
- *Students should already be familiar with the basic concepts involved in hypothesis testing.*
- *The problems in this activity use both the critical value method and the P-value method of hypothesis testing.*
- *Notes for using the TI-Nspire™ Navigator™ System are included throughout the activity. The use of the Navigator System is not necessary for completion of this activity.*
- ***To download the student and solution TI-Nspire documents (.tns files) and student worksheet, go to education.ti.com/exchange and enter "10286" in the keyword search box.***

Associated Materials
- *Goodness-Of-Fit_Tests_Student.doc*
- *Goodness-Of-Fit_Tests.tns*
- *Goodness-Of-Fit_Tests_Soln.tns*

Suggested Related Activities
To download any activity listed, go to education.ti.com/exchange and enter the number in the keyword search box.
- *Candy Pieces (TI-Nspire technology) — 9997*
- *Chi-Square Distributions (TI-Nspire technology) — 9738*
- *Testing Goodness-of-Fit and Two-Way Table Chi-Square (TI-84 Plus family) — 4590*

Problem 1 – The Test Statistic

Use pages 1.2 through 1.3 to introduce the goodness-of-fit test and the test statistic used.
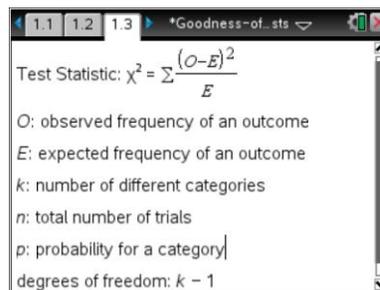
Students may recognize the positively skewed chi-square distribution from earlier lessons on estimating population variance and standard deviation.

They can change the degrees of freedom and observe how the distribution changes.
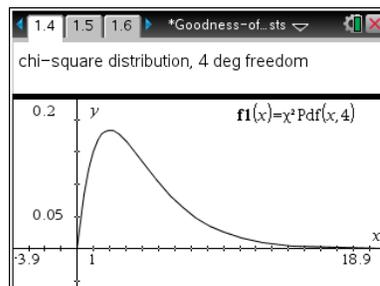
Have students discuss the answers to the questions on pages 1.5, 1.6, and 1.7. Based on the formula for χ^2, the test statistic will be large when the observed and expected values are very different and small when the differences are small.

A small χ^2 would suggest that the values are a good fit with the theoretical distribution.

All goodness-of-fit tests are right-tailed because the leftmost value, 0, occurs when there are no differences.
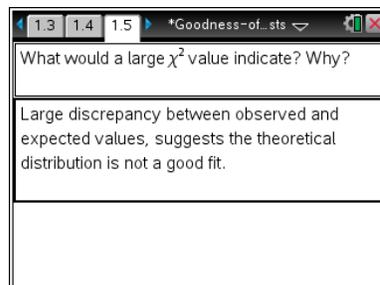
Problem 2 – Same Proportion Throughout

The scenario presented on page 2.1 tests a claim that every category has the same proportion.

Tell students to list H_0 and H_A on their worksheets.

H_0: $p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = p_7$

H_A: At least one of the proportions is different than the others.
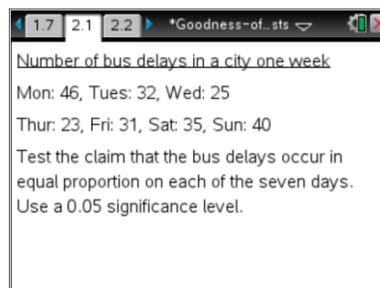
After students read the scenario, they are to move to page 2.2 to find E, the expected value for each day. Assuming the null hypothesis is correct, the probability of a delay on any given day is $\frac{1}{7}n$. The value of n is the total number of trials, the sum of the delays.
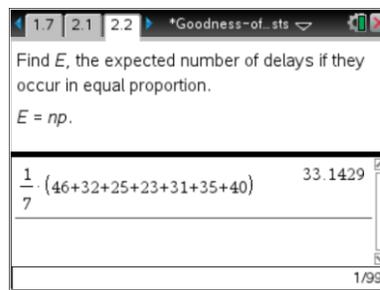
TI-Nspire Navigator Opportunity: *Live Presenter*

See Note 1 at the end of this lesson.

Students need to list this value, the expected value, in cell B1 on page 2.3. Students are to copy the formula down through cell B7 by pressing **MENU > Data > Fill** and clicking the down arrow key until they reach cell B7 where they press enter.
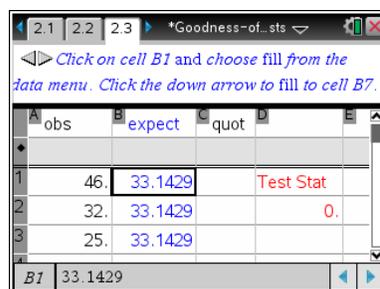
The observed values are already listed in Column A, the $\frac{(O-E)^2}{E}$ quotient for each category will be found in Column C and the test statistic, the sum of the quotients, will be found in cell D2.

In cell C1, students need to type $=\frac{(A1-B1)^2}{B1}$. When they press enter, the value will appear in the cell; the formula will appear at the bottom of the page.
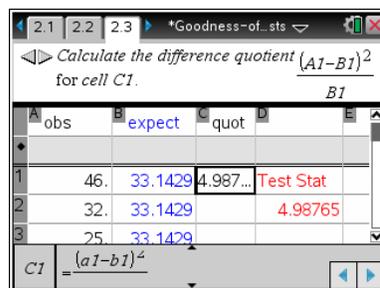
Students are to copy the formula down through cell C7 by pressing **MENU > Data > Fill** and clicking the down arrow key until they reach cell C7 where they press enter. Students can scroll through each value and see that the formula was computed for each observed value.
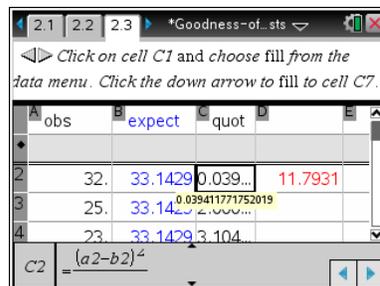
Students can then observe the test statistic in cell D2.

Have students advance to page 2.4 and find the critical value for the chi-square value with 0.95 of the area to the left (**MENU > Statistics > Distributions > Inverse χ^2**).

Remind students that the number of degrees of freedom is one less than the number of categories, so it is 6 for this problem.

The critical value is 12.5916. Ask students if they should reject or fail to reject the null hypothesis and why.

(Fail to reject; if the test statistic is less than the critical value, it is not to the right of the critical value (not in the critical region).)
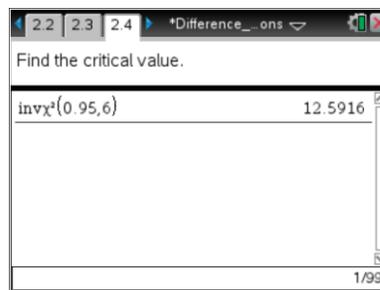
TI-Nspire Navigator Opportunity: *Quick Poll*

See Note 2 at the end of this lesson.

Students will use page 2.5 to find the *P*-values. First, they need to change the *x*-value of the plotted point to the test statistic, 11.7941, by double clicking on the *x*-coordinate value and entering 11.7941
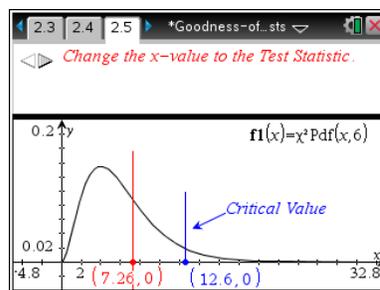
Then students need to use the **Integral** tool (**MENU > Analyze Graph > Integral**) to find the area to the right of the test statistic. Once the Integral tool is highlighted, they should select the leftmost value (the test statistic) and then the rightmost value, which for practical purposes can be around *x* = 32.
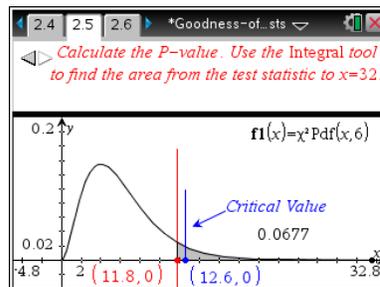
The *P*-value is about 0.066712 (to observe the *P*-value more accurately, have students move the cursor to the calculated value and press ⊞ a few times.). Have students discuss what the *P*-value means and why it confirms their decision based on the critical value. (*P*-value is greater than 0.05.) Students should also answer the question on page 2.6.

This is an excellent opportunity to visually illustrate why we fail to reject the null hypothesis (the critical value is inside the critical region.)

TI-Nspire Navigator Opportunity: *Live Presenter*

See Note 3 at the end of this lesson.

Students can verify their calculation on page 2.7 by pressing **MENU > Statistics > Stat Tests >** χ^2 **GOF**. They should select **obs** for the Observed List, **expect** for the Expected List and **6** for Deg of Freedom, *df*.
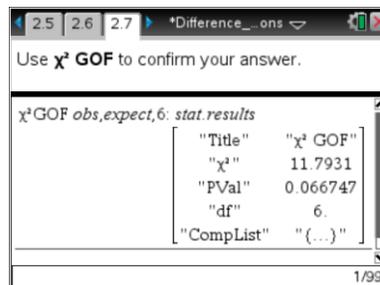
Problem 3 – Different Proportions Throughout

Unlike Problem 2, the expected values in Problem 3 will not all be the same. Students are to list H_0 and H_A on their worksheet.

H_0: $p_{fruit} = 0.35$, $p_{nuts} = 0.25$, $p_{choc} = 0.2$, $p_{seeds} = 0.2$

H_A: At least one of the proportions is different than what is listed above.

On page 3.2, the sum of the observed values listed is 450. Students are to use this and the percents to calculate the expected values in Column C. For example, in cell C1, they should type **= 0.35 * 450**.
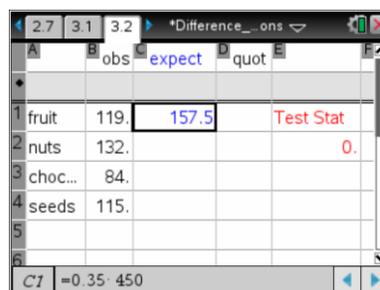
Students can then find each $\dfrac{(O-E)^2}{E}$ quotient and the sum of the quotients as they did in Problem 2.

The test statistic equals 20.1356.
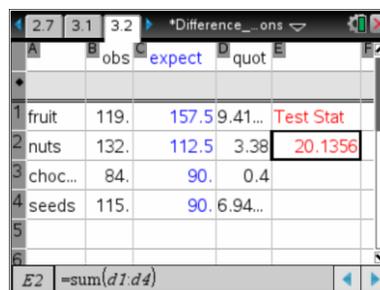
Students are to move to page 3.3 and determine the critical value. Ask students if they should accept or reject the null hypothesis and why.

(Reject; the test statistic is greater than the critical value, it is to the right of the critical value (in the critical region).)
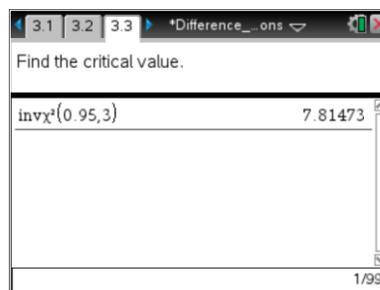
Have students find the corresponding *P*-value on page 3.4 and interpret its meaning.

(*P*-value = 0.000159)

They should also answer the question on page 3.5 about why they should reject the null hypothesis. (Because the *P*-value is much less than 0.05.)

Students can then move to page 3.6 and verify their calculations by using the **chi-square goodness-of-fit** command in the *Calculator* application.
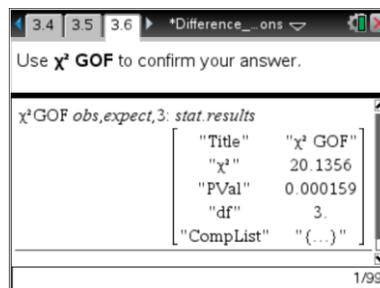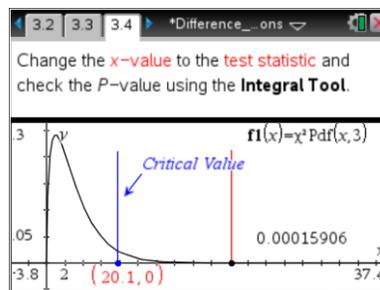
TI-Nspire™ Navigator™ Opportunities

Note 1
Question 1, *Live Presenter*
For page 2.3, use Live Presenter to help students see how to click the slider for directions, enter and fill down the data and calculated values. Also use Live Presenter to illustrate that the test statistic is the sum of the quotients in Column C by clicking on the test statistic value and displaying the formula.

Note 2
Question 2, *Quick Poll*
Send students a Quick Poll asking if they reject or fail to reject the null hypothesis. For those answering incorrectly (fail to reject) inform the students that the next page will further illustrate the correct answer.

Note 3
Question 3, *Live Presenter*
Use Live Presenter to demonstrate how to change the *x*-value to the critical test statistic as well as how to use the integral tool for finding the *P*-Value. Also use this page to help those students who incorrectly answered the Quick Poll why we fail to reject the null hypothesis - the critical value is INSIDE the critical region.