# Goodness-Of-Fit Tests

**Student Activity**

**Name** _____

**Class** _____

A **goodness of fit test** tests the hypothesis that observed frequencies in different categories fit a theoretical distribution. The test statistic is $\chi^2$.

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

*O*: observed frequency of an outcome

*E*: expected frequency of an outcome

*k*: number of different categories

*n*: total number of trials

*p*: probability for a category

degrees of freedom: $k - 1$

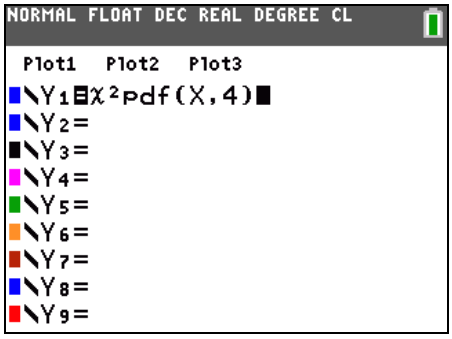
**Problem 1 – The Test Statistic**

Press WINDOW and set the values equal to the following.

Xmin = –4 Xmax = 20 Xscl = 1

Ymin = –0.1 Ymax = 0.25 Yscl = 1

Then press Y= and enter $\chi^2$ **pdf (X, 4)** in **Y1**. The $\chi^2$ **pdf** command is found by pressing 2nd [DISTR].




Press ENTER. This is the chi-square distribution with 4 degrees of freedom. Study the formula for $\chi^2$ to answer the following.

**1.** What would a large $\chi^2$ value indicate? Why?

**2.** What would a small $\chi^2$ value indicate? Why?

**3.** Would a goodness-of-fit test be left-, right-, or two-tailed? Why?

**Problem 2 – Same Proportions Throughout**

Number of bus delays in a city in one week:

| Mon | Tues | Wed | Thur | Fri | Sat | Sun |
|---|---|---|---|---|---|---|
| 46 | 32 | 25 | 23 | 31 | 35 | 40 |

Test the claim that the bus delays occur in equal proportion on each of the seven days at the 0.05 significance level.

**4.** Write the null and alternative hypotheses.

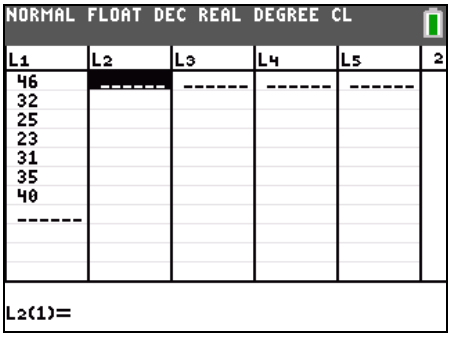
**5.** Find *n*, the total number of trials (the sum of the delays).

**6.** What is *p*, the probability of a delay on any given day, assuming the null hypothesis is correct?

**7.** Find *E*, the expected number of delays on a given day, if they occur in equal proportion, given that $E = np$. Round to three decimal places.

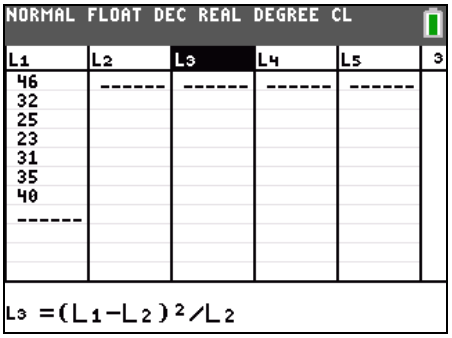
Now that you know *E* for each day (found in Question 7), and *O* for each day (in the chart), you will find the value of the test statistic, $\chi^2$.

Press [STAT] [ENTER]. Enter the observed frequencies in **L1**. Enter the expected frequency for each day in **L2**.



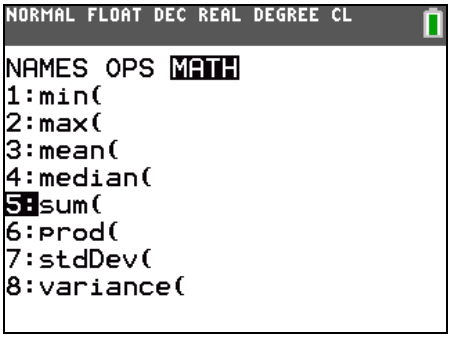

Move the cursor to the top of the third column (over the title of **L3**) and type the expression for chi-square using **L1** for *O* and **L2** for *E*.

Press [ENTER]. The seven individual quotients are shown in the list.

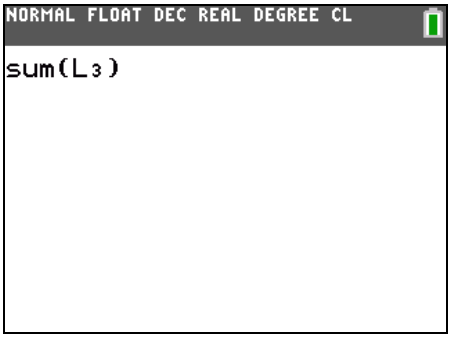



Exit out of the List Editor by pressing [2nd] [QUIT].

Press [2nd] [LIST]. Move to the **MATH** menu and choose **sum(**.




Enter **L3**, then press [ENTER]. The value shown is the chi-square test statistic.



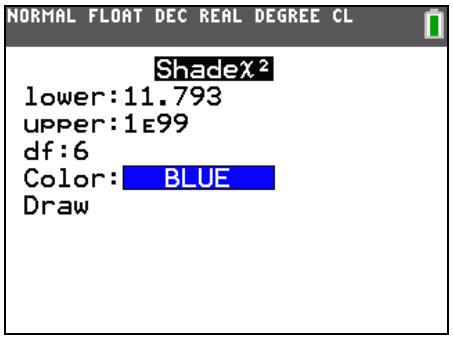

**8.** What is the value of the chi-square test statistic?

**9.** Use a Chi-Square distribution chart to find the critical value for this test.

**10.** Do you reject or fail to reject the null hypothesis? Why?

Press [Y=] and clear the entry for Y1.

The *P*-value can be found by using the Shade$\chi^2$ command. It is located in the **DRAW** menu which is accessed by pressing [2nd] [DISTR]. The format is *lower bound, upper bound, number of degrees of freedom*. Use 1E99 for positive infinity. (E is entered by pressing [2nd] [EE].)

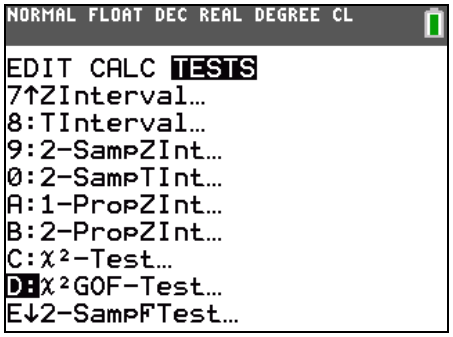



**11.** What is the area to the right of the test statistic?

**12.** What is the *P*-value?

**13.** Explain how this value confirms your decision to reject or fail to reject the hypothesis.

To check your work, press [STAT], choose **TESTS**, and scroll down to $\chi^2$ **GOF-Test**. Keep the default of **L1** for the Observed and **L2** for the Expected. Update **df**, the number of degrees of freedom to 6. Press [ENTER] when the cursor is on **Calculate**.



**Problem 3 – Different Proportions Throughout**

Test if the observed values fit the claimed percents at the 0.05 significance level.

| | Dried Fruit | Nuts | Chocolate | Seeds |
|---|---|---|---|---|
| Claimed Percent | 35% | 25% | 20% | 20% |
| Observed values (g) | 119 | 132 | 84 | 115 |

**14.** Write the null and alternative hypotheses.

**15.** What is *n*, the sum of the observed values?

**16.** Use the answer to Question 15 and the percents in the chart to find the expected values.

**17.** Enter the observed values in **L1** and the expected values in **L2**. Create **L3** by writing the formula at the top of **L3** and using the sum command. What is the test statistic?

**18.** Use your chart to find the critical value for this test.

**19.** Do you reject or fail to reject the null hypothesis? Why?

**20.** Use the Shade$\chi^2$ command to find the *P*-value.

Check your work by using the $\chi^2$ **GOF-Test**.