NUMB3RS Activity: Candy Pieces
Episode: “End of Watch”

Topic: Chi-square test for goodness-of-fit **Grade Level:** 11 - 12
Objective: Use a chi-square test to determine if there is a significant difference between the proportions observed in sample data and the theoretical proportions expected for that sample.
Materials: TI-83 Plus/TI-84 Plus graphing calculator
Time: 20 - 30 minutes

Introduction
In “End of Watch,” the FBI identifies a suspect for the murder of a police officer. When they go to question him, they find that he has been drugged and killed. Charlie believes that from the specific chemical makeup of the drug, he can use a statistical identification method to trace it back to a specific drug dealer who may be involved in the crimes. While the method is not named in the dialog, the chalkboard work in the scene is similar to this activity.

This activity introduces the statistical concept of hypothesis testing. Students are given information on the number of pieces by color in a bag of candy. They are asked whether they think that the bag could have come from a manufacturing process designed to produce equal proportions of each color. Students will then use a chi-square (pronounced “kī”) test for goodness-of-fit to determine if there is really a significant difference between the proportions they find in the sample and the proportions they would expect if the manufacturer produced equal proportions of each color.

Discuss with Students
Tell students that this activity is only an informal introduction to the concept of hypothesis testing, a major topic in statistical inference. Hypothesis tests work in a manner similar to a jury in a criminal trial. In the American system of justice, a defendant is presumed innocent. Only if the evidence convinces the jury beyond any reasonable doubt does the jury decide that the defendant is guilty. A jury never finds a defendant innocent—just guilty or not guilty. Statisticians use hypothesis tests to make inferences about a population based on random samples. With hypothesis tests, statisticians determine whether there is enough evidence to reject the hypothesis that the difference can be due to chance, or decide there is not enough evidence to reject it.

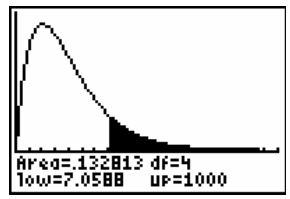
Calculating chi-square is one of many procedures statisticians use for testing hypotheses. The starting hypothesis (called a null hypothesis) in this case assumes there is no difference in the distribution of colors. The data are inspected to see if they support this assumption and chi-square is the statistical tool used to make a decision. The alternate hypothesis is that the proportions of colors are not all equal. In this brief activity, it is not possible to cover all of the nuances of a chi-square test. In courses like AP® Statistics, students learn that three assumptions need to be checked before using a chi-square test for goodness-of-fit: the samples need to be chosen randomly, the samples need to be independent, and the sample should be large enough so that the expected values should each be at least 5. For example, if you toss a 6-sided die 12 times, the expected numbers for each outcome would be 2; if you tossed the die 60 times, the expect outcome for each face would be 10).

In Question #2, some students will realize that the sum of the differences is 0 because the sum of the observed values and the sum of the expected values are equal: both 85. In Question #3,

John F. Mahoney, Benjamin Banneker Academic High School, Washington, DC

one way of explaining the use of $\frac{(O-E)^2}{E}$ is that it makes the values positive and gives a way of measuring the relative sizes of differences. In Question #4, the question is whether the calculated chi-square value is large or small. The value could range from 0 (if the observed numbers were the same) to a very large number if the observed values were all one color. A chi-square distribution is the theoretical distribution of the probabilities that a chi-square is greater than or equal to any given value. The question to consider is where does the chi-square calculated from the sample data fit in this distribution? The chi-square cumulative distribution function on the TI 83 Plus/TI 84 Plus graphing calculator can be used to determine the probability that a sample chi-square would be as large or larger than one that occurred by chance.

The shape of the distribution depends on the number of categories under consideration which can affect the size of the chi-square. If a chi-square value is way out in the tail of the distribution, it is not likely to have occurred by chance, and it may be appropriate to reject the hypothesis that there is no difference. The graph of the chi-square distribution with 4 degrees of freedom is displayed.



A degree of freedom can be compared to the number of free choices there are before the variables are determined by the situation. If the sum of four numbers is 8, three of them can be any value, but the fourth number is determined. This means there are 3 degrees of freedom. The area of the shaded region between the graph and the x-axis from 7.0588 to the right is the probability that a chi-square is greater than 7.0588 by chance.

Students are instructed to enter the command **Shadeχ^2(ans, 1000, 4)**. Note that "chi" (χ) is a Greek letter. The first parameter in the command is the chi-square test statistic computed in Question #3. The second parameter is a large number so students can find the probability that a chi-square value would fall between their chi-square and that large number by chance. The final parameter is the number of degrees of freedom. In this activity it is 4. Since there are 85 candies and 5 colors, if the number of candies for four of the colors is known, the number for the fifth color is determined because the total has to sum to 85.

Actual bags of candy can be used for problems 6 and 7.

Student Page Answers:
1. *There are 85 pieces of candy. The expected number is 17 pieces of each color. Answers will vary for the third question.* **2.** *The sum of the differences is 0, and thus this value doesn't appropriately represent the total difference*. **3.** *Yellow: 2.1176, Red: 0.2353, Blue: 3.7647, Orange: 0, Green: 0.9412; sum = 7.0588.* **4.** *The p-value is 0.1328* **5.** *Since the p-value is greater than 0.05 there is not enough evidence to reject the hypothesis that the colors were manufactured in equal numbers.* **6.** *Chi-square of 16.729, p-value 0.00504. Since the p-value is less than 0.05 there is enough evidence to reject the hypothesis that the colors were manufactured in equal numbers.* **7.** *Chi-square of 2.407, p-value 0.7904. Since the p-value is greater than 0.05 there is not enough evidence to suggest that the bag of candy did not come from a process that produced the stated proportion of colors.*

John F. Mahoney, Benjamin Banneker Academic High School, Washington, DC

Name: ___________________________________ Date: ________________

NUMB3RS Activity: Candy Land

In “End of Watch,” the FBI identifies a suspect for the murder of a police officer. When they go to question him, they find that he has been drugged and killed. Charlie believes that from the specific chemical makeup of the drug, he can use a statistical identification method to trace it back to a specific drug dealer who may be involved in the crimes. In any situation, the outcomes (like the proportions of ingredients in a sample of heroin) will vary due to chance. If you toss a coin 1,000 times, you would expect to get somewhere around 500 heads. The *chi-square test for goodness-of-fit* can be used to determine if there is a significant difference between the proportions you would expect to get in a sample and the proportions you actually get. This activity contains an informal introduction to the mechanics of this test.

Suppose a certain popular brand of candy pieces comes in five colors. A student counted the number of pieces of each color in a bag and found the results shown in the table below.

Based on these data, is it likely that this bag of candy came from a manufacturing process that was designed to produce equal proportions of each color? One way to answer this question is to perform a statistical procedure called a chi-square test for goodness-of-fit, checking to see how well the sample distribution fits the theoretical distribution.

1. How many pieces of candy were in the bag? If the proportions of each color were the same, how many pieces of each color would you expect to find? These values are called the *expected values*. The actual counts of each color of candy in the table are called the *observed values.* Complete the next two columns of the table below. Do you think that there is a considerable difference between the observed values and the expected values?

Color	**Observed**	**Expected**	**Observed – Expected**	$\frac{(O-E)^2}{E}$
Yellow	11			
Red	19			
Blue	25			
Orange	17			
Green	13			

2. This chi-square test involves quantifying the extent of the difference between the observed and expected values for each of the colors. Does it make sense to find the sum of the differences (Observed – Expected) to describe the total difference? Why or why not?

3. For each color, compute “(Observed– Expected)^2 / Expected” and enter this value in the last column. Find the sum of these five values. This is called the *chi-square*, represented by χ^2.

The question now is whether the chi-square found is large or small. To find out, you can compare your observed chi-square to the theoretical distribution of the probabilities that a chi-square would be greater than or equal to any given value. The chi-square density function can be graphed on a graphing calculator. Set the calculator's window to: **Xmin = 0**, **Xmax = 20**, **Xscl = 1**, **Ymin = –0.05**, **Ymax = 0.2**, **Yscl = 1**, **Xres = 1**. Press [2nd] [DISTR] to get the distribution menu. Scroll to the right to select **DRAW**, choose **Shade**χ^2**(**, and enter the following: **Shade**χ^2**(Ans, 1000, 4)**. “Ans” is the chi-square computed in Question #3. The number 1,000 is used as a large number, so you can find the probability that a chi-square would fall between

John F. Mahoney, Benjamin Banneker Academic High School, Washington, DC

your value and 1,000 by chance. The final number in the command is the number of independent variables in the problem, called the degrees of freedom. In this example it is 4. Since there are 85 candies and 5 colors, the number of candies possible for the fifth color is determined by the number of candies for each of the other four colors. That is, there are only four independent variables; the last color depends on how many of the other four there are. The “**Shade**χ^2” command computes the area of the region under the graph from the chi-square value to, in this case, 1,000. This area is called the *p*-value, the probability that a chi-square would be as large as or larger than the observed value simply by chance.

4. What is the *p*-value for the bag of candy?

A chi-square of 7.77 would yield a *p*-value of 0.10. This value means that if the colors of the candy were in fact distributed equally, a chi-square of 7.77 or higher would occur by chance about 10% of the time. Something that happens 10% of the time is not considered too unusual, so a statistician would conclude there is not enough evidence to state that the bags of candy did not come from a process that produced equal numbers of the colors. Statisticians usually make the same judgment if the *p*-value is more than 5%. If the *p*-value is less than 5%, a statistician would conclude that there is sufficient evidence to reject the assumption that the bags of candy came from a process that produced equal numbers of the colors.

5. Based on the answer to Question #4, is there sufficient evidence to reject the hypothesis that the bags of candy came from a process that produced equal numbers of the colors? Why or why not?

6. Another student opened a bag of a different brand of candy, counted the number of pieces of each color and found the results shown in the table below.

Color	**Observed**	**Expected**	$\frac{(O-E)^2}{E}$
Brown	15		
Yellow	14		
Red	16		
Blue	35		
Orange	29		
Green	24		

Repeat Questions #1–5, making any necessary modifications, and use a chi-square test to determine if it is likely that this bag of candy came from a manufacturing process that was designed to produce equal numbers of each color.

7. A third student looked at the manufacturer's Web site for this second brand of candy and saw the claim that the candy is made in the proportions: Brown 13%, Yellow 14%, Red 13%, Blue 24%, Orange 20%, and Green 16%. Because of the distribution of colors in this bag, she was suspicious of the accuracy of the manufacturer’s claim. Repeat Questions #1–5 and use the chi-square test to determine if it is likely that this bag of candy came from a manufacturing process that was designed to produce these proportions of each color.

John F. Mahoney, Benjamin Banneker Academic High School, Washington, DC

The goal of this activity is to give your students a short and simple snapshot into a very extensive mathematical topic. TI and NCTM encourage you and your students to learn more about this topic using the extensions provided below and through your own independent research.

Extensions

For the Student

1. Explore the color distribution of various brands of candy, such as Kissables®, varieties of M&Ms®, Skittles®, Reese's Pieces®, etc., using the techniques of this activity. Is there enough evidence to accept that the distribution of colors is uniform?

2. The administrative guidelines for the distribution of grades for a certain high school are: As 10%, Bs 20%, Cs 40%, Ds 20%, and Fs 10%. A mathematics teacher assigned the following grades to students in her statistics classes: 15 As, 25 Bs, 25 Cs, 15 Ds and 6 Fs. Her principal argued that the teacher did not follow the guidelines. Provide statistical evidence to support or refute the principal's statement.

3. Survey at least 20 students and ask each which of these colors is their favorite: blue, red, green, or yellow. You might expect the distribution of preferences to be equal; however, it usually is not and this can be shown using a chi-square test.

Additional Resources

1. The M&M® website gives counts of colors of various products: **http://us.mms.com/us/about/products**

2. For a full explanation of the chi-square goodness-of-fit test and a series of simulation exercises, see: **http://www.math.uah.edu/stat/hypothesis/ChiSquare.xhtml**

3. AP® Statistics textbooks each include sections on the chi-square goodness of fit test. A few of these are:

- Yates, D., Moore, D., & Starnes, D. (2003) The Practice of Statistics. New York: W. H. Freeman and Company.
- Bock, D., Velleman, P., & DeVeaux, R. (2004) Stats: Modeling the World. Boston: Pearson Education, Inc.
- Peck, R., Olsen, C., & Devore, J. (2005) Introduction to Statistics and Data Analysis, Second edition. Belmont: Thomson, Brooks/Cole.

John F. Mahoney, Benjamin Banneker Academic High School, Washington, DC