**In-Class Example of Random Sampling and Testing Hypotheses with Statistics**

Center for Microbial Oceanography: Research and Education (2013). C-MORE Science Kits: Random Sampling [Educational materials]. Retrieved from <http://cmore.soest.hawaii.edu/education/teachers/science_kits/materials/Random_Sampling/Random_Sampling_Binder_Materials.pdf>

C-MORE Science Kits: <http://cmore.soest.hawaii.edu/education/teachers/science_kits.htm>



Grade Level: This kit is appropriate for students in grades 6–12.

Standards: This kit is aligned with state science and math content standards for Hawai‘i, California and Oregon, as well as national Ocean Literacy Principles.

Overview: This three-lesson kit introduces random sampling, one of the key concepts employed by scientists to study the natural environment. In Lesson 1, concepts of random sampling are introduced. Working in groups, students randomly selects samples from a population and record the data. The students then compare the composition of the samples they obtained to that of the entire population. In Lesson 2, technology is introduced. Students enter and graph data from Lesson 1 using Excel (a Microsoft Office program). In Lesson 3, statistics provide a means for students to assess how well their samples represent the total population.

Lessons 1 and 2 are suitable for grades 6–12, and Lesson 3 is geared toward grades 9–12. All of the supplies in this kit are in support of Lesson 1. Computers (not provided) are required for Lesson 2 and may be optionally used for Lesson 3. If computers are not available, Lesson 2 may be omitted, and Lesson 3 can be taught directly following Lesson 1. Each of the three lessons requires approximately 50 minutes of class time.

Suggestions for Curriculum Placement: These lessons can be used to introduce the scientific method or parts of the process at the beginning of the year. Skills such as forming and revising hypotheses, collecting meaningful data, setting up data tables, and graphing and analyzing results are taught through this exercise. Sampling and probability are important concepts in all fields of science, including ecology, genetics, and oceanography. The use of the Excel program to create data tables and graphs provides a means of incorporating technology into the science classroom, and gives the students exposure to a spreadsheet program that is commonly used in upper level science classes.

**Hawai‘i Content & Performance Standards (HCPS III)**

**Science Standard 1: The Scientific Process: SCIENTIFIC INVESTIGATION: Discover, invent, and investigate using the skills necessary to engage in the scientific process.**

Grades 6–8 Benchmarks for Science:

SC.6.1.1 Formulate a testable hypothesis that can be answered through a controlled experiment.

SC.6.1.2 Use appropriate tools, equipment, and techniques to safely collect, display, and analyze data.

SC.7.1.2 Explain the importance of replicable trials.

SC.7.1.3 Explain the need to revise conclusions and explanations based on new scientific evidence.

SC.8.1.1 Determine the link(s) between evidence and the conclusions(s) of an investigation.

SC.8.1.2 Communicate the significant components of the experimental design and results of a scientific investigation.

Grades 9–12 Benchmarks for Physical Science, Biological Science & Earth and Space Sciences:

SC.PS/BS/ES.1.1 Describe how a testable hypothesis may need to be revised to guide a scientific investigation.

SC.PS/BS/ES.1.3 Defend and support conclusions, explanations, and arguments based on logic, scientific knowledge, and evidence from data.

SC.PS/BS/ES.1.4 Determine the connection(s) among hypotheses, scientific evidence, and conclusions.

**Science Standard 2: The Scientific Process: NATURE OF SCIENCE: Understand that science, technology, and society are interrelated.**

Grades 6–8 Benchmarks for Science:

SC.8.2.2 Describe how scale and mathematical models can be used to support and explain scientific data.

**Science Standard 3: Life and Environmental Sciences: ORGANISMS AND THE ENVIRONMENT: Understand the unity, diversity, and interrelationships of organisms, including their relationship to cycles of matter and energy in the environment.**

Grades 6–8 Benchmarks for Science:

SC.7.3.2 Explain the interaction and dependence of organisms on one another.

**Math Standard 11: Data Analysis, Statistics, and Probability: FLUENCY WITH DATA: Pose questions and collect, organize, and represent data to answer those questions.**

Grades 9–12 Benchmarks for Statistics:

MA.S.11.1 Develop a hypothesis for an investigation or experiment.

MA.S.11.3 Select appropriate display for a data set (e.g., frequency table, histogram, line graph, bar graph, stem-and-leaf plot, box-and-whisker plot, scatter plot).

MA.S.11.5 Recognize sampling, randomness, bias, and sampling size in data collection and interpretation.

MA.S.11.6 Describe the purpose and function of a variety of data collection methods (e.g., census, sample surveys, experiment, observation).

**Lesson 1: Introduction to Random Sampling**

**Overview**:

This lesson introduces random sampling, one of the key concepts employed by scientists to study the natural environment.

M&M’s in a bag are then used to represent different samples, with the bag itself representing the population. Students collect random samples of M&M’s from the population and record the data.

To learn about the inherent variability of random sampling, students then compare the composition of their individual samples, their group’s pooled sample data, and that of the entire population.

**Random Sampling Lesson**

Random Sampling

1. In scientific research, why do we collect samples?
   1. It may not be practical or possible to count the entire population.
   2. For example, if we are interested in the different kinds of fish species in an area of tide pools, or counting the number of Native vs. Invasive species of plants in an area, it is unlikely that we would be able to count all of the fish or all of the plants. Therefore, we take samples to estimate the population.
2. Why is it important that we collect the samples randomly?
   1. This eliminates potential bias.
   2. For example, if you collect a sample from only one spot, other factors could be influencing your findings.

M&M’s Activity

Let’s imagine that we are interested in studying the prevalence of different colors of M&M’s.

We can imagine that the different colors of M&M’s represent different kinds of endemic or native fish in a set of tide pools or different kinds of native plants in an area.

From this activity, a simple research question might be: “What color of M&M’s is most prevalent?”

Environmental scientists often study the effects of invasive species on the native species, such as native fish or native plants. Therefore, let’s imagine that I am an invasive species who loves to eat M&M’s. However, I’m particular to only certain colors of M&M’s.

How can we rephrase our research question: “Does the frequency of M&M’s differ by colors?” “Does the abundance of M&M’s differ in the population by color?”

How can we determine what color of M&M’s I like to eat? We can collect random samples of M&M’s from the population to make inferences about the frequency of different colors of M&M’s in the population. From there, we can apply inferential statistics to determine if there is a significant difference in the frequency of the different colors of M&M’s.

Collect sample of 10 M&M’s from the population and record the number for each color here.

\_\_\_\_\_: Red

\_\_\_\_\_: Orange

\_\_\_\_\_: Yellow

\_\_\_\_\_: Green

\_\_\_\_\_: Blue

\_\_\_\_\_: Brown

According to the sample data you collected, what is the most abundant color of M&M’s? What is the least abundant color?

Most frequent color M&M: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Percentage = \_\_\_\_%

Least frequent color M&M: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Percentage = \_\_\_\_%

HYPOTHESIS: The least abundant type of M&M in the bag (population) is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

Summarize group data collected.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sample** | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** | **Sample Total** |
| 1 |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |
| 11 |  |  |  |  |  |  |  |
| 12 |  |  |  |  |  |  |  |
| 13 |  |  |  |  |  |  |  |
| 14 |  |  |  |  |  |  |  |
| 15 |  |  |  |  |  |  |  |
| 16 |  |  |  |  |  |  |  |
| 17 |  |  |  |  |  |  |  |
| 18 |  |  |  |  |  |  |  |
| 19 |  |  |  |  |  |  |  |
| 20 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **M&M Color** | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** | **Sample Total** |
| **Group Total** |  |  |  |  |  |  |  |
| **Group Percentage** |  |  |  |  |  |  |  |

1. According to the group data collected, what is the most abundant color of M&M’s? What is the least abundant color?
   1. Most frequent color M&M: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Percentage = \_\_\_\_%
   2. Least frequent color M&M: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Percentage = \_\_\_\_%
2. REVISED HYPOTHESIS: Based on the group data collected, should you change your hypothesis? If so, write your revised hypothesis in the space below. If not, explain why you chose to retain your hypothesis.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_.

1. How did the data from the sample that you personally collected compare with the data collected by the entire group? Is this what you would expect? Why?
2. How did you ensure that sampling was random and not biased toward collecting any one particular color of M&M?
   1. What would happen if you ate the M&M’s and didn’t put them back in the bag?
   2. Does it matter if the same M&M’s were sampled more than once?
3. Can you think of other examples, research questions, or issues that this example can be applied to?

**Lesson 1b: Graphing Data in Excel**

Walk everyone through graphing the data in Excel.

In the Science Kits materials, there is a lesson plan and instructions to use this example to introduce Microsoft Excel to teach your students.

I didn’t include them here because, based on the student presentations, it looks as if everyone knows how to create graphs, charts, and figures in Excel.

However, if anyone has any questions at this time, feel free to ask now and I can begin to walk through various procedures in Excel.

Questions to think about when graphing data:

What type of graph, chart, or figure is appropriate for the type of data that I have?

In this case we have frequency data. We could use a bar chart.

We also have percentages that naturally total 100%, therefore we could utilize a pie chart.

Any other questions at this time?

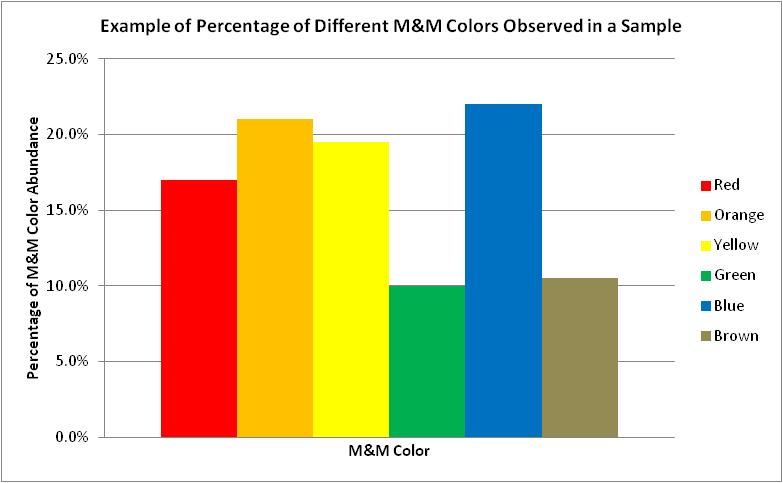
*I bet you’re beginning to wonder, yeah, we already know how to do all of this stuff, but how can we tell if there is a statistically significant difference between the frequency of colors of M&M’s in the population based on our sample?*

**Lesson 2: Using Statistics to Test Hypotheses**

**Overview**: In this lesson, we learn how to use statistics to assess how well random samples represent the total population.

Research question: Do the different colors of M&M’s differ in abundance?

*We believe that we observe a difference in the percentage of M&M’s in the random sample that we collected, but how do we know that the differences are statistically significant?*



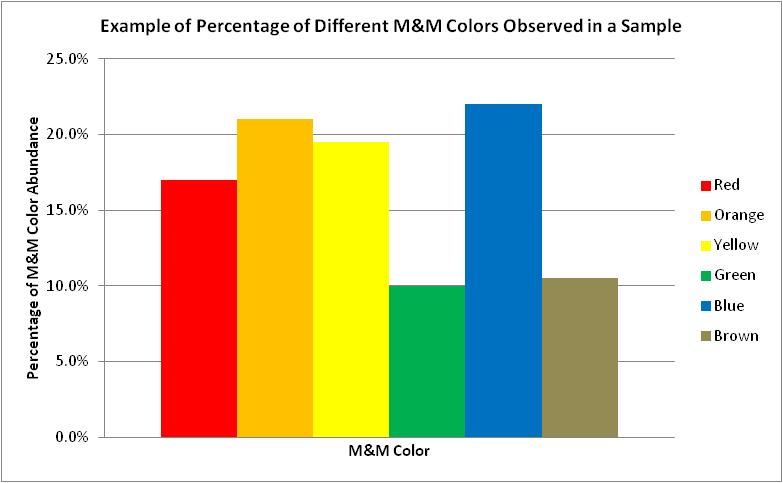
**Lesson 2 Example**

If data is collected in class, then sample data can be filled in here:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sample** | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** | **Sample Total** |
| 1 |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |
| 11 |  |  |  |  |  |  |  |
| 12 |  |  |  |  |  |  |  |
| 13 |  |  |  |  |  |  |  |
| 14 |  |  |  |  |  |  |  |
| 15 |  |  |  |  |  |  |  |
| 16 |  |  |  |  |  |  |  |
| 17 |  |  |  |  |  |  |  |
| 18 |  |  |  |  |  |  |  |
| 19 |  |  |  |  |  |  |  |
| 20 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| **M&M Color** | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** | **Sample Total** |
| **Group Total** |  |  |  |  |  |  |  |
| **Group Percentage** |  |  |  |  |  |  |  |
| **Mean** |  |  |  |  |  |  |  |
| **SD** |  |  |  |  |  |  |  |

Or, sample data can be utilized to demonstrate an example:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sample** | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** | **Sample Total** |
| 1 | 2 | 3 | 3 | 0 | 1 | 1 | 10 |
| 2 | 5 | 0 | 0 | 0 | 4 | 1 | 10 |
| 3 | 0 | 2 | 3 | 3 | 0 | 2 | 10 |
| 4 | 4 | 0 | 5 | 1 | 0 | 0 | 10 |
| 5 | 2 | 0 | 1 | 0 | 5 | 2 | 10 |
| 6 | 4 | 2 | 0 | 1 | 3 | 0 | 10 |
| 7 | 3 | 3 | 0 | 0 | 3 | 1 | 10 |
| 8 | 0 | 2 | 4 | 0 | 3 | 1 | 10 |
| 9 | 6 | 0 | 0 | 3 | 0 | 1 | 10 |
| 10 | 1 | 1 | 5 | 2 | 0 | 1 | 10 |
| 11 | 3 | 2 | 2 | 2 | 1 | 0 | 10 |
| 12 | 0 | 1 | 6 | 0 | 2 | 1 | 10 |
| 13 | 0 | 2 | 1 | 4 | 3 | 0 | 10 |
| 14 | 1 | 4 | 0 | 1 | 4 | 0 | 10 |
| 15 | 2 | 0 | 2 | 0 | 5 | 1 | 10 |
| 16 | 1 | 3 | 1 | 0 | 2 | 3 | 10 |
| 17 | 0 | 4 | 2 | 0 | 0 | 4 | 10 |
| 18 | 0 | 5 | 0 | 3 | 2 | 0 | 10 |
| 19 | 0 | 4 | 2 | 0 | 3 | 1 | 10 |
| 20 | 0 | 4 | 2 | 0 | 3 | 1 | 10 |
|  |  |  |  |  |  |  |  |
| **M&M Color** | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** | **Sample Total** |
| **Group Total** | 34 | 42 | 39 | 20 | 44 | 21 | 200 |
| **Group Percentage** | 17.0% | 21.0% | 19.5% | 10.0% | 22.0% | 10.5% | 100.0% |
| **Mean** | 1.7 | 2.1 | 1.95 | 1 | 2.2 | 1.05 | 10 |
| **SD** | 1.89 | 1.62 | 1.88 | 1.34 | 1.67 | 1.05 | 0.00 |



**Lesson 2: Using Statistics to Test the Hypotheses**

**Informational Overview**

**I. What is Statistics?**

“**Statistics**” refers to a group of numbers describing a certain collection of “things”. We collect statistics when we record our **observations** (data) for some subset (**sample**) of the entire population. A more precise definition of a statistic is a number representing a particular characteristic of a sample of the population. **Descriptive statistics** are used to summarize a sample of a population.

“Statistics” example: test scores of freshmen biology students, average body temperatures.

Descriptive statistics example: the mean (average) body temperature of 100 adults; frequency distributions of different M&M colors.

Statistics can also be used to test hypotheses to draw (or **infer**) certain conclusions about the population from which the sample is drawn. For example, we can use the frequencies of colors of M&M’s sampled earlier to make an inference about the frequency distribution of the different colors of all the M&M’s. This is called **inferential statistics** or **hypothesis-testing**.

However, we must be very careful to follow certain rules to avoid making incorrect inferences about the population. Each statistical method is applied for a specific purpose and works only if certain conditions are met. In this lesson, we will learn one statistical method for testing hypotheses, and we’ll apply it to a population of M&M candies.

**II. Testing a Hypothesis about M&M’s**

Research Question, Null Hypothesis

Mars, Inc. (the company that makes M&M’s) claims that M& M’s milk chocolate candies are currently produced in the following frequencies (percentages):

* Brown – 30; Yellow – 20; Red – 20; Orange – 10; Green – 10; Blue – 10
* In other words, the manufacturer claims that, in the whole population, the probability of getting a brown M&M is 30%, the probability of getting a yellow M&M is 20%, etc.

Procedures

In this lesson, we will demonstrate **how to test this claim about this population using statistics**. To do this, we will first need to take a sample and record some data.

* Because we are very dedicated to this cause, we selflessly went out and bought a 14 oz. bag of M&M’s milk chocolate candies.
* Without peeking into the bag, we took out 100 candies without trying to preferentially select our favorite colors.
* Another way of saying this is we took a random sample of 100 candies.
* We observed the following color distribution in our sample:
* Brown – 33; Yellow – 26; Red – 21; Orange – 8; Green – 7; Blue – 5

Observations

Note that the color distribution of our sample does not exactly match the theoretical color distribution of the population.

* For example, the frequencies for the color brown are 30 (expected) vs. 33 (observed); for the color blue, there are 10 (expected) vs. 5 (observed).
* Are the color distributions of the sample and population “close enough”, such that the differences can be attributed to the natural variations that you might expect when you take a sample?
* Or are they sufficiently different to cast doubt on the company’s stated color distribution?

Utilizing Statistics to Test the Null Hypothesis

To answer these questions, we will use a statistical test called the Chi-Squared (pronounced kai-squared) Goodness-of-Fit Test. This test will help us measure how well our data fits the claim by Mars, Inc. and allow us to infer, within the rules of statistics, if the data support the manufacturer’s claim.

However, before we can use this test, we have to make sure our sample data meets the following conditions:

*Take a few minutes and check if the M&M’s data fit all five conditions. If one or more of the conditions are not met, you may not use this test. If all of the above conditions are met, you may proceed to set up a hypothesis and test it with the Chi-Squared Goodness-of-Fit Test.*

1. The sample data are in categories.

Yes – in this case, the categories are colors.

1. Each observation must be classified into exactly one category.

Yes – Each M&M sampled will be exactly one of the six colors. It’s not possible for it to be no color, or two colors.

1. The sample data have been collected randomly.

Yes – We picked a sample without peeking (we didn’t pick our favorite colors or avoid the colors we disliked).

1. The sample data are listed as frequency counts for each category.

Yes – of the 100 candies selected, 33 are brown, 26 are yellow, etc.

1. For each category, the expected frequency is at least 5.

The expected frequency is the frequency of each color from the theoretical color distribution for the population. These are the percentages provided by Mars, Inc. – that is, Brown – 30%; Yellow – 20%; Red – 20%; Orange – 10%; Green – 10%; Blue – 10%.

Using a sample of 100, the lowest expected frequency is 10% \* 100 = 10, which is greater than 5, so this condition is met.

**III. Writing Statistical Hypotheses**

To use the Chi-Squared Goodness-of-Fit Test, we first have to form a hypothesis (so we have something to test). Actually, in statistics, we form two hypotheses:

1. The first is called the null hypothesis – this is the hypothesis we are testing. The null Hypothesis (H0) must always be stated in terms of equality (=, ≥, or ≤).
2. The second hypothesis (H1) is called the alternate hypothesis and it is essentially the opposite of the null hypothesis. If the null hypothesis is false, then the alternate hypothesis is true.

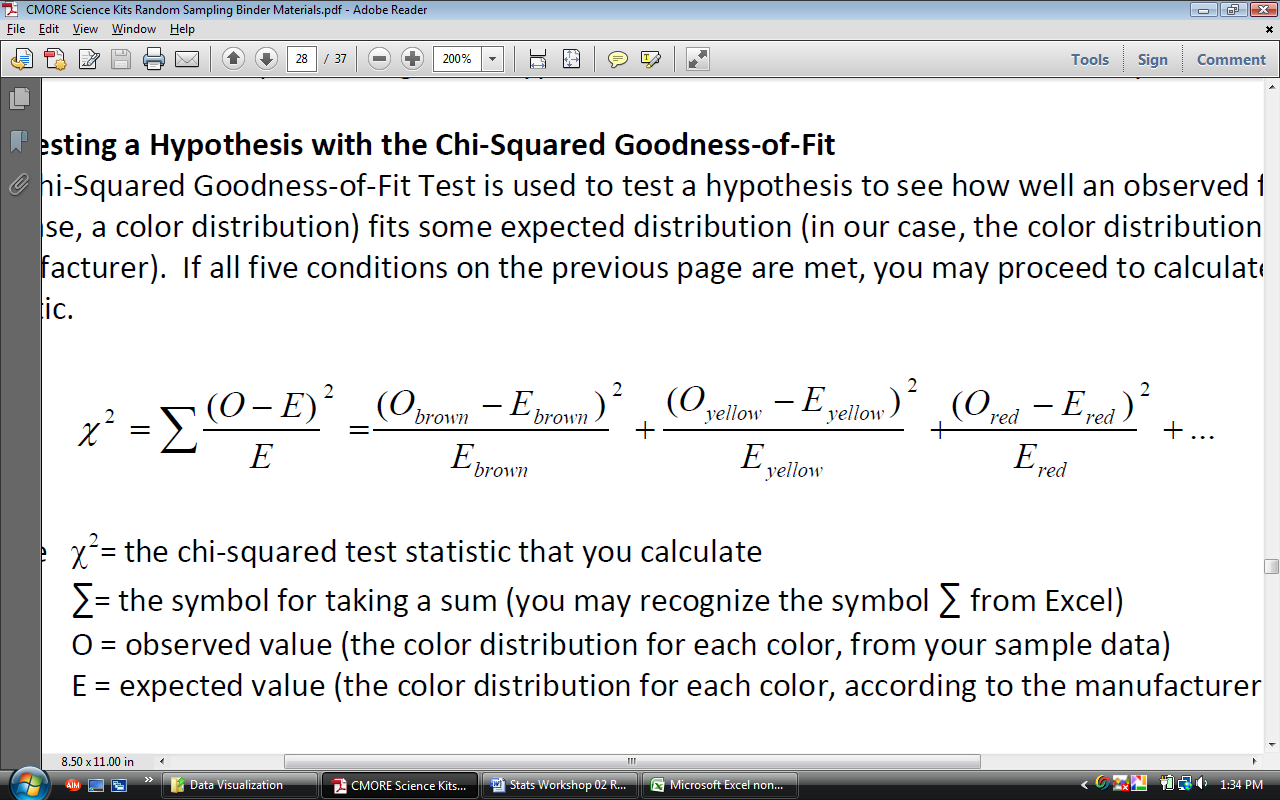
In our M&M’s example, writing a null hypothesis can be tricky because there are so many colors. For now, be thinking about what the null and alternate hypotheses are and we will come back to this in a minute.

**IV. Testing a Hypothesis with the Chi-Squared Goodness-of-Fit**

The Chi-Squared Goodness-of-Fit Test is used to test a hypothesis to see how well an observed frequency distribution fits some expected distribution.

* In the M&M example, the Chi-Squared Goodness-of-Fit Test can test the hypothesis to see how well the observed sample color distribution fits the color distribution claimed by the manufacturer.
* If all five conditions on the previous page are met, you may proceed to calculate the Chi-Squared test statistic.

The formula for the Chi-Squared Goodness-of-Fit Test statistic is:



Where:

* χ2 = the chi-squared test statistic that you calculate
* Σ = the symbol for taking a sum (you may recognize the symbol Σ from Excel)
* *O* = observed value (the color distribution for each color, from your sample data)
* *E* = expected value (the color distribution for each color, according to the manufacturer’s claimed distribution)

In other words, this formula tells you to calculate the difference between O and E for each color, square each difference, and divide by E. Do this for every color and then add them all together. The squaring of the difference gives us a positive number. The greater the differences between the observed and expected value, the bigger the test statistic χ2 will be, and the more likely that we will reject the null hypothesis. To decide if the test statistic is big enough to reject, read on…..

Some important things to consider while looking at the statistical formula are:

* We’re looking at the difference between the “expected” and the “observed”. Large differences in the “expected” and the “observed” will be signficiant.
* Minor “random” variation will be cancelled out. If I observe 1 or 2 more than expected here, and 1 or 2 less than expected here, they will be taken together additively to equal zero, or non-significant.
* The squaring in the formula is where the magnitude of the effect is detected.
  + If the squaring wasn’t included, then everything would equal zero.
  + However, what do we know about squaring numbers (12 = 1; 22 = 4; 32 = 9; 42 = 16)? They get bigger exponentially. Therefore, if there is a large deviation in the observed versus the expected, we will be able to detect an effect. It also cancels out the negative values (-12 = 1; -22 = 4; -32 = 9; -42 = 16).

**V. Using the test statistic to evaluate the null hypothesis**

How big does the test statistic have to be to reject the null hypothesis? It depends on two things:

1. The number of **degrees of freedom**, which is simply one less than the number of categories. In our case, we have 6 categories (6 colors), so the number of degrees of freedom is 5.
2. The **significance level**, which is something we choose. Think of the significance level as a measure of the amount of error that you can live with – because we can never be 100% certain when we reject a null hypothesis that it is really false. It’s common to accept a significance level of 0.05 (which means that 5% of the time that we reject a null hypothesis, we will do so in error).

Once you have these two pieces of information, you can look up the “critical value” on a Table of Chi-Square Probabilities. If the test statistic is bigger than the critical value, you REJECT the null hypothesis. If the test statistic is less than the critical value, you ACCEPT the null hypothesis.

Are you confused yet? Reading about statistics can get confusing!! It’s much easier to learn statistics by doing statistics. So, let’s follow the steps in the STUDENT WORKSHEET – Lesson 3: Testing Hypotheses with Statistics to apply this statistical test to M&M’s.

**Lesson 2: Using Statistics to Test the Hypotheses**

**Worksheet**

1. Fill in Table 1 below with the **observed frequency** (O) for the actual sample of 100 M&M’s and the **expected frequency** (E) that Mars, Inc. claims to manufacture for each candy color, based on the information provided in the INFORMATIONAL OVERVIEW SHEET – Lesson 2: Utilizing Statistics to Test Hypotheses. Complete the data table below for all colors, and give a descriptive title to the table.

Table Title: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** |
| **Observed frequency (O)** |  |  |  |  |  |  |
| **Expected frequency (E)** |  |  |  |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Red** | **Orange** | **Yellow** | **Green** | **Blue** | **Brown** |
| **Observed frequency (O)** | 21 | 8 | 26 | 7 | 5 | 33 |
| **Expected frequency (E)** | 20 | 10 | 20 | 10 | 10 | 30 |

1. In order to visualize these data, make a bar chart or histogram to compare the observed vs. expected frequencies. Use one color for the observed values and another color for the expected values. Label your axes and give the graph a title. If you have computer access, you may do this in Microsoft Excel.
2. How similar are the observed and expected frequencies? Do you think the Goodness-of-Fit test will reveal that the observed sample data supports the company’s claim? Explain your answer.
3. Writing Statistical Hypotheses
   1. Write a null hypothesis to test the claim that Mars, Inc. is making for the color distribution of M&M’s. Remember the null hypothesis is what you are attempting to prove or disprove and that it must be expressed in terms of equality. The first part of the null hypothesis is written for you; complete the rest.
   2. Let P= the population proportion of a given color
   3. H0: Pbrown = 30% and Pyellow = \_\_\_ and Pred = \_\_\_ and Porange = \_\_\_ and Pgreen = \_\_\_ and Pblue= \_\_\_
   4. Write the alternate hypothesis. This is the hypothesis that must be true whenever the null hypothesis is false.
   5. H1: At least one of the above proportions is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ from the claimed value.
   6. H1: Pbrown ≠ 30% or Pyellow ≠ \_\_\_ or Pred ≠ \_\_\_ or Porange ≠ \_\_\_ or Pgreen ≠ \_\_\_ or Pblue ≠ \_\_\_
4. Write out the formula for the Chi-Square Test Statistic and define all terms.

Have students complete this part in class.

1. Calculate and fill in the values in Table 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Color Category | Observed Frequency (O) | Expected Frequency (F) | O - E | (O - E)2 | (O - E)2  E |
| Red |  |  |  |  |  |
| Orange |  |  |  |  |  |
| Yellow |  |  |  |  |  |
| Green |  |  |  |  |  |
| Blue |  |  |  |  |  |
| Brown |  |  |  |  |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Color Category | Observed Frequency (O) | Expected Frequency (F) | O - E | (O - E)2 | (O - E)2  E |
| Red | 21 | 20 | 1 | 1 | 0.05 |
| Orange | 8 | 10 | -2 | 4 | 0.40 |
| Yellow | 26 | 20 | 6 | 36 | 1.80 |
| Green | 7 | 10 | -3 | 9 | 0.90 |
| Blue | 5 | 10 | -5 | 25 | 2.50 |
| Brown | 33 | 30 | 3 | 9 | 0.30 |

Complete using the means from the earlier example in-class.

1. Use your answers from the above table to calculate the χ2.

χ2 = \_\_\_\_\_\_\_\_\_\_ 5.95

1. Now that we have determined the χ2 test statistic, we want to compare this value to the critical χ2 value. To do this, we will need to know the number of degrees of freedom.
   1. Degrees of Freedom (df) = \_\_\_\_\_\_ 5
   2. How did you determine the number of degrees of freedom?
2. Using the degrees of freedom (df) determined above and the 0.05 significance level, look up the critical χ2 value on a Table of Chi-Square Probabilities (an abbreviated version of this table, up to df=10, is given below).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **df** | **0.10** | **0.05** | **0.025** | **0.1** |
| **1** | 2.706 | 3.841 | 5.024 | 6.635 |
| **2** | 4.605 | 5.991 | 7.378 | 9.210 |
| **3** | 6.251 | 7.815 | 9.348 | 11.345 |
| **4** | 7.779 | 9.488 | 11.143 | 13.277 |
| **5** | 9.236 | 11.070 | 12.833 | 15.086 |
| **6** | 10.645 | 12.592 | 14.449 | 16.812 |
| **7** | 12.017 | 14.067 | 16.013 | 18.475 |
| **8** | 13.362 | 15.507 | 17.535 | 20.090 |
| **9** | 14.684 | 16.919 | 19.023 | 21.666 |
| **10** | 15.987 | 18.307 | 20.483 | 23.209 |

* 1. Critical Value of χ2 = \_\_\_\_\_\_\_\_\_\_\_ 11.070

1. Compare the χ2 test statistic to the critical value, and circle the correct answers below:
   1. The χ2 test statistic is (larger, smaller) than the critical value, so we (reject, do not reject) the null hypothesis.
   2. The χ2 test statistic is (larger, smaller) than the critical value, so we (reject, do not reject) the null hypothesis.
2. Interpret your results.

**Glossary of Terms Used in This Lesson**

**alternate hypothesis (H1)**: The opposite of the null hypothesis. If the null hypothesis is T = 98.6°, then the alternate hypothesis is T ≠ 98.6°.

**biased**: Favoring or preferring one result over other results (not random).

**chi‐square goodness-of-fit test (χ2)**: (kai-square); A statistical method used to test a hypothesis.

**dependent variable**: The variable you measure in an experiment; it is dependent on another variable (the **independent variable**).

**frequency**: The number of items in a particular category.

**hypothesis**: A prediction that can be tested.

**independent variable**: The variable that is not influenced by other variables in the experiment.

**infer**: To draw certain conclusions.

**null hypothesis (Ho)**: The hypothesis being tested. The null hypothesis must be stated in terms of equality (=, ≥, or ≤). Example: T = 98.6°F.

**probability**: The chance that a particular outcome will happen.

**random sampling**: Selecting items purely by chance, without any favor, preference or prejudice. Random sampling is used to estimate the population sizes of species within an ecosystem when it is not possible or practical to count all the individuals.

**Statistics**: A collection of mathematical methods used to describe data and draw conclusions.