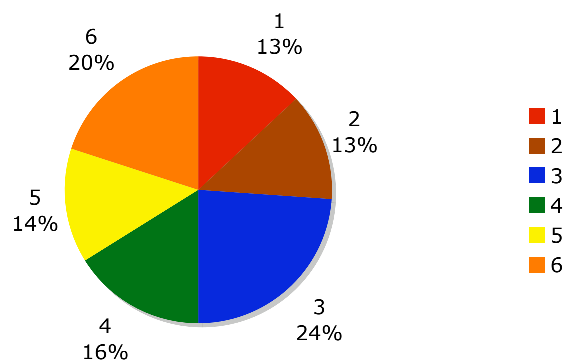
"Sweet Statistics: A Chi Square Activity that Melts in Your Mouth…" *

(Revised July 2010)

Have you ever wondered why the package of M&Ms you just bought never seems to have enough of your favorite color? Or, why is it that you always seem to get the package of mostly brown M&Ms? What's going on at the Mars Company? Is the number of the different colors of M&Ms in a package really different from one package to the next, or does the Mars Company do something to insure that each package gets the correct number of each color of M&M? I'll bet like me, you've stayed up nights wondering about this!

Well, according to candy folklore, each package of milk chocolate M&Ms that the Mars Company sells has the following percentages of each color of M&M. (Wow…that must be some job!)

Fig. 1 Milk Chocolate M&M Colors



1 = red
2 = brown
3 = blue
4 = green
5 = yellow
6 = orange

One way that we could determine if the Mars Co. is true to its word is to sample a package of M&Ms and do a type of statistical test known as a "goodness of fit" test. These type of statistical tests allow us to determine if any differences between our observed measurements (counts of colors from our M&M sample) and our expected (what the Mars Co. claims) are simply due to chance sample error or some other reason (i.e. the Mars Co.'s sorters aren't really doing a very good job of putting the correct number of M&M's in each package). The goodness of fit test we will be doing today is called a **Chi Square Analysis**. This test is generally used when we are dealing with discrete data (i.e. count data, or non-continuous data). Just like in a t-test, we will be calculating a test statistic (this time it is called the chi square or X^2, where X is a Greek letter pronounced "kye."). In addition, just like in a t-test we will be using a table to determine a probability of getting this particular X^2 value. Remember that, just like in a t-test, our probability values tell us what the chances are that the differences in our data are due simply to chance alone (sample error).

Let's start by stating our null hypothesis. Assuming the Mars Company M&M sorters are doing their job correctly…

> *There is no difference in M&M color ratios between actual store-bought bags of M&Ms and what the Mars Co. claims are the actual ratios.*

*Author: J. Peters, College of Charleston, SC, as shared by C. Hilvert, Glenbrook South HS, IL; many similar activities exist.

To test this null hypothesis we will need to calculate the X^2 statistic, which is calculated in the following way:

$$X^2 = \Sigma \frac{(O-E)^2}{E}$$

O is the observed value, E is the expected value

After the quantity $\mathbf{(O-E)^2 / E}$ is calculated for each category, values for all categories are summed.

The main thing to note about this formula is that, when all else is equal, the value of X^2 increases as the difference between the observed and expected values increase.

On to the sample!

1. Wash your hands. You will be handling food that you *may* want to munch on later!
2. Your teacher will open a bag of M&Ms and split it up between the groups of your class.
3. DO NOT EAT ANY OF THE M&Ms (for now)!
4. Separate the M&Ms into color categories and count the number of each color of M&M you have.
5. Record your observed counts under each category on the "master candy list" on the board.
6. Once the entire class has counted and recorded their data, sum the observed counts for each color category.
7. Now determine the total number of M&Ms in the bag.
8. Calculate the expected numbers of M&Ms in each color category using the information in figure 2.
9. Record both the total observed counts and the total expected counts in the table below.

Fig. 2 Class Data for Calculating Chi Square (X^2)

	Color Categories						
	Brown	**Blue**	**Orange**	**Green**	**Red**	**Yellow**	**Total**
Observed number (O)							
Expected number (E)							
Deviation = (O – E)							
D^2 = $(O-E)^2$							
D^2 / E = $(O-E)^2/E$							
X^2							

Now determine the probability that the difference between the observed and expected values (as summarized by the calculated value of the chi-square) occurred simply by chance (sample error). The procedure is to compare the calculated value of the chi-square to the appropriate value in Figure 3. But first -- note the term "degrees of freedom". For this statistical test the degrees of freedom equal the number of classes (i.e. color categories) minus one.

Degrees of freedom = number of categories – 1

In our M&M data the number of color categories is 6, so degrees of freedom = 6 –1 = 5.

The reason why it is important to consider degrees of freedom is that the value of the chi-square statistic is calculated as the sum of the squared deviations for all classes. The natural increase in the value of chi-square with an increase in classes must be taken into account.

Scan across the row of Figure 3 corresponding to 5 degrees of freedom. Values of the chi-square are given for several different probabilities, ranging from 0.90 on the left to 0.01 on the right.

Fig. 3 Chi Square vs. Probability Values

Degrees of Freedom	**Probability**					
	Not significant				Significant	
	0.90	**0.50**	**0.25**	**0.10**	**0.05**	**0.01**
1	0.016	0.46	1.32	2.71	3.84	6.64
2	0.21	1.39	2.77	4.61	5.99	9.21
3	0.58	2.37	4.11	6.25	7.82	11.35
4	1.06	3.36	5.39	7.78	9.49	13.28
5	1.61	4.35	6.63	9.24	11.07	15.09

Notice that a chi-square value of 1.61 would be expected by chance in 90% (0.90) of the cases, whereas one as large as 15.09 would only be expected by chance in 1% (0.01) of the cases. The column that we need to concern ourselves with is the one under "0.05". Scientists, in general, are willing to say that if their probability of getting the observed deviation from the expected results by chance is greater than 0.05 (5%) then we can accept the null hypothesis. In other words, there is no significant difference in M&M color ratios between actual store-bought bags of M&Ms and what the Mars Co. claims are the actual ratios. Stated another way…any differences we see between what Mars claims and what is actually in a bag of M&Ms just happened by chance sampling error. Five percent! That's not much, but it's good enough for a scientist!

If, however, the probability of getting the observed deviation from the expected results by chance is less than 0.05 (5%) then we should reject the null hypothesis. In other words, for our study, there is a statistically significant difference in M&M color ratios between actual store-bought bags of M&Ms and what the Mars Co. claims are the actual ratios. Stated another way…any differences we see between what Mars claims and what is actually in a bag of M&Ms did not just happen by chance sampling error.

Based on our sample bag of M&Ms, should we accept or reject our null hypothesis? You decide!

If we did reject our null hypothesis, what might be some possible explanations for this outcome?

If we did accept our null hypothesis, what might be some possible explanations for this outcome?