M&M Statistics/A Chi Square Analysis

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?

Here's some information from the M&M website:

% color	Plain	Peanut	Crispy	Minis	Peanut Butter	Almond
Brown	13%	12%	17%	13%	10%	10%
Yellow	14%	15%	17%	13%	20%	20%
Red	13%	12%	17%	12%	10%	10%
Green	16%	15%	16%	12%	20%	20%
Blue	24%	23%	17%	25%	20%	20%
Orange	20%	23%	16%	25%	20%	20%

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. This type of statistical test 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 or some other reason (i.e. the Mars company's sorters aren't 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 using is called a Chi Square ($\mathbf{X}^2$) Analysis.

We will be calculating a statistical value and using a table to determine the probability that any difference between observed data and expected data is due to chance alone.

We begin by stating the null hypothesis. A null hypothesis is the prediction that something is not present, that a treatment will have no effect, or that there is no difference between treatment and control. Another way of saying this is the hypothesis that an observed pattern of data and an expected pattern are effectively the same, differing only by chance, not because they are truly different.

What is our null hypothesis for this experiment?

To test this hypothesis we will need to calculate the $\mathbf{X^2}$ statistic, which is calculated in the following way:

$$X^2 = \text{Sum of } \frac{(o-e)^2}{e}$$

where o is the observed (actual count) and e is the expected number for each color category. The main thing to note about this formula is that, when all else is equal, the value of $\mathbf{X^2}$ increases as the difference between the observed and expected values increase.

On to it!

1. Lay out a large sheet of paper—you'll be sorting M&Ms on this.
2. Open up a bag of M&Ms and split them between the members of lab group.
3. DO NOT EAT ANY OF THE M&M'S (for now!)
4. Separate the M&M's into color categories and count the number of each color of M&M you have.
5. Record your counts in Data Chart 1.
6. Determine the Chi square value for your data.

Data Chart 1	**Color Categories**						
	Brown	Blue	Orange	Green	Red	Yellow	**Total**
Observed (o)							
Expected (e)							
Difference (o-e)							
Difference Squared $(o-e)^2$							
$(o-e)^2/e$							
$\Sigma (d^2/e) = X^2$							

Now you must determine the probability that the difference between the observed and expected values occurred simply by *chance*. The procedure is to compare the calculated value of the chi-square to the appropriate value in the table below. First examine the table. 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 your M&M experiment, what is the number of degrees of freedom? ___________

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 corresponding to your 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. Note that the chi-square increases as the probability decreases. If your exact chi-square value is not listed in the table, then estimate the probability.

Accept the null hypothesis ← | → Reject

Degrees of Freedom	**Probability**					
	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 as large as 1.61 would be expected by chance in 90% of the cases, whereas one as large as 15.09 would only be expected by chance in 1% of the cases. Stated another way, it is more likely that you'll get a little deviation from the expected (thus a lower Chi-Square value) than a large deviation from the expected. 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 really no difference in actual ratios…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 is 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 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.

The following information should be in your conclusion.

Based on your individual sample, should you accept or reject the null hypothesis? Why?

If you rejected your null hypothesis, what might be some explanations for your outcome?

Now that you completed this chi-square test for your data, let's do it for the entire class, as if we had one huge bag of M&Ms. Using the information reported on the whiteboard, complete Data Chart 2.

Data Chart 2	Color Categories					
	Brown	Blue	Orange	Green	Red	Yellow	**Total**
Observed							
Expected (e)							
Deviation (difference between expected and observed)							
Deviation Squared (d^2)							
d^2/e							
$\Sigma (d^2/e) = X^2$							

The following also should be discussed in your conclusion.

Based on the class data, should you accept or reject the null hypothesis? Why ?

If you rejected the null hypothesis based on the class data, what might be some of the explanations for your outcome?

If you accepted the null hypothesis, how do you explain it—particularly if you rejected the null based on individual group data? What is the purpose of collecting data from the entire group?