A study in Statistical Fluctuations

Today, we will look at techniques that are useful in determining how confident we can be in answering questions related to measurement. In particular, we will look at how many dark brown M&Ms are typically found in a bag.

To begin, your only source of data is a single bag of M&Ms.

1. Open your bag of M&Ms and record all available data:
   number of dark brown M&Ms;
   number of red M&Ms;
   number of orange M&Ms;
   number of yellow M&Ms;
   number of green M&Ms;
   number of blue M&Ms;
   total number of M&Ms.

   Enter your data in the Google document found at:
   http://www.delicious.com/koutsoliotas/
   (Look for the M&M link.)

INDIVIDUAL DATA

Task 1: Write down your best guess for the typical number of dark brown M&Ms in a bag. What, if anything, can you say about how close the 'true' answer is to your guess?

SMALL GROUP DATA

2. Find a group of 3 or 4 other people, and share your results.

Task 2: Revise your answers to the previous task based on the additional information provided by your group. What level of confidence can you place on this number? Discuss this with your group.

Task 3: Complete the sentence,
"I am almost certain that the 'true' answer is within ________ of our estimate."

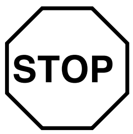
3. What if we were to collect more data? Go to the board, and write down your individual result under the category of "Math" group or "Science" group. Depending on whether you are from the Math or Science group, take the collective data from your group and create a graph of 'person' on the *x*-axis versus "no. of dark brown M&Ms" on the *y*-axis.

LARGE GROUP DATA (MATH OR SCIENCE)

Task 4: Write down your best guess of the typical number of dark brown M&Ms in a bag, based on all the information available to your Math/Science group.

Task 5: Identify this value on your graph by drawing a dashed line to represent this 'best guess.'

Task 6: In your small groups, devise a method that captures the range of fluctuations observed between different bags of M&Ms.



Discuss your method with one of the instructors before continuing.

Develop a set of rules for determining the variation inherent in individual measurements of the same quantity. This set of rules should be applicable in all conceivable cases in a standardized way.

We call this range of fluctuations the ***standard deviation*** (s_{indiv}).

4. To better visualize the data, we can make a graph that tells us how frequent each result is. Such a graph is called a *histogram*.

Task 7: Create a table of the following form that contains all the results of your Math/Science group.

Number of dark brown M&M's	Number of bags
0	0
⋮	⋮
8	3
9	4
⋮	⋮

Histogram data (an example)

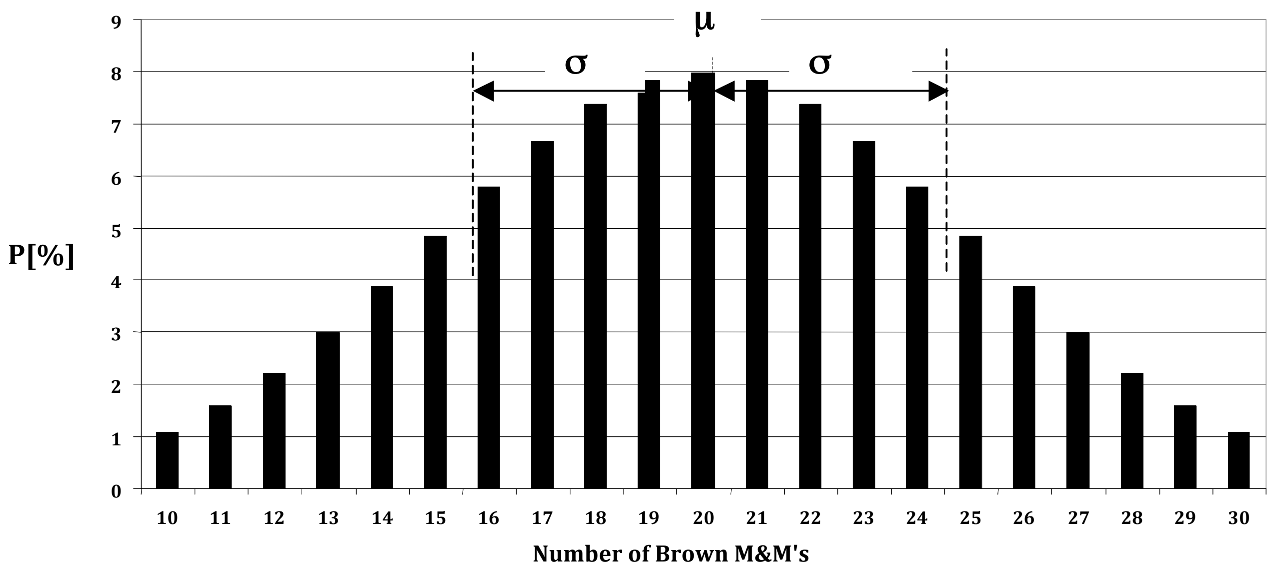
Check your results with the other members in your small group, and then plot the histogram on the graph paper provided. Please note that the histograms may look different for the Math and the Science groups.

THE ULTIMATE DATA SET

So far we have only opened one M&M bag for each student (*about 25* M&M bags for each group). In order to fully characterize the complete distribution of dark brown

M&Ms in this batch, our goal would be to open every single M&M bag produced in this batch. Besides getting a stomach pain from eating all these M&Ms, this task is obviously impossible. Even though that task is impossible, we don't give up. Instead, we imagine what the probability of getting a particular number of dark brown M&Ms would look like.

With enough bags, our histogram may begin to look like:



Probability Distribution of the universe of possible measurements.

This distribution commonly arises when making measurements in the presence of uncertainty. It is known by various names: the normal distribution; the bell-shaped curve; and the Gaussian distribution.

We call this distribution, *the distribution of the universe of possible measurements*. Just as before it would have a mean value, the true mean

Features of the curve:

- The peak of this curve represents the most likely value, and it occurs at the average (or *mean*) of the distribution. (symbol: μ)

- The breadth of the curve provides an indication of how large the fluctuations of individual measurements may be. This is called the *standard deviation* of the individual result. (symbol: σ)

- approximately 68% of the measurements lie within one standard deviation from the true mean,
- approximately 95% within two standard deviations of the true mean
- approximately, 99.74% within three standard deviations of the true mean.

Task 8: On your histogram, identify the mean (or average) with a dashed line and show the limits for +/- one standard deviation away from your mean. Determine what proportion of the total number of bags lies within one standard deviation of your average. Compare this value with the expected result and comment.

COMBINING DATA SETS

5. Now, this standard deviation represents the size of typical fluctuations of an individual bag. However, each of the Math/Science groups can call on information gathered from 25 or so bags.

We would like to know how far our average is from the true average. Because it is an average of *N* measurements, and is calculated from measurements that lie both above and below the true average, you might correctly guess that the average of N results is likely to be closer to the true average than a typical individual measurement. *In other words, if we were to repeat the measurement with another (different) set of 25 bags, our average value would not fluctuate as wildly as that of another individual bag.*

The answer turns out to be that the size of the fluctuations is *smaller* than the individual fluctuations by an amount that is $\sqrt{N}$.

$$\text{standard deviation of the average} = \frac{s_{indiv}}{\sqrt{N}}.$$

Task 9: For your Math/Science group, quote your best estimate for the number of dark brown M&Ms in a typical bag, along with your guess for the standard deviation.

COMPARING DATA SETS

The two groups have made independent measurements of the same quantity: the number of dark brown M&Ms in a typical bag. Compare the two results and comment on their consistency.

Task 10. Is there any overlap in the predictions of the two groups? Is this surprising? What might happen if the entire experiment was repeated?

For the Interested:

Q: Describe why the distribution of the number of brown M&Ms in a bag is *not* actually a Gaussian.