COLOR BY NUMBERS

1. Engage

a. Open your bag of M&Ms and record the distribution on the frequency table.



b. What color's sample proportion seems the most unusual to you?

According to Mars, Inc., the distribution of the population of milk chocolate M&Ms are:



c. Assuming Mars, Inc.'s distribution is correct, use the binomial distribution to determine the probability of observing a sample proportion like yours or one more severe.

2. Explore

a. Write appropriate hypotheses and choose an alpha level.

H₀:

H_A:

Alpha:

b. Check the necessary assumptions and conditions.

Random:

10%:

S/F:

Is our sample large enough to test the company's claim that $p_0 = .2$?

What sample size would allow us to use the normal model without being cut off for at least 3 standard deviations?

 $np \ge 10$, so when I divide both sides by p I get $n \ge 10p$

Our minimum sample size to be able to use a **1-Proportion Z-Test** would be **n** = ____.





3. Explain

Input first trial data into the data sheet to calculate totals and generate a scatterplot of the cumulative relative frequency of orange M&Ms vs. sample size.

Record the total number of orange M&Ms and the total number of M&Ms sampled after the first trial.

Orange Freq. =

What is your sample proportion? $\hat{\mathbf{p}} =$

n =

There are supposed to be 20% orange candies. Does our data provide statistically significant evidence to contradict the company's claim?

- a. Perform the mechanics of the test (find the z-score and p-value) and fill in the sampling distribution model to the right.
- b. Explain carefully what the p-value means in context.
- c. What is your conclusion?



4. Extend

Let us eat more chocolate! I mean, let's gather more data! Input your data by your name in the data sheet for the second trial.

Record the total number of orange M&Ms and the total number of M&Ms sampled after the second trial.

Orange Freq. = n =

What is your sample proportion? $\hat{\mathbf{p}} =$

There are supposed to be 20% orange candies. Does our data provide statistically significant evidence to contradict the company's claim?

- a. Perform the mechanics of the test (find the z-score and p-value) and fill in the sampling distribution model to the right.
- b. Explain carefully what the p-value means in context.
- c. What benefit do we get from increasing our sample size?
- d. Did increasing our sample size change our conclusion?
- e. Write a 90% confidence interval from our data.



5. Evaluate

Let's investigate: At which point, if any, did our sample cross the threshold of significance?

a. At what point, if any, could we have had a p-value significant enough to reject our null hypothesis?

We are going to graph the margin of error from our $p_0 = .2$ to see if at any point during our simulation we could have rejected the null.

If we are doing a 2-sided test with an alpha level of $\alpha = .1$ then we will need a confidence interval with 5% on each tail. This would be a **90% CI** with a **z* = 1.65**

Put the following equations into your calculator. Then press Zoom + 9.

- b. At any point would we have had enough evidence to reject the H_0 ?
- c. Try it again with an alpha level of $\alpha = .05$ then we will need a confidence interval with 2.5% on each tail. This would be a 95% CI with a critical value $z^* = 1.96$.
- d. Does our sample proportion appear to be converging to the $p_0 = .2$ as we opened more bags of candy?

