FFM 8/19/14:

1.) Among various ethnic groups, the population standard deviation of heights is known to be approximately 3 inches. We wish to construct a 95% confidence interval for the average height of men from Sweden. 48 Swedish men are surveyed. The sample mean is 71 inches. What will happen to the width of the confidence interval obtained if 1000 Swedish men are surveyed instead of 48? Why?

a.	It wil	l be	wider because	
b.	It wil	l be	narrower because	
C.	no ch	ang	e because	_

2.) In a study of the length of time that students take to earn bachelor's degrees, a sample of 1,200 students took an average of 5.9 years to complete their degree. Assume that the population standard deviation is known to be 2.3 years. What would happen to the width of the confidence interval if 300 students were surveyed instead of 1,200? Why?

a.	It	would	be	narrower because	¥		
b.	It	would	be	wider because			
C.	. No change because						

Unit 1, Task 1 - Sample Proportions with Simulation

I. Reviewing some basic definitions:

a. Think about a single bag of M&M's. Does this single bag represent a sample of M&M's or the population of M&M's?

b. We use the term statistic to refer to measures based on samples and the term parameter to refer to measures of the entire population.

If there are 62 M&M's in your bag, is 62 a statistic or a parameter? Statistic

n=62

If Hershey claims that 24% of all M&M's are blue, is 24% a statistic or a parameter? To remark 7

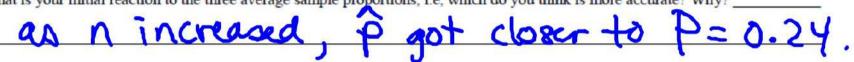
Blue: +6%

II. How many blue candies should	ld I expect in a bag of	M&M's?			
First, random samples of 10 c samples were recorded by e			M's. The count and	proportion of each colo	r in the
Binomials		Blue	Not Blue		
Dillo	Count	3	7		
	Proportion	330=3	4= 7		
The average of the sample propo	ortions obtained by each	lass for sample size	10 wa recorded and	complied.	
a. Do you think their averages are				_	_
his/her sample?Closer	Why do you think that?	average-	> cente	ur of dai	a.
Next, random sample of 25 car obtained by each class for samp	ndies were taken and re	ecorded by each Math			
Finally, random samples of 40 obtained by each class for samp			ath IV student. The a	verages of the sample p	proportions
b. Do you think the averages of the	e 40 candies are closer	or farther from the true	e proportion of blue l	M&Ms than each indivi	idual
student found in his/her sample of	size 10 and 25? 0		Sm	iller erro	r,
		e.	stimate	is more a	cchade

Here is the data obtained in one of the class periods:

		6th period	
TOTALS:	n= 10	n= 25	n= 40
count:	46	105	165
# of samples:	170	425	680
sample proportion:	27.06%	24.71%	24.26%
			M

c. What is your initial reaction to the three average sample proportions, i.e, which do you think is more accurate? Why? _



III. Sampling Distribution of \hat{p}

We have been looking a number of different sampling distributions of \hat{p} , but we have seen that there is great variability in the distributions. We would like to know that \hat{p} is a good estimate for the true proportion of blue M&M's. However, there are guidelines for when we can use the statistic to estimate the parameter. This is what we will investigate in the next section.

First, however, we need to understand the center, shape, and spread of the sampling distribution of \hat{p} .

We know that if we are counting the number of M&M's that are blue and comparing with those that are not blue, then the counts of blues follow a binomial distribution (given that the population is much larger than our sample size). Formulas for the mean and standard deviation of a binomial distribution were discussed in Math III.

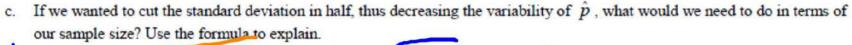
Given that $\hat{p} = \frac{X}{n}$, where X is the count of blues and n is the total in the sample, how might we find $\mathcal{H}_{\hat{p}}$ and $\sigma_{\hat{p}}$? Find formulas for each statistic.

This leads to the following statement of the characteristics of the sampling distribution of a sample proportion:

The Sampling Distribution of a Sample Proportion:

The ose a simple random sample of size n from a large population with population parameter p having some characteristic of interest. Let \hat{p} be the proportion of the sample having that characteristic. Then:

- The mean of the sampling distribution is
- The standard deviation of the sampling distribution is ____
- a. Let's look at the standard deviation a bit more. What happens to the standard deviation as the sample size increases?
- b. Show a few examples to verify you conclusion. Then use the formula to explain why your conjecture is true:





*Caution: We can only use the formula for the standard deviation of \hat{p} when the population is at least 10 times as large as the n.

d. For each of the samples taken in part 3, determine what the population of M&M's must be for us to use the standard deviation formula derived above. Is it safe to assume that the population is at least as large as these amounts? Explain.

$$\frac{N}{10} \times 10 \frac{N}{100}$$

$$\frac{250}{400} \times 100 \frac{N}{100}$$

$$\frac{N}{100} \times 100 \frac{N}{100}$$

CLT for Proportions:



characteristic of interest

* IF SRS, size n, from large population ($N \ge 10n$, $np \ge 10$, $nq \ge 10$) with population parameter, p, having some

* THEN Sampling distribution of sample proportion, \hat{p} ,

- is approximately normal

- has $\mu_{\hat{p}} = p$



* NOTE: This approximation becomes more and more accurate as the sample size, n, increases.

* ALSO NOTE: The CLT allows us to use normal calculations to determine probabilities about sample proportions obtained from populations that are not normally distributed.