SAMPLING REESE'S PIECES

1. Take a guess at the proportion of all Reese's Pieces that are orange.

2. Mr. Tyson will give you a sample of 25 Reese's Pieces. Record the number and proportion of your sample that are yellow, orange, and brown. Who or what are the observational units? What variable are you measuring? Is this variable categorical or quantitative?

	Orange	Yellow	Brown
Count			
Proportion			

3. Do you suppose that everyone in the class obtained the same proportion of orange candies? With your classmates, construct a dotplot of the sample proportions of orange candies. Label the number line appropriately.
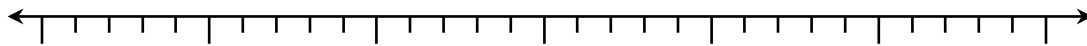
4. Does a pattern appear to be emerging? To see the pattern more clearly, get a laptop and go to http://www.rossmanchance.com/applets and open the Reese's Pieces applet. Set the sample size to 25, number of samples to 1, and the population proportion (π in this applet) to 0.45. Click "Draw Samples" 5 times to simulate taking 5 samples of 25 candies. Were the sample proportions ($\hat{p}$) reasonably close to the population proportion?

5. Now uncheck "Animate," enter 495 as the number of samples (NOT the sample size), and click "Draw Samples." Record the mean, standard deviation, and shape of the resulting distribution of sample proportions in the table on the back of this page.

6. Predict how the distribution would change if the sample size was n = 75. Make this change in the applet, uncheck "Animate," enter 500 as the number of samples, and click "Draw Samples." Was your prediction correct? Record the mean, standard deviation, and shape of this distribution in the table on the back of this page.

7. Continue filling in the table below using this applet. Make sure that "Animate" is unchecked and that you draw 500 samples each time.

Population Proportion	Sample Size	Mean of $\hat{p}$ distribution	Standard Deviation of $\hat{p}$ distribution	Shape of $\hat{p}$ distribution
p = 0.45	n = 25			
p = 0.45	n = 75			
p = 0.45	n = 300			
p = 0.10	n = 25			
p = 0.10	n = 75			
p = 0.10	n = 300			
p = 0.90	n = 25			
p = 0.90	n = 75			
p = 0.90	n = 300			

8. Are the distributions you examined the exact sampling distributions of the sample proportion $\hat{p}$? Why or why not?

To summarize what we've seen to this point, the sampling distribution of the sample proportion $\hat{p}$ has the following characteristics:

- **Shape:** the shape of the sampling distribution of $\hat{p}$ is sometimes well-approximated by a Normal density curve.
- **Center:** the mean of the sampling distribution of $\hat{p}$ is $\mu_{\hat{p}} = p$.
- **Spread:** for a fixed value of p, the standard deviation $\sigma_{\hat{p}}$ of the sampling distribution of $\hat{p}$ gets smaller as *n* increases.

THE SAMPLING DISTRIBUTION OF A SAMPLE PROPORTION

Suppose we take an SRS of size n from a population of size N with a proportion p of successes. Then if $\hat{p}$ is the sample proportion of successes, the sampling distribution of $\hat{p}$ has these characteristics:

- The **mean** of the sampling distribution of $\hat{p}$ is
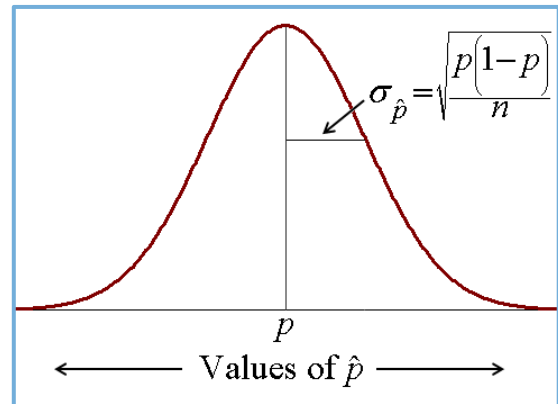 $\mu_{\hat{p}} = p$.

- The **standard deviation** of the sampling distribution of $\hat{p}$ is $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$ when sampling with replacement. **Note**: when sampling without replacement, we must satisfy the 10% condition: $n \leq \dfrac{1}{10}N$.

$$\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$$

p

Values of $\hat{p}$

- The **shape** of the sampling distribution is approximately Normal when $np \geq 10$ and $n(1-p) \geq 10$.

9. According to a recent Harris Interactive poll, 33% of 2236 people that were randomly surveyed agreed that "finding and picking up a penny" was good luck.

   a. Is 0.33 a parameter or a statistic? Explain. What is the appropriate symbol for this value?

   b. Suppose that Mr. Tyson claims that the true proportion of all Americans that "finding and picking up a penny" is good luck is really 30%. What is the appropriate symbol for this value?

   c. Describe the sampling distribution of the sample proportion who think finding/picking up a penny is good luck, for random samples of 2236 Americans.

   d. Based on your answer to part c, would you be surprised to see a sample proportion as high or higher than 0.33? Explain.

   e. Based on your answer to part d, does the sample from the Harris Interactive provides evidence against Mr. Tyson's claim? Explain.

10. Turn to p.462 and complete Exercise 36 in your textbook.

11. A Turn to p.448 and complete Exercise 38 in your textbook.

A MATTER OF CONDITIONS

Remember that two conditions must be met for our claims about the sampling distribution to be correct:

1. ***Normal condition***: the sampling distribution of $\hat{p}$ is approximately Normal only when $np \geq 10$ and $n(1 - p) \geq 10$.
2. ***10% condition***: when sampling without replacement, the sampling distribution of $\hat{p}$ has a standard deviation that is approximately equal to the formula above only if we sample less than 10% of the population.

You may wonder what happens if we violate these conditions. Here are the consequences:

1. ***Violating the Normal condition***: the sampling distribution of $\hat{p}$ will have a shape we do not know and we cannot calculate probabilities with any degree of accuracy.
2. ***Violating the 10% condition***: to calculate the correct standard deviation, we need to adjust the formula we use. That adjusted formula is not given in your text and is not required in AP Statistics. If you would like to know it, you should ask Mr. Tyson.