

## Tuesday, April 14: The Central Limit Theorem

**Group Members:** List the names of your group mates below.

---



---

We've been creating sampling distributions by hand, which is *very* tedious. Today, we're going to work with bigger populations and bigger samples so we'll use an applet. Go to the website:

[http://spark.rstudio.com/minebocek/CLT\\_prop/](http://spark.rstudio.com/minebocek/CLT_prop/)



1. According to <http://nbatattoos.tumblr.com/>, 55% of the 442 players in the NBA had tattoos in the 2013-2014 season.

- a) What is the parameter  $p$  in this problem? Which of the 3 sliders in the upper left of the applet controls the parameter? Set this slider to the appropriate value.
  
- b) At the top is a *relative frequency distribution*. What does the blue bar above the number 1 represent?

- c) Below this are 8 smaller frequency distributions. What do these represent? Are they all the same?
- d) Of the 8 samples shown, which one has the highest estimator  $\hat{p}$ ? What is its value?
- e) What is the green diagram on the bottom?
- f) What is the standard error for this sampling distribution? Where is  $SE$  shown in the diagram?
- g) Calculate the standard error using the formula.

$$SE = \sqrt{\frac{p(1-p)}{n}}$$

(Hint: Remember that  $n$  is the sample size. Where is this shown in the applet?)

How does it compare to the standard error shown in the diagram? Are they the same?

- h) How many samples of size 20 were chosen to create this sampling distribution?

2. Scroll back to the top of the page and decrease the number of samples down to 10.

- a) How would you describe the shape of this sampling distribution? Is it unimodal or multimodal? Symmetric or skewed? Does it have a recognizable shape?
- b) What is the standard error of this sampling distribution?
- c) **Make a prediction:** What do you think happens as we increase the number of samples? How will this affect the shape of the distribution?

3. We're going to slowly increase the number of samples we collect to make our sampling distribution. First, it'll be 10 samples, then 17, then 25, and so on. Fill in the table below as you do this.

Number of samples (each of size 20)	Sampling Error, <i>SE</i> (shown in diagram)	Notes about the shape of the distribution. Symmetric or (left/right) skewed? Unimodal or multimodal? Spreadout or little variation?	$\sqrt{\frac{p(1-p)}{n}}$
10			0.111243
17			0.111243
25			0.111243
41			0.111243
56			0.111243
70			0.111243
101			0.111243
151			0.111243
200			0.111243
450			0.111243
1000			0.111243

What happens as you increase the number of samples??

4. Now, set the number of samples to 100, and set the sample size to 2. Again, we're going to slowly increase the sample size and see what happens. Fill in the table below as you do this.

Sample Size (there are 100 of these samples)	Sampling Error, <i>SE</i> (shown in diagram)	Notes about the shape of the distribution. Symmetric or (left/right) skewed? Unimodal or multimodal? Spreadout or little variation?	$\sqrt{\frac{p(1-p)}{n}}$
2			0.35178
4			0.2487
6			
9			
13			
15			
18			
22			
31			
40			
55			

What happens as you increase the sample size??

The **Central Limit Theorem for Sample Proportions** says...

If each the following conditions are true:

- Condition 1.
  
  
  
- Condition 2.
  
  
  
- Condition 3.

Then the sampling distribution for  $\hat{p}$  is approximately \_\_\_\_\_ with mean \_\_\_\_\_ and standard deviation equal to \_\_\_\_\_, which is given by the formula:

To check condition 1, we have to....

To check condition 2, we have to....

To check condition 3, we have to....

5. Of the 15 players who played for the Detroit Pistons last season, 80% had tattoos. What's the sample size and does this satisfy conditions 1, 2, and 3 above? **Explain your answer.**

6. I gave every player in the NBA a random number and then used a pseudo-random number generator to select a sample of 20 NBA players. 30% of them had tattoos. What's the sample size and does this satisfy conditions 1, 2, and 3 above? **Explain your answer.**

7. I gave every player in the NBA a random number and then used a pseudo-random number generator to select a sample of 27 NBA players. 40% of them had tattoos. What's the sample size and does this satisfy conditions 1, 2, and 3 above? **Explain your answer.**