Objectives:

- To understand the ***Central Limit Theorem (CLT)*** as describing the (approximate) sampling distribution of a sample proportion and of a sample mean.
- To recognize the connection between CLT calculations and the situations that you performed and analyzed in earlier topics.
- To become proficient with using the normal approximation to calculate probabilities of sample statistics falling in given intervals.
- To discover and appreciate the effect of the sample size on the sampling distribution and on relevant probabilities.
- To gain some insight and experience concerning the use of the CLT for drawing inferences about a population parameter based on a sample.

So far we have seen…

We have seen that over the long run sample means fall into a particular pattern. This pattern is directly in line with what we refer to as the normal distribution! Thus we can use our knowledge of the Central Limit Theorem to calculate probabilities that a sample statistic falls within a certain range of a population proportion.

CLT and Sample proportions – sample proportions can be thought of just the same as sample means – therefore the CLT theorem holds for proportions also.

Statistically significant result – A result that is unlikely to occur by random variation alone. The Central Limit Theorem allows us to make probability calculations to assess how “Unlikely” the result is.

Confidence Even if you do not know what the population parameter is – you can be confident that the sample statistic will fall within a certain distance of that unknown parameter value. The CLT allows us to calculate the probability that a sample statistic falls with in a certain distance of the population parameter.

There are conditions that must be true before the CLT may be used with validity.

For sample proportions:

$$n\pi \geq 10 \text{ and } n(1-\pi) \geq 10$$

For example

if $\pi = .25$ then $.25*40 = 10$

therefore ***n*** must be at least ***40*** for the CLT to be valid.

But what if $\theta > .5$ that is lets say that is $\theta = .9$

if $\pi = .9$ then $.9*12 > 10$, but on the other side we have ***$1-\pi = .1$ and $.1 * 100 = 10$***

So, since ***n*** must be at least ***100*** for both conditions to be true we must have an ***n*** of at least 100 for CLT to be valid.

IMPORTANT: If we do not have $n\pi \geq 10$ and $n(1-\pi) \geq 10$ we cannot use the CLT with any validity!

For Sample means:

If the population is itself normal then the CLT is exact

If the population is ***not*** normal then the CLT holds approximately for large sample sizes, as a rule of thumb we use ***$n \geq 30$***

IMPORTANT: If we do not have $n \geq 30$ in a non-normal distribution, we cannot use the CLT with any validity!