## Objectives:

- To understand the purpose of ***confidence intervals*** for estimating population parameters
- To learn how to construct confidence intervals for estimating a population proportion
- To appreciate the distinctions between correct and incorrect interpretations of confidence intervals
- To explore some properties of confidence intervals for a variety of genuine applications and interpret the results

***Statistical Confidence*** Our sample statistic provides a reasonable approximation of our population parameter. It will almost never be exact - but we can make some assumptions based on the statistic and the level of confidence we are looking for. We expect the statistic to be "*in the ballpark*" therefore we can use the statistic (instead of a probably unknown parameter) to get an idea of the value of the true parameter. This value will be in the form of an interval as follows:

The confidence interval is expresses: $\hat{p} \pm z^* \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

the $\hat{p}$ is the sample statistic found using our sample, z* is the critical value from the standard normal distribution found in the z-table for the confidence level we desire.

If you are looking for a *confidence level* 95% you look in the body of the chart for the value that is half of the difference between 1 and the confidence level in decimal form. *i.e.* (1 - .95)/2 = .05/2 = .025 The reason for this is that you are looking for an area under the curve of .95, so you must have equal amounts cut off this area on both sides: .025 cut off on each side leaves .95 under the curve. Once you have found this value in the body of the chart read backwards to the z* value that gives you this .025 which is 1.96 (notice this is very close to 2 as per the empirical rule)

***Statistical intervals***:

- In a 95% confidence interval we mean that we are 95% sure that the population parameter lies in the interval stated.
- If one takes repeated random samples from the population and constructs 95% confidence intervals for each sample, then in the long run 95% of these confidence intervals will capture the true proportion.
- We must realize that the interval is based on the randomness of the sample not on the parameter. The parameter must be stable but it can not be found directly normally - so we make an assumption based on our samples
- **It is incorrect to say that the probability is .95 of falling within the 95% confidence interval**. It is not a probability because θ is a fixed non-random, probabilities are used for instances when you are trying to predict the outcome of a random event
- The confidence interval predicts a parameter, it **does not** make any inference on the value of any particular value or any sample
- We use the same rule of thumb as we did for CLT that is:
  $\hat{p}(n) \geq 10$ and $(1-\hat{p})(n) \geq 10$

***Standard Error:*** When we do not know the standard deviation, since we usually do not know $\theta$, we estimate it as $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ and we call it the standard error instead of the standard deviation. (notice that we are basically allowing the sample proportion to approximate $\theta$ and so we don't confuse it with the s.d. we call it the standard error.0

***It is important to understand that we treat the standard error much like we do the standard deviation –*** we will simply be working backwards.

***Critical Values*** - The Z-score (denoted z*) that gives us the proper Confidence level we are seeking. If we want 95% confidence then we need to have a Z-score of 1.96. We start by selecting how confident we want (or need) to be and then find the Z-score that gives this to us.

***Common Confidence levels and their Critical values***:

| **Confidence level** | **80%** | **90%** | **95%** | **99%** | **99.9%** |
|---|---|---|---|---|---|
| Area Left | .90 | .95 | .975 | .995 | .9995 |
| Area Right | .10 | .05 | .025 | .005 | .0005 |
| Critical Value (z*) | 1.282 | 1.645 | 1.960 | 2.576 | 3.291 |

***These three points are extremely important to know!!!***

******* *As one increases the confidence level desired, the width of the confidence interval increases.*

******* A Larger sample size produces a narrower confidence interval whenever other factors remain the same.

*** If a sample is Biased then results from confidence intervals can be very misleading and definitely invalid

Using your calculator to find a confidence interval for a proportion:

Go to STAT : TESTS : 1 – PropZInt

Enter the number in sample that satisfied condition for **X**

Enter the total number in the sample for **n**

Enter the confidence level you are using for **C-Level**

***Interval*** The range of values from the lower to the upper bound

***Midpoint*** The center of the interval ( $\hat{p}$ )

***Half width*** The distance from the midpoint to either bound (or half the distance between the upper and lower bound)