

## Confidence Intervals

Standard Deviation of Sampling Distributions

Sample means

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

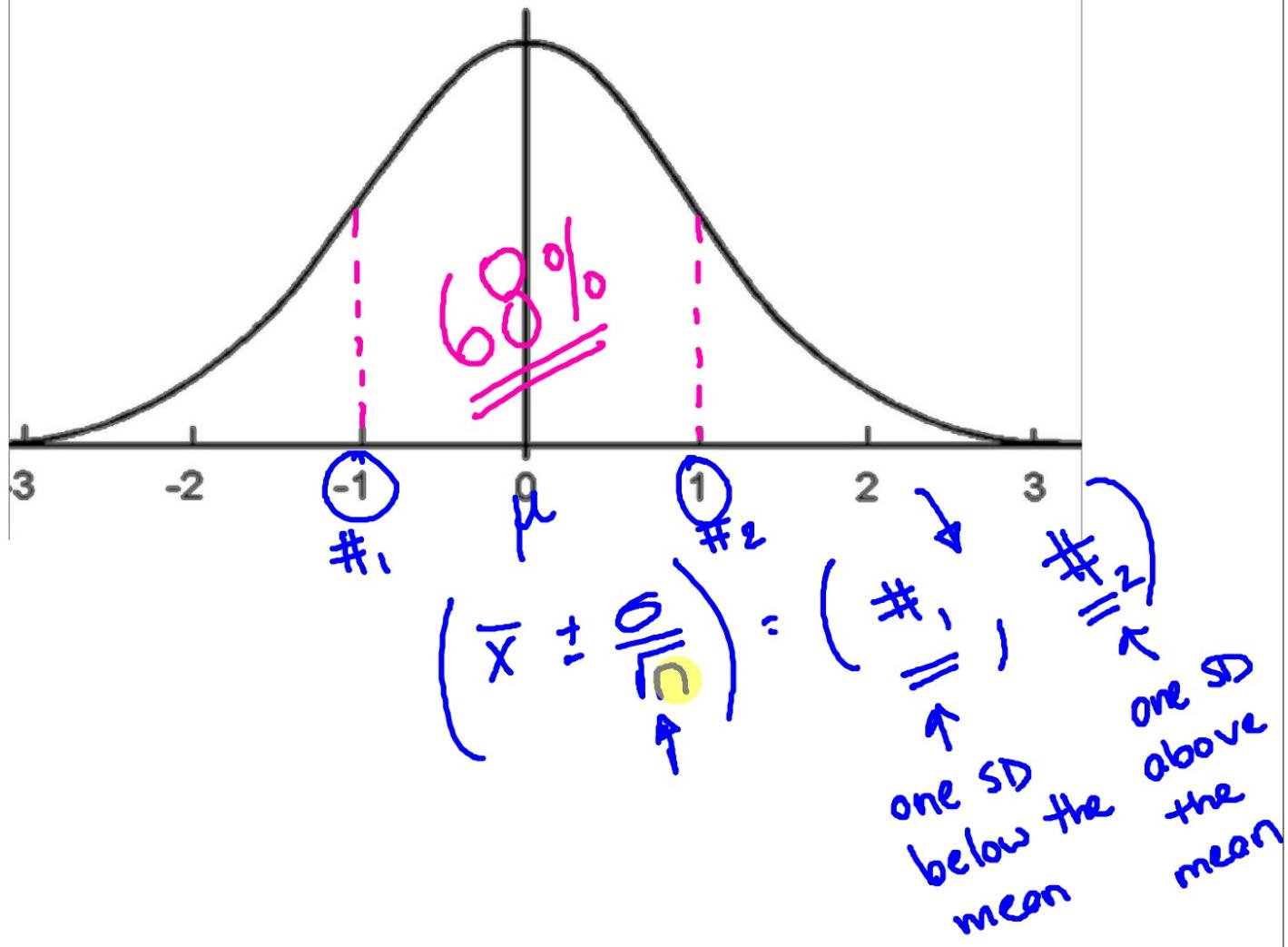
Sample Proportions

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

standard ERROR (when we use statistics to approximate parameters)

$$SE_{\bar{x}} = \frac{s}{\sqrt{n}}$$

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

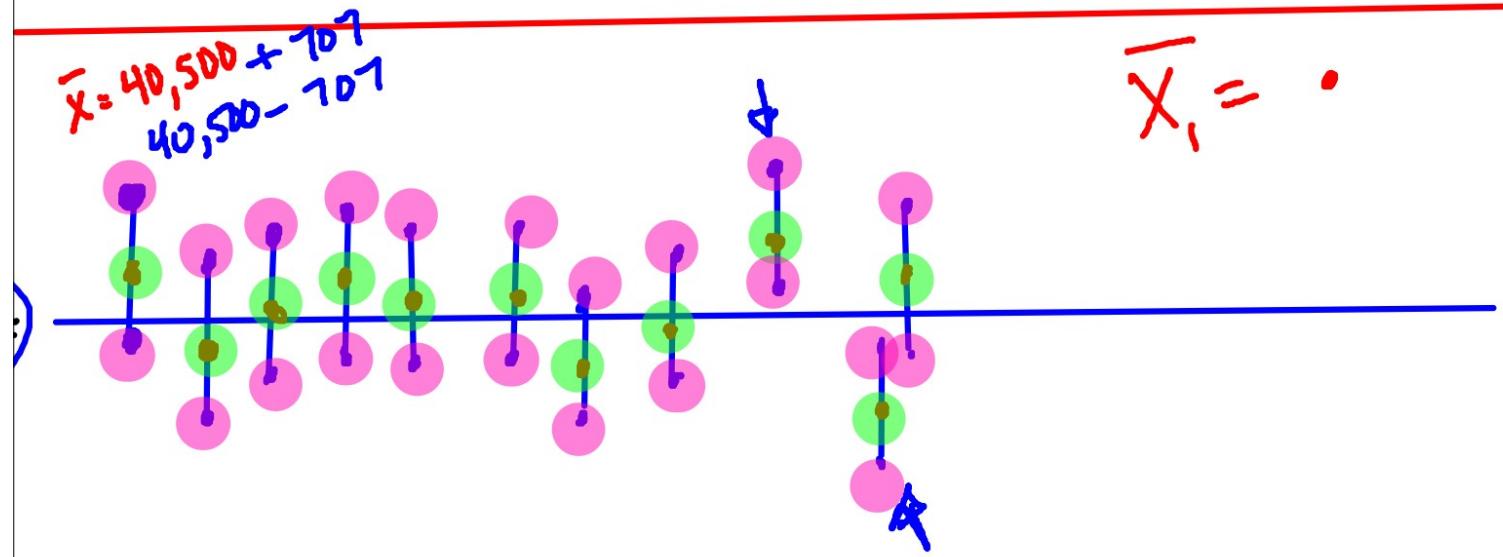


When we create an interval we are saying:

"We are \_\_\_ % confident that the (unknown parameter) (mean of the population or proportion of the population) is somewhere between \_\_\_ & \_\_\_ .  
A ↑  
confidence interval

To become more confident in your interval  
you can:

- 1) Use a wider (larger) interval
- 2) Collect more data "increase"  $\underline{n}$ .



What we mean by #% confident  
If we were to create many  
different intervals using different  
samples, then #% of intervals  
would contain the unknown parameter.

Create a 95% confidence interval

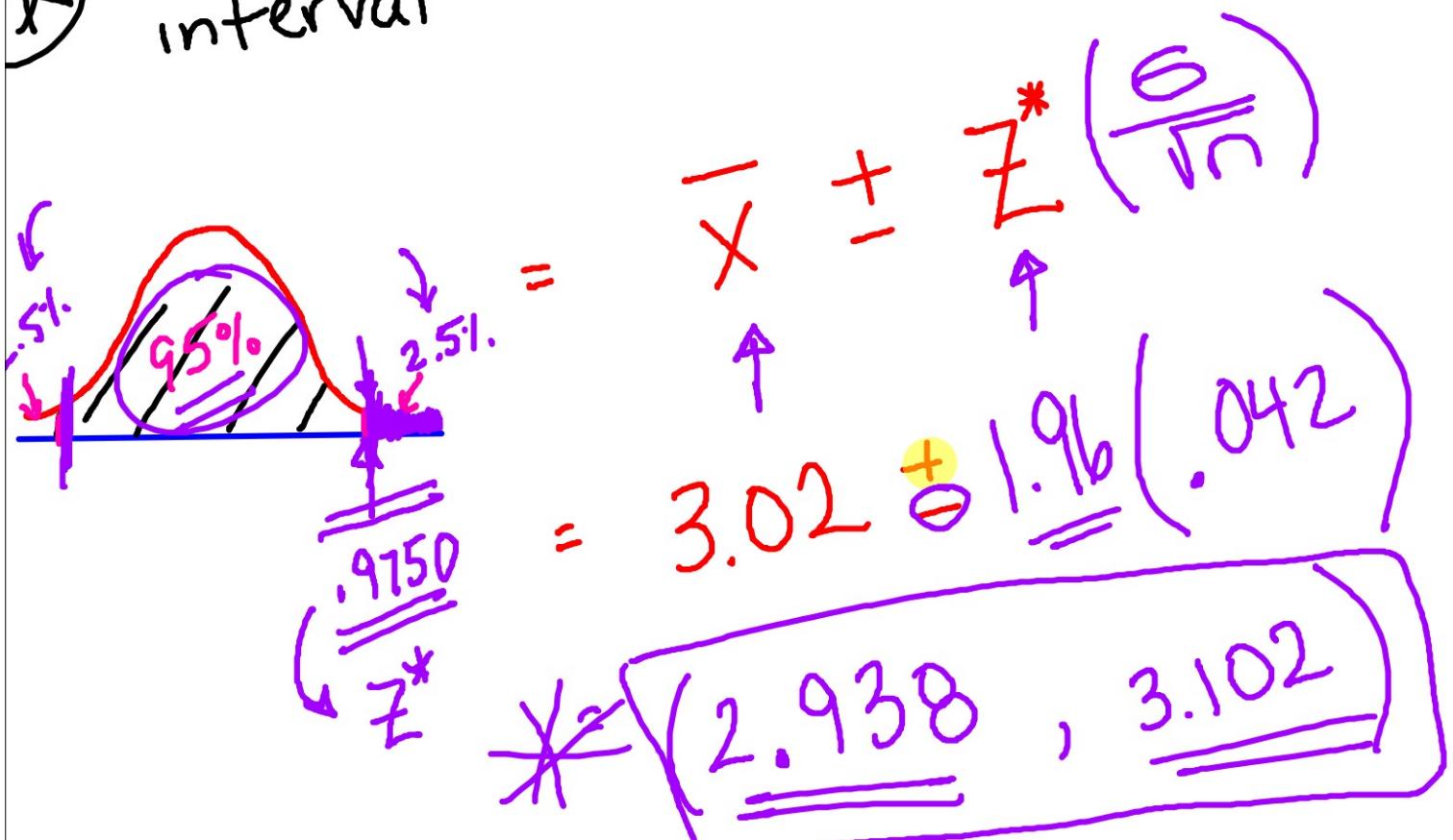
Avg. G.P.A at CCBC

- Random Sample of 100 students and we calculate their mean to be 3.02. It is known that the standard deviation for all students at CCBC is .42.

$$*\quad s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{.42}{\sqrt{100}} = .042$$

- Random Sample ✓
- $10n < \text{Pop. Size}$  ✓
- $n \geq 30$  ✓

\* Confidence interval = statistic  $\pm$  (critical value) ( $SD$  of statistic)



2.938, 3.102

\* Interpret the confidence interval \*

\* { We are 95% confident that the  
average G.P.A at CCBC is somewhere  
between 2.938 and 3.102.

## Confidence Intervals with Proportions

\* Formula:

$$\hat{p} \pm z^* \left( \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

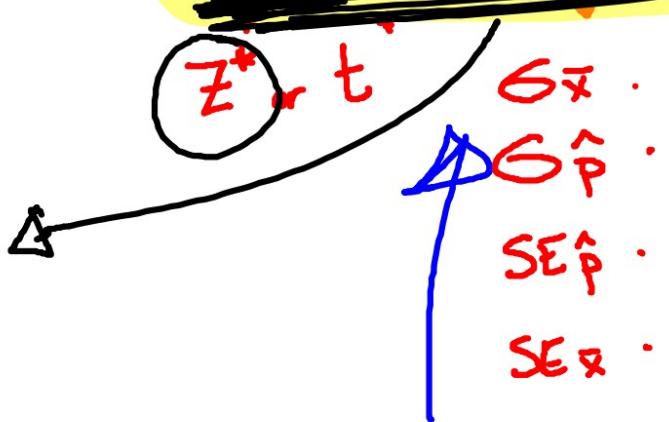
SE $\hat{p}$

\* Proportions

ALWAYS use  $z^*$

$$\text{Confidence Interval} = \text{statistic} \pm (\text{critical value}) (\text{SD of the statistic})$$

( $\bar{x}$  or  $\hat{p}$ )



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

Margin of Error

The countries of Europe report that 46% of the labor force is female. The United Nations wonders if the percentage of females in the labor force is the same in the United States. Representatives from the United States Department of Labor plan to check a random sample of over 10,000 employment records on file to estimate a percentage of females in the United States labor force.

1. The representatives from the Department of Labor want to estimate a percentage of females in the United States labor force to within  $\pm 5\%$ , with 90% confidence. How many employment records should they sample?

$$ME = Z \left( \sqrt{\frac{P(1-P)}{n}} \right)$$

$\leq .05$

$P = .46$   
 $Z \approx 1.645$   
 $* ME \text{ less than or } = 5\%$

$1.645 \left( \sqrt{\frac{.46(.54)}{n}} \right) \leq .05$

$ME$

$$\frac{1.645 \cdot \left( \sqrt{\frac{.46(.54)}{n}} \right) \leq .05}{1.645}$$

$$\sqrt{\frac{.46(.54)}{n}} \leq \left( \frac{.05}{1.645} \right)^2$$

$$A: \frac{.46(.54)}{n} \leq .0009239(n)$$

$$\frac{(.46)(.54)}{.0009239} \leq \frac{.0009239(n)}{.0009239}$$

Sample size must be at least 269

sample size  
↓  
 $268.9 \leq n$

The countries of Europe report that 46% of the labor force is female. The United Nations wonders if the percentage of females in the labor force is the same in the United States. Representatives from the United States Department of Labor plan to check a random sample of over 10,000 employment records on file to estimate a percentage of females in the United States labor force.

3. Interpret the confidence interval in this context.
1. The representatives from the Department of Labor want to estimate a percentage of females in the United States labor force to within  $\pm 5\%$ , with 90% confidence. How many employment records should they sample?

\* 2 They actually select a random sample of 525 employment records, and find that 229 of the people are females. Create the confidence interval

Random Samp ✓       $\text{CI} = \hat{P} \pm Z \left( \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \right)$        $n = 525 *$   
 10n < Pop. Size ✓       $\hat{P} = \frac{229}{525} *$        $(.4006, .4718) *$        $(1)$   
 1.  $np \geq 10$   
 2.  $n(1-p) \geq 10$

we are 90% confident that the proportion of the labor force in the U.S that is female is between .4006 & .4718.

## Parts of Confidence Interval

- 1. Conditions ✓
- 2. Type of Interval ✓
- 3. Formula & Statistics ✓
- 4. Interval (From STAT CRUNCH) ✓
- 5. Interpret Interval ✓

## t - Distributions (means) Quantitative

- (\*) We use "t" distributions when we do not know the standard deviation of the population. ( $\sigma$ ). We use "s" the standard deviation of the sample to approximate  $\sigma$ . ( $SE_{\bar{x}}$  instead of  $SE_x$ )

(Quantitative)

$\bar{x}$

Difference between  
one-sample  $Z$  & One-Sample T

$\sigma$  is known  
SD of pop.

$\sigma$  is unknown  
SD of sample (s)

## Formulas

\* 1-Proportion :  $\hat{P} \pm Z \left( \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \right)$  Categorical Data  
 $(Z)$  Interval

\* 1-Sample  $Z$  :  $\bar{X} \pm Z \left( \frac{s}{\sqrt{n}} \right)$  Quantitative sample means

\* 1-Sample  $t$  :  $\bar{X} \pm t \left( \frac{s}{\sqrt{n}} \right)$

1) A company that produces batteries is concerned about the distribution of the life expectancy of their batteries. The company takes a simple random sample of 50 batteries and computes the sample mean to be 800 hours per battery and a standard deviation of 25 hours. Construct and interpret a 90% confidence interval for the unknown mean life expectancy.

1-Sample T Interval ✓

$$D.F (n - 1) \rightarrow$$

$$n = 50$$

- Random Sample ✓
- $10n < \text{Pop. Size}$  ✓
- .  $n \geq 30$  ✓

$$\bar{x} \pm t^* \left( \frac{s}{\sqrt{n}} \right)$$

$$(794.0725 \quad 805.9275)$$

$$\textcircled{4}$$

$$\bar{x} = 800$$

$$\textcircled{5} s = 25$$

90% CI

We are 90% confident the mean life expectancy of All batteries produced by a company is somewhere between 794.0725 hours & 805.9275 hours.

1) In a simple random sample of 200 students taken at a very large university, 64 stated that they enjoy strawberry ice cream. Construct and interpret a 99% confidence interval for the percent of students at the university who enjoy strawberry ice cream.

### 1-Proportion Z Interval

- Random Sample ✓
- $10n < \text{Pop. Size}$  ✓
- $n\hat{p} \geq 10$  ✓

$$n(1-\hat{p}) \geq 10$$

\* Since we do not know " $p$ " we use  $\hat{p}$

$$n = 200 \quad \text{99% CI}$$

$$\hat{p} = \frac{64}{200} \Rightarrow \frac{\# \text{ of successes}}{\# \text{ of observations}}$$

$$\hat{p} \pm Z \left( \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)$$

$$(0.23503667 \quad 0.40496333) \quad (.78, .80)$$

We are 99% confident that the actual proportion of students at the university who enjoy strawberry ice cream is somewhere between .235 & .405.

