chapter 8 Confidence Intervals

We finally have the tool set to infer our descriptive statistics onto our population

8.1 Conf. Int'vl for a population mean {S.D. is known}

THE SET-UP PROBLEM

Ex A simple random sample of 100 fourth graders is selected to take part in a reading program.

- At the end of the program the mean score was $\bar{x} = 67.30$ wpm* We know from previous testing that the standard deviation is $\sigma = 15$ wpm.

Q: We now want to be able to apply our sample results to the pop. of all 4th graders in the state.

A: T.B.D., first some basics...

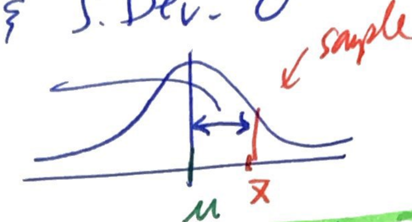
Def A point estimate is a single number used to estimate the value of an unknown parameter of a population: Typically a samples' statistic.

It is very unlikely that the point est. $\bar{x}$ is exactly equal to the population mean due to the random nature of sampling.

* wpm = words per minute read

- Recall the z-score was a number assigned to a data point from a populat'n that possess a mean μ & S. Dev. σ

$$\boxed{Z = \frac{\bar{X} - \mu}{\sigma}}$$



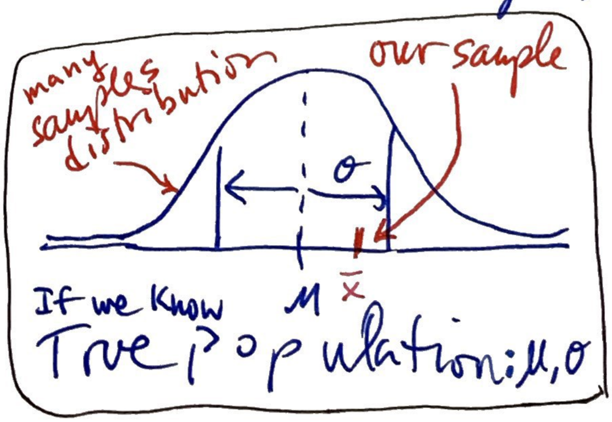
it tells us how many std. dev. the sample is from the population parameter

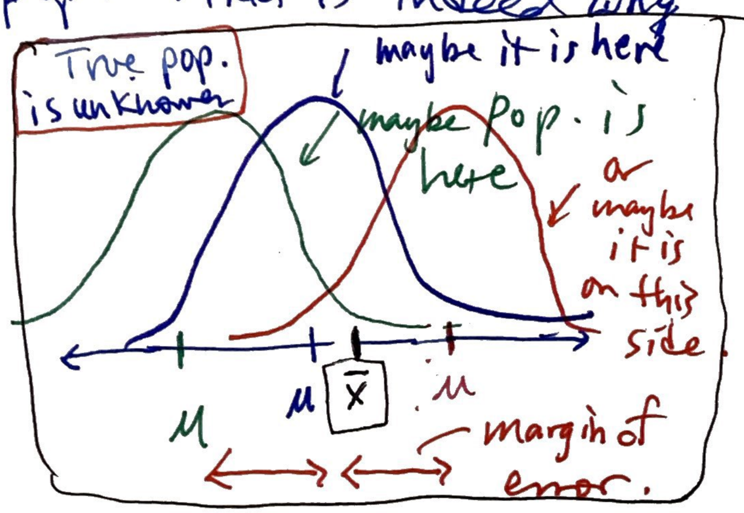
ex: $z = 1.8$ means that the sample value is one and one-half S. Dev. away from the pop. mean.

we can solve for μ to get

$$\mu = \bar{X} - \sigma Z$$

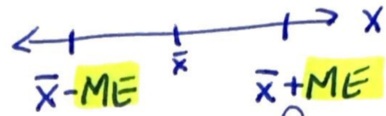
- When we sample we do not know the true value of the pop. That is indeed why we seek samples.





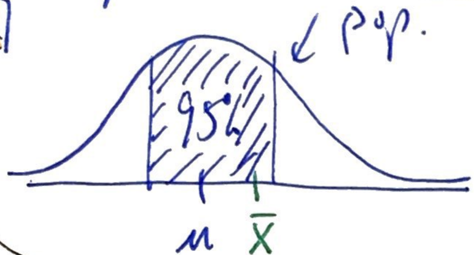
The plus or minus nature of the difference between the sample and the populatn is called the Margin of Error

$$\mu = \bar{X} \pm ME$$



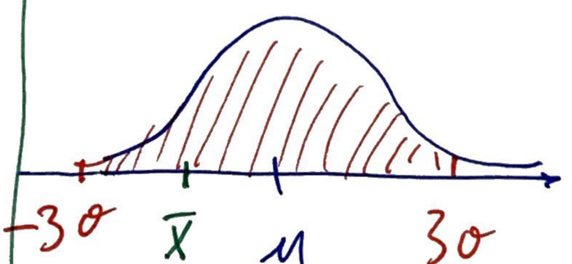
• The C.L. Thm tells us that, for a group of size "n", we can be assured that our sample falls between certain limits.

Ex

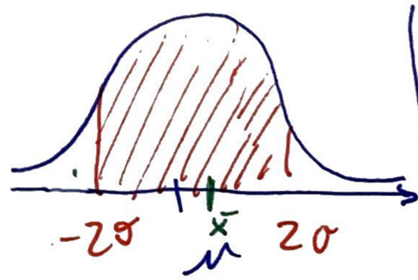


We are "confident" that our sample falls between ($\bar{X} - ME$ to $\bar{X} + ME$)

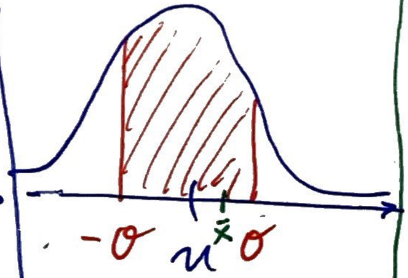
We call these regions "the Confidence Interval." They are based on a Confidence Level, which are typically 90%, 95% or maybe 98%. • If we did know the pop. parameter μ...



then we are 98% conf. that our sample mean $\bar{x}$ is with $\mu \pm 3\sigma$



then we are 95% conf. that $\bar{X}$ is within $\mu \pm 2\sigma$



then we are 68% conf. that $\bar{X}$ is within $\mu \pm \sigma$

Now for sample means, if the std. dev. is assumed known, we have the model

- Margin of Error = $\boxed{Z_c} \cdot \boxed{\text{Standard Error}}$

where Standard Error = $\dfrac{\sigma_{pop}}{\sqrt{n}}$

and Z_c = is the critical value associated with a desired confidence Interval → see next page

EX Cont.

If, for the reading method, we determined that $\bar{x} = 67.30$ wpm, and we know from previous testing that our std. dev. of the pop is historically $\sigma = 15$ words per minute,

then $S.E. = \dfrac{15}{\sqrt{100}} = 1.5$ words per minute

- Lets say we desire 95% of samples (of size 100) to produce a confidence interval that captures the population mean, then $Z_c = 1.96$ and so $ME = (1.96)(1.5) = 2.94$ wpm

This is to say, "there is a 95% chance that the interval we constructed captured the true population parameter"

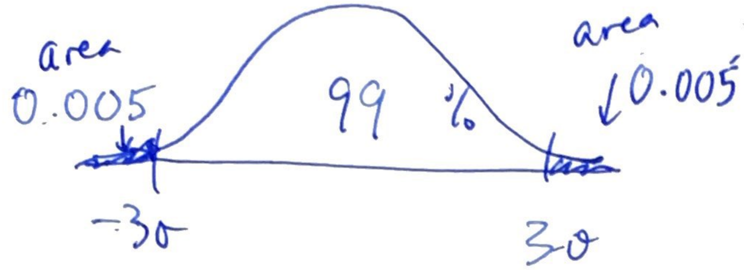
- Next page shows various values of Z_c corresponding to various confidence Levels

ex: 95% is $\pm 2\sigma$ → z-tables tell us...

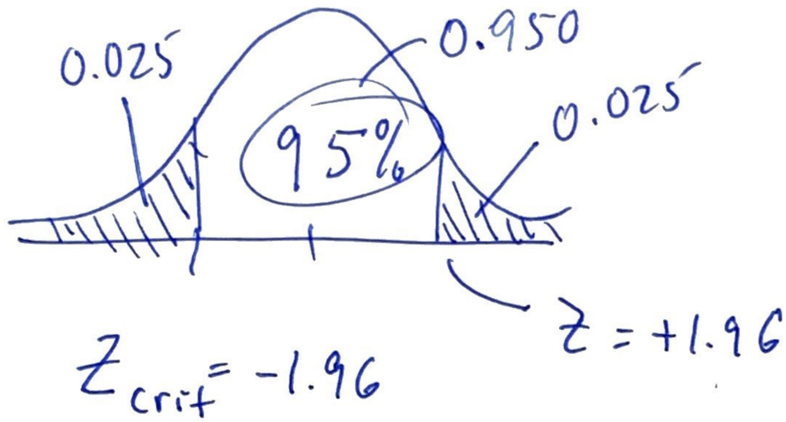
NOTE:

Conf. Levels

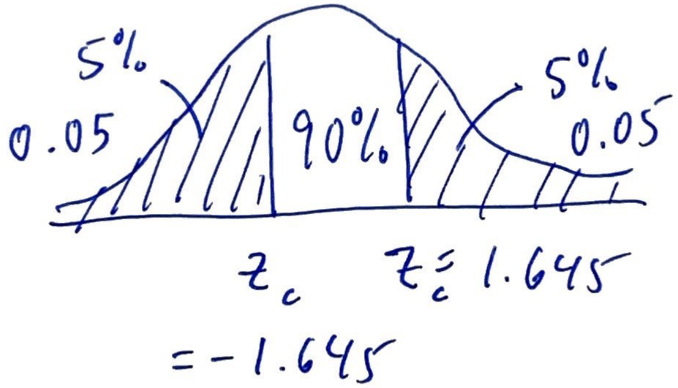
Critical value



$$\boxed{\begin{array}{c} 99\ \% \\ Z = 2.575 \end{array}}$$



$$\boxed{\begin{array}{c} 95\% \\ Z = 1.96 \end{array}}$$



$$\boxed{\begin{array}{c} 90\% \\ Z = 1.645 \end{array}}$$

EX Completing to construction of the C. Intvl

$$\boxed{SE = \frac{15}{\sqrt{100}} = 1.5}$$

$$ME = Z_c \cdot SE$$

$ME = (1.96)(1.5)$ for 95% conf.

$$\boxed{ME = 2.94}$$

...

EX Cont.

- We now form the confidence Interval from

$$\boxed{\bar{x} \pm ME}$$

Notation: $\bar{x} \pm ME$ means that

Conf. Interval.

$$\boxed{\bar{x} - ME < \mu < \bar{x} + ME}$$

So here 67.30 ± 2.94

becomes $67.3 - 2.94 < \mu < 67.3 + 2.94$

or simply $\boxed{64.36_{wpm} < \mu < 70.24_{wpm}}$ C. Int'vl

Statement: We are 95% confident that the population of 4th graders in Santa Clarita have a reading speed of between 64.36 wpm to 70.24 wpm

What this means is that samples of size 100 will produce a confidence Interval that, 95% of the time, captures the true population parameter.

- It is possible that our sample is further from the population parameter that the stated interval, this could happen 5% of the time.

Steps for Confidence Intervals

Step 0. Are the $\boxed{\text{condition met}}$ {C.L.Thm}

- check "assumptions and SE formulas" from mymathmantra.com → statistics reserve

Step 1: Calculate (or state) $\boxed{\bar{x}\text{, the point statistic.}}$

Step 2: Find the $\boxed{\text{critical value}}$ associated with the $\boxed{\text{desired confidence}}$ Level

Step 3: Calculate $\boxed{S.E:}$ $\frac{\sigma_{pop}}{\sqrt{n}}$

Step 4: Calculate the $\boxed{M.E} = z_c \cdot S.E.$

Step 5: Form the $\boxed{\text{confidence Interval}}$

$$\boxed{\bar{x} - M.E. < \mu_{pop} < \bar{x} + M.E.}$$

Step 6: state / interpret $\boxed{\text{the results}}$.

(step 7: Verify on statdisk.com → Analysis
↳ Conf. Intvl)

EX Reading Program (Using the Steps this time but use Conf. Level of 98%)

The reading program produced $\bar{x} = 67.30$ with the historical S.Dev. of 15 wpm. for 100 students.
Calculate a confidence interval for 98% Conf. Level

- Step 1. State $\bar{x} = 67.30$
- Step 2. C. Level is 98%, so $Z_c = 2.326$
- Step 3. $SE = \frac{\sigma_{pop}}{\sqrt{n}} = \frac{15}{\sqrt{100}} = 1.5$
- Step 4. $ME = z_c \cdot SE = (2.326)(1.5) = 3.489$
- Step 5: Conf. Intvl:

$$\bar{x} - ME < \mu_{pop} < \bar{x} + ME$$

$$67.30 - 3.489 < \mu_{pop} < 67.30 + 3.489$$

$$\boxed{63.81\text{ wpm} < \mu < 70.79\text{ wpm}}$$

- Step 6. Interpret the result:

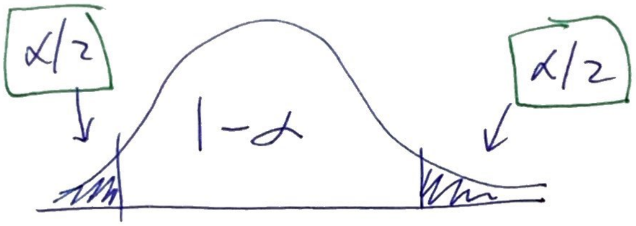
"We are 98% confident that the mean reading speed of 4th graders in SCV using the new method falls between 63.81 & 70.79 words per minute"

98% confidence mean that 98% of samples of size 100 will produce population conf. Intv'ls that contain the true parameter.

about Confidence levels ...

In general if we desire that our constructed Confidence level be R % successful at capturing the populations true para we need to find a Z-score that captures (100−R)/2 area in the wings.

- Lets let α = area total in the two wings
- Then



Z_c or $\boxed{Z_{\alpha/2}}$ as some authors use.

↑ critical value

EX If Conf. Level is to be 70% then

$\alpha = 30\%$ or 0.30 so $\alpha/2 = 0.15$

- The z-tables show that $Z = 1.04$ separates the lowest 15% of area from the upper.

$$\boxed{Z_c = Z_{\alpha/2} = 1.04}$$

this To be used in the $ME = Z_c SE$ calculation

Some critical z-scores...

Confidence Level	Critical Value
70%	$z_c = 1.04$
90%	$z_c = 1.645$
95%	$z_c = 1.96$
98%	$z_c = 2.326$
99%	$z_c = 2.576$

* If you desire to have a larger Conf. Level you will need to have a wider margin of Error, which is usually not desireable.

To remedy the situation you will need larger sample sizes.

Recall $ME = z_c \frac{\sigma}{\sqrt{n}}$ so $\boxed{n = \left(\frac{z_c \cdot \sigma}{ME} \right)^2}$

EX Say we want a ME of 2 wpm in our reading method Conf. Intvl. Then for a 95% Conf Level

$$n = \left(\frac{(1.96) \cdot 15\,\text{wpm}}{\pm 2\,\text{wpm}} \right)^2 = 216.09$$

So a sample of $\boxed{n = 217}$ will produce a ME = 2.0 @ a Conf Level of 95%... go get more data!! add to the existing n = 100 sample values...

EX A machine fills Cheerio boxes to 20oz. A simple random sample of 6 boxes is found to have a mean of 20.25 oz. Historically the machine has a std. dev of $\sigma_{pop} = 0.2$ oz.

Task: Construct a 90% C. Int. for the mean fill value.

- Step 0: To Be determined
- Step 1: point estimate: $\begin{cases} \bar{x} = 20.25 \text{ oz} \\ \sigma_{pop} = 0.2 \text{ oz} \end{cases}$
- Step 2: $Z_c = 1.645$ for 90% C. Intvl
- Step 3: S.E: $\sigma_{pop}/\sqrt{n} = 0.2/\sqrt{6} = 0.08166$
- Step 4:

$$\begin{aligned} M.E. &= Z_{crit} \cdot SE \\ &= (1.645)(0.08166) \\ &= 0.1343 \end{aligned}$$

- Step 5: C. Intvl

$$20.25 - 0.1343 \leq \mu_{pop} \leq 20.25 + 0.1343$$

$$\boxed{20.01 \leq \mu_{pop} \leq 20.38} \text{ oz}$$

- Step 6: We are 90% confident that the mean weight of the 20oz boxes is actually between 20.01 and 20.38 oz.

90% of s of size 6 will within this Intvl

* Conditions to be met before we use the steps to build the Conf. Intvl.

Assumptions to be met to construct a Conf. Intvl for μ when σ is known

1. Simple Random Sampling (SRS)
2. Sample size is $n > 3$

-OR-

examining the samples shape (histogram or dot plot or box plot) shows that the data is unimodal & symmetric

If these conditions are not met stop!

- You may Proceed but call out the violation(s)

MATH ______ Statistics Confidence Interval Worksheet Name ____________

STEP 0: **(a) Type of problem (circle the line or part therein)**

- 1- pop | 2 pop for proportion (z-table)
- 1- pop | 2 pop for means (t-table)

(b) Assumptions (state the general and justify your application's)

- ________________ ____________________________
- ________________ ____________________________
- ________________ ____________________________
- ________________ ____________________________

STEP 1: **Compute the point estimate**

STEP 2: **(a) State the Confidence Level:** ____
(b) Find the corresponding critical value from the tables

row ____________ column ____________

critical value (circle one): z t = ______________

STEP 3: **Compute the standard error.**

(a) Formula SE: $\sqrt{\frac{\hat{p}\hat{q}}{n}}$ $\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}$ $\frac{s}{\sqrt{n}}$ $\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ $\frac{s_d}{\sqrt{n}}$

SE Value = ____________________

STEP 4: **Compute the Margin of Error = critical value * SE**

ME = ____________ * ____________ = ____________

STEP 5: **Construct the Confidence Interval: point estimate ± ME**

________ **-** ________ **<** **<** ________ **+** ________

STEP 6: **Interpret the results**