Meeting 8 - 3/25/2020

Standard Normal Dist'n

$Norm(\mu = 0, \sigma = 1)$

$$f(x \mid \mu = 0, \sigma = 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Central Limit Theorem

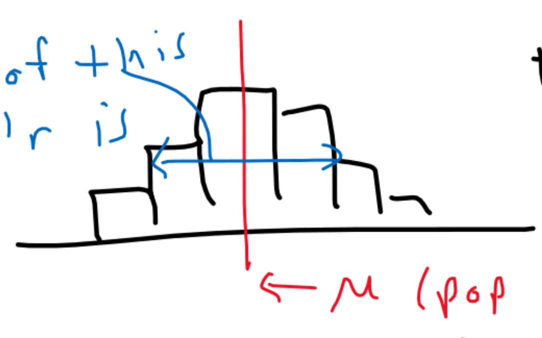
prob dist of means of samples from any population is normally distributed

→ the mean of a large # of sample means is approximately the population mean

Side Bar
- population - all of the things in our group of interest.
- sample - subset of those things that we observe.

Dist'n of sample means

SD of this dist'n is the SE.



Hist of sample means

← μ (pop mean)

standard error of the mean

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

← population s.d.

← sample size

$$S_{\bar{y}} = \frac{s}{\sqrt{n}}$$

← sample s.d.

Standard Error "tells us about the error in using $\bar{Y}$ to estimate μ (population mean)"

Confidence Intervals

can transform any normal to a standard normal using a z-transformation

$$Z_i = \frac{Y_i - \mu}{\sigma}$$

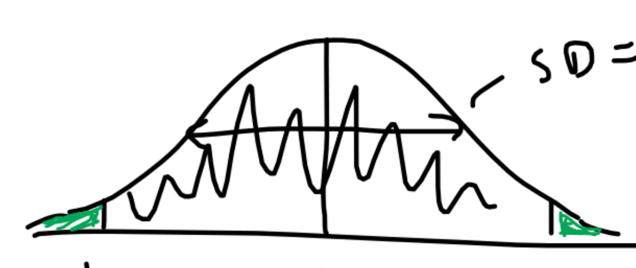
z-scores

Apply this to dist'n of sample means.

$$Z = \frac{\bar{Y} - \mu}{\sigma_{\bar{Y}}}$$

why?

↳ we can say something about the probability of observing specific ranges of values



Can say something about How confident we are that a sample mean represents the true mean.

$$P(\bar{Y} - 1.96\,\sigma_{\bar{Y}} \leq \mu \leq \bar{Y} + 1.96\,\sigma_{\bar{Y}}) = 0.95$$

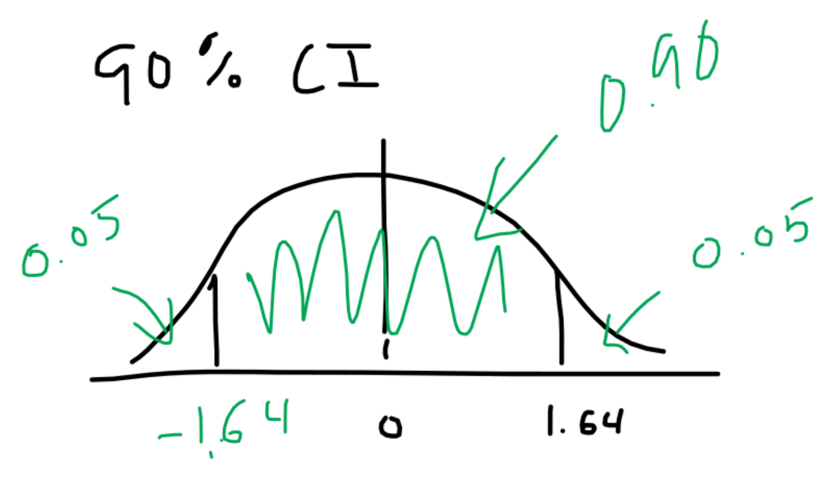
formula for 95% CI

$\bar{Y}$ = sample mean

$\sigma_{\bar{Y}}$ = population SE

± 1.96 = critical values (95%)

90% CI



Reality Check

dont usually know σ, $\sigma_{\bar{Y}}$

use sample sd as an estimate

$s \sim \sigma$

using estimate, no longer using a standard normal, but rather, we use a t-distribution.

?dt ⇒ what are the parameters t

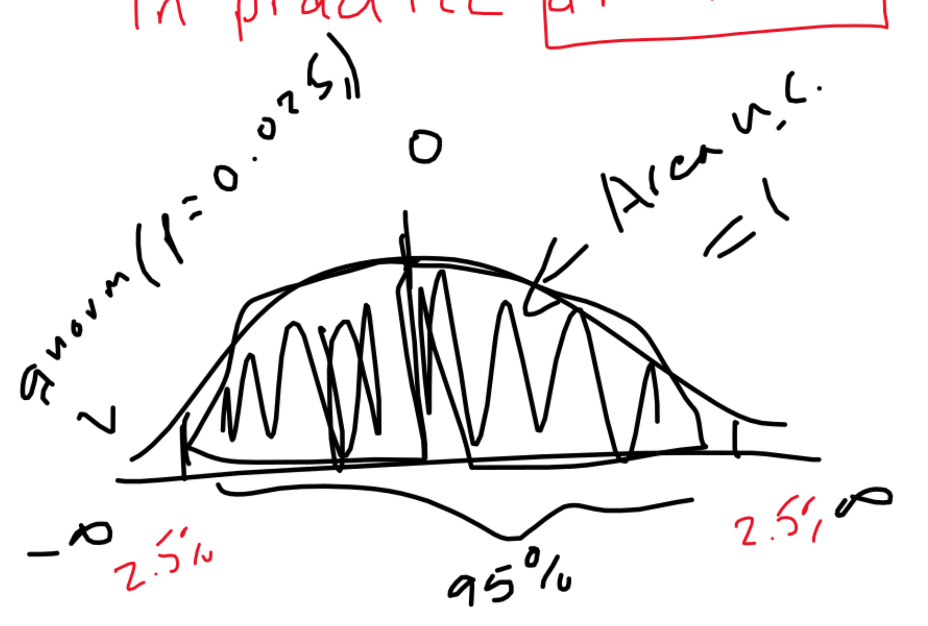
Parameters of n r al? mean, sl

Parameter for t-dist = degrees of freedom

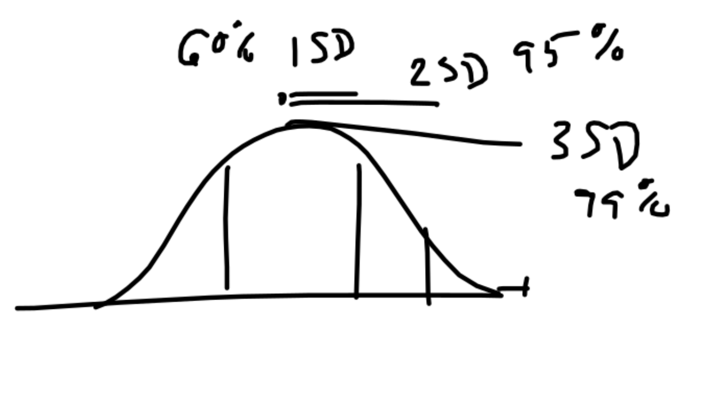
"the number of observations that are "free to vary" when estimati a statistic"

$\{1, 2, 3\} \Rightarrow$ mean = 2

in practice $\boxed{df = n - 1}$



60-95-99 Rule



Z-score is essentially the distance in standard deviations that a point is from a mean value.