Chapter 9: Estimating the Value of a Parameter

Section 9.2: Estimating a Population Mean

INTRO

We want to estimate the true population mean, μ. We can do the same thing we did for proportions and get: $\mu \pm E$, where $E = z_{\alpha/2}\sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$. Since the standard error for proportion is $\sigma_{\hat{p}} = \sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$, and the standard error for the mean is, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, we might guess that the error for the mean is $\boxed{E = z_{\alpha/2}\cdot\frac{\sigma}{\sqrt{n}}}$. $\sigma \to s?$

(ch8)

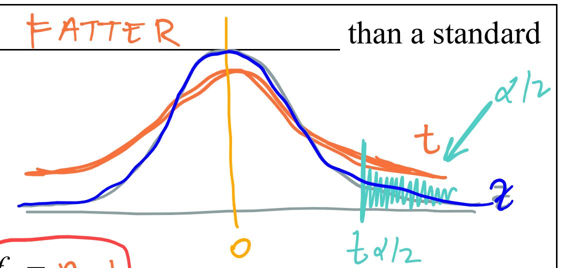
There's two problems with this approach. 1) We don't know the mean, μ, so we also don't know σ! 2) We could replace σ by s but, unfortunately, there's still a problem! The problem is with $z_{\alpha/2}$!

We need a new model to help us estimate CIs for the mean. We will need to use a different distribution.

STUDENT T DISTRIBUTION

(VIDEO: https://www.youtube.com/watch?v=32CuxWdOlow)

STUDENT'S t- DISTRIBUTION RULES	• BELL - shaped but the tails will be FATTER than a standard normal curve. • Centered around $\mu_t =$ 0 and $\sigma_t = \frac{s}{\sqrt{n}}$ • Characterized by the *degree of freedom* FORMULA: $\boxed{df = n-1}$ • We change $\boldsymbol{z_{\frac{\alpha}{2}}}$ to * $t_{\alpha/2}$



Rounding Rule for $t_{\alpha/2}$: 2 decimal places (same as for z-scores)

Calculator
tcdf → prob
invT → t scores

CONFIDENCE INTERVAL FOR THE POPULATION MEAN

CI: (Alternative Forms) $\bar{x} - E < \mu < \bar{x} + E$ or $\bar{x} \pm E$ or $\boxed{(\bar{x} - E,\ \bar{x} + E)}$

Point Estimate: $\bar{x} = \frac{\sum x}{n}$ (1-VAR STATS)

Critical Value: $t_{\alpha/2} =$ invT(α/2, df)

Error: $E = t_{\alpha/2}\cdot\frac{s}{\sqrt{n}}$

←distribution used for critical value: t-dist

$t_{\alpha/2} = \text{invT}(1-\alpha/2,\ df)$

NOTE invT is always Left Tailed

Requirements

1. The sample is a simple random sample (SRS)
2. The sample is independent or $n \leq 0.05N$ (i.e. less than 5% of population size)
3. The value of the population standard deviation σ is not known.

* 3. Either or both of the given conditions are satisfied: { The population is normally distributed or $\boxed{n > 30}$ }

To create these confidence intervals, we won't have σ but we can get a sample standard deviation S. The larger the sample, the closer the sample s will get you to σ by the CLT. (Ch8)

Ex1: Find the critical t –value that corresponds to 98% confidence with 15 degrees of freedom.

df

t-dist

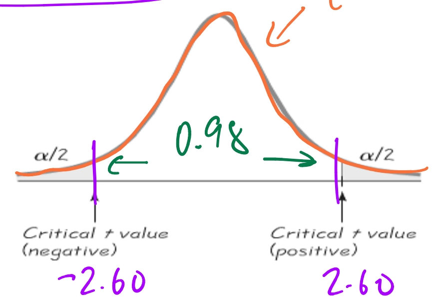
$t_{\alpha/2} = t_{0.01} = |\text{invT}(0.01, 15)| = |-2.60|$

$= \boxed{2.60}$

$CL = 0.98$

$\alpha = 1 - CL = 0.02$

$\alpha/2 = 0.01$



Def ***Point Estimate*** A single value used to approximate a population parameter.

Note: The sample mean $\bar{x}$ is the point estimate of the population mean μ.

SAMPLE MEAN: $\bar{x} = \dfrac{\sum x}{n}$

Def ***Margin of Error*** The maximum likely difference between the observed sample mean $\bar{x}$ and the true value of the population mean μ.

MARGIN OF ERROR FORMULA: $\boxed{E = t_{\alpha/2} \cdot \dfrac{s}{\sqrt{n}}}$

USE t-dist

$n = 34$ $df = 33$

mean

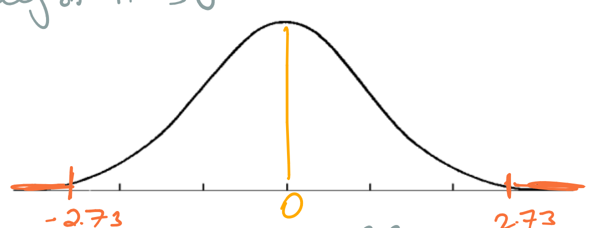
Ex2: A 2017 sample of 34 randomly selected cities in the US found that the average cost of a wedding was \$36,000 with a standard deviation of \$14,114. Determine a **99% confidence interval** for the corresponding population mean cost of a US wedding. Assume that wedding costs are normally distributed.

Check requirements

① SRS ② Indep ③ σ unknown ④ normally dist or n>30

Identify point estimate (sample mean)

$\boxed{\bar{x} = \$36000}$



Determine critical value $t_{\alpha/2}$

$t_{\alpha/2} = t_{0.005} = |\text{invT}(\underset{\alpha/2}{0.005}, \underset{df}{33})| = \boxed{2.73}$

$CL = 0.99$

$\alpha = 1 - CL = 0.01$

$\alpha/2 = 0.005$

Find margin of error

$E = t_{\alpha/2} \cdot \dfrac{s}{\sqrt{n}} = (2.73) \cdot \dfrac{14114}{\sqrt{34}} = \boxed{\$6608.05}$

Construct confidence interval

$\bar{x} \pm E$: 36000 ± 6608.05 : $\boxed{CI: (\$29391.95, \$42608.05)}$

Interpretation of CI (*note: you do need units here!*) yes?

(template answer) "*We are* 99% *confident that the true mean* cost of a wedding in US cities *is between* \$29,391.95 *and* \$42,608.05."

Graphing Calculator (TI-83 or 84)

Instructions: (a) STAT $\Rightarrow$ TESTS $\Rightarrow$ $\boxed{\text{TInterval}}$

(b) Enter $\begin{cases} \text{Data} \rightarrow \text{if individual values are entered in list} \\ \text{Stats} \rightarrow \text{if using summary statistics} \end{cases}$

Ex3: A Statistics instructor wanted to find out the mean amount of time GCC students spent on social media per day. She randomly sampled 12 students and found out how many minutes per day they used social media. Those results are below:

$n = 12$ $df = 11$

want 1VARSTATS

80	112	95	72	150	120	85	67	92	102	50	115

$\bar{x} = 95$ min/day

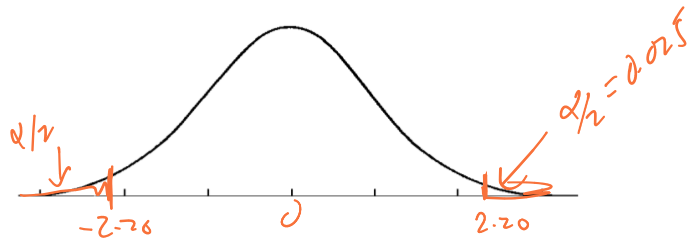
$s = 27.1$

Assuming that such times for all GCC students are normally distributed, make a **95% confidence interval** for the corresponding population mean amount of time spent on social media for all students at GCC.

Check requirements ① SRS ② indep ③ σ unknown ④ normal or ~~n > 30~~

Identify $\boxed{\textit{point estimate}}$ *(sample mean)*

$$\boxed{\bar{x} = 95 \text{ min/day}}$$



$CL = 0.95$

$\alpha = 1 - CL = 0.05$

$\alpha/2 = 0.025$

Determine critical value $t_{\alpha/2}$

$$t_{\alpha/2} = \left| \text{invT}\left(\tfrac{\alpha}{2}, df\right) \right|$$

$$t_{\alpha/2} = t_{0.025} = \left| \text{invT}(\underset{\alpha/2}{0.025}, \underset{df}{11}) \right| = 2.20$$

Find margin of error

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \qquad E = 2.20 \cdot \frac{27.1}{\sqrt{12}} = \boxed{17.2}$$

Construct confidence interval

$$CI: (\bar{x} - E, \bar{x} + E)$$

$$\bar{x} \pm E: 95 \pm 17.2$$

μ

$$\boxed{CI: \left(77.8 \tfrac{\text{min}}{\text{day}}, 112.2 \tfrac{\text{min}}{\text{day}}\right)}$$

Interpretation of CI #1

"We are 95% confident that the true population mean amount of time spent on social media by all GCC students is 77.8 and 112.2 minutes per day."

Interpretation of CI #2 "If we obtained 100 samples of size n=12 from the population of GCC students*, then we would expect about* 95 *of the samples to result in confidence intervals that* contain *μ."*

Let's say we want to be 95% confident that the true mean will be within 5 minutes. What sample size is needed?

E

Remember: $\boxed{E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}}$. Need to solve for n.

DETERMINING SAMPLE SIZE

Population Standard Deviation σ ***Known***
$\boxed{n = \left(\frac{z_{\alpha/2} \cdot s}{E}\right)^2}$

Wait? Why are we using $z_{\alpha/2}$ instead of $t_{\alpha/2}$? Because we don't know n, hence df! So it's our best guess for it.

Rounding Rule for n: always up!

Ex: The president of GCC Dr. Viar is concerned about the amount of time that students spend on jobs. He would like to estimate the mean number of hours worked per week by these students. He knows that the standard deviation of the times spent per week on such jobs by all students is 2.5 hours. What sample size should he choose so that he can be 90% confident that the estimate is within .75 hours of the population mean?

- parameter: mean amount time spent working μ → t-dist
- looking for: sample size n

Givens
- $s = 2.5$
- $CL = 0.9$
- $E = 0.75$

Want
$n = ?$

$$n = \left(\frac{z_{\alpha/2} \cdot s}{E}\right)^2 = \left(\frac{1.64 * 2.5}{0.75}\right)^2 = 29.884\ldots$$

$\boxed{n = 30 \text{ students}}$

$CL = 0.9$
$\alpha = 0.1$
$\alpha/2 = 0.05$

$z_{\alpha/2} = \text{invNorm}(0.05, 0, 1, \text{RIGHT}) = 1.64$

How to tell to use z-distribution or t-distribution?	
If: 1) You are looking for a CI 2) You are estimating proportions p THEN use z-distribution for critical values $z_{\alpha/2}$	If: 1) You are looking for a CI 2) You are estimating means μ THEN use t-distribution for critical values $t_{\alpha/2}$
???? If anyone comes up with something, let me know!	"Mr. T is Mean"