Statistics

Spring 2023

Lecture 13

Feb 19-8:47 AM

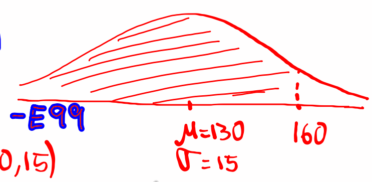
Class QZ 5

Given $\boxed{N(\overset{\mu}{130}, \overset{\sigma}{15})}$

1) Find $P(x < 160)$

$= \text{normalcdf}(-E99, 160, 130, 15)$

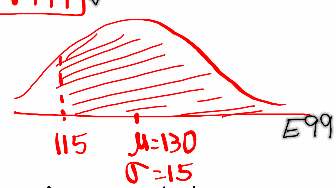
$= \boxed{.977}$ ✓



2) Find $P(x > 115)$

$= \text{normalcdf}(115, E99, 130, 15)$

$= \boxed{.841}$ ✓

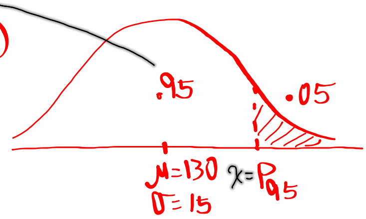


3) Find $x = P_{95}$, Round to whole #.

$x = \text{invNorm}(.95, 130, 15)$

$= 154.673$

$\approx \boxed{155}$ ✓



May 2-9:12 PM

t-Dist

Graph is bell-shape, symmetric, total Area = 1

$\mu = 0$, σ unknown

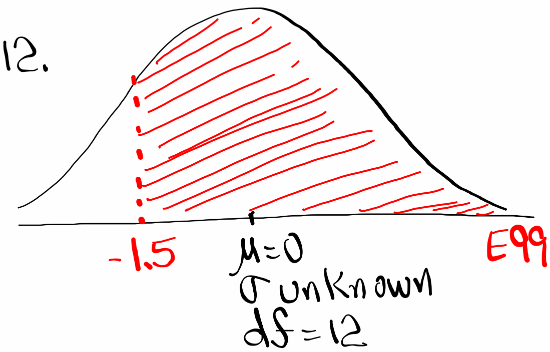
It comes with df.

$P(t > -1.5)$ with df = 12.

$= tcdf(L, U, df)$

$= tcdf(-1.5, E99, 12)$

$= \boxed{.920}$

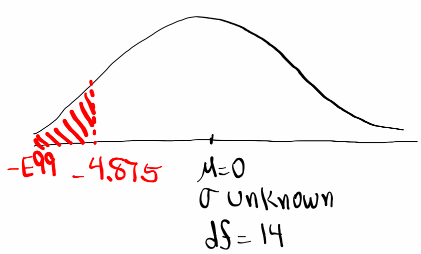


May 9-6:53 PM

Find twice the area to the left of $t = -4.875$ with df = 14.

$2 * tcdf(-E99, -4.875, 14)$

$= \boxed{2.455 \times 10^{-4}}$



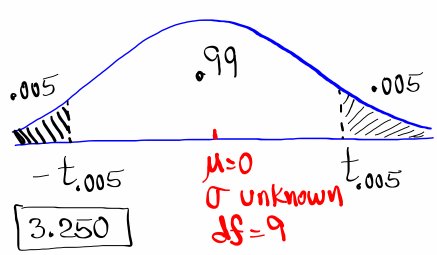
Find $t_{\alpha/2}$ for 99% C-level with df = 9.

Middle Area = .99

$1 - .99 = .01 \leftarrow \alpha$

$\alpha/2 = .005$

Left Area



$t_{.005} = invT(.995, 9) = \boxed{3.250}$

May 9-6:56 PM

24% of 275 randomly selected drivers have texted while driving.

$n=275$
$\hat{p}=.24$

1) How many of them have texted while driving?

$x = n\hat{p} = 275(.24) = \boxed{66}$

If decimal ⇒ Round up.

2) Find conf. interval for the proportion of all drivers that texts while driving.

⇒ No C-level ⇒ Use .95

1-PropZInt

$<p<$

$(.190, .290)$

$.190 < p < .290$

$\boxed{19\% < p < 29\%}$

3) Give the margin of error.

$E = \dfrac{.290 - .190}{2} = \boxed{.05}$

May 9-7:04 PM

I randomly selected 25 cars on the freeway, their mean speed was 74.8 mph.

$n=25$, $\bar{x}=74.8$

It is known that standard deviation of speeds of all cars on the freeway is 10.5 mph.

$\boxed{\sigma=10.5}$

1) Find 98% conf. interval for the mean speed of all cars on the freeway.

$<\mu<$

C-level: .98

$\boxed{\text{If } \sigma \text{ known} \Rightarrow \text{ZInterval}}$ ← inpt: $\boxed{\text{Stats}}$

If σ unknown ⇒ TInterval

$\sigma=10.5$

$\bar{x}=74.8$

$n=25$

C-level: .98

Since $\bar{x}$ is 1-decimal ⇒ Round to 1-dec.

$\boxed{69.9<\mu<79.7}$

2) Find the margin of error.

$E = \dfrac{79.7 - 69.9}{2} = \boxed{4.9}$

May 9-7:12 PM

12 randomly Selected Cars on the FWY had a mean speed of 72.5 mph with standard deviation of 9.5 mph.

$n = 12$
$\bar{x} = 72.5$
$S = 9.5$

1) Find conf. interval for the mean speed of all cars on the freeway.

→ No C-level ⇒ Use .95

$< \mu <$

If σ Known ⇒ Z Interval

$\boxed{\text{If } \sigma \text{ Unknown} \Rightarrow \text{T Interval}}$ ← inpt: $\boxed{\text{Stats}}$

2) Find its margin of error.

$\bar{x} = 72.5$
$S = 9.5$
$n = 12$
C-level: .95

$E = \dfrac{78.5 - 66.5}{2} = \boxed{6}$ $\quad \boxed{66.5 < \mu < 78.5}$

May 9-7:21 PM

Minimum Sample Size for constructing Confidence Interval:

1) Pop. Proportion

$E = Z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p}\hat{q}}{n}} \xrightarrow[\text{Algebra}]{\text{with Some}} n = \hat{p}\hat{q}\left(\dfrac{Z_{\alpha/2}}{E}\right)^2$

If decimal ⇒ Round-up

If $\hat{p}$ & $\hat{q}$ are unknown, we use .5 for each $\quad n = .25\left(\dfrac{Z_{\alpha/2}}{E}\right)^2$

Find minimum Sample Size needed if we wish to build 95% conf. interval with margin of error not to exceed 4%.

1) Assume $\hat{p} = .4$, $\hat{q} = .6$ $\quad n = \hat{p}\hat{q}\left(\dfrac{Z_{\alpha/2}}{E}\right)^2$

$Z_{\alpha/2} = \text{invNorm}(.975, 0, 1) = 1.960$

$= (.4)(.6)\left(\dfrac{1.960}{.04}\right)^2$

$= 576.24$ $\quad \boxed{n = 577}$

2) If $\hat{p}$ & $\hat{q}$ unknown

$n = .25\left(\dfrac{1.960}{.04}\right)^2$

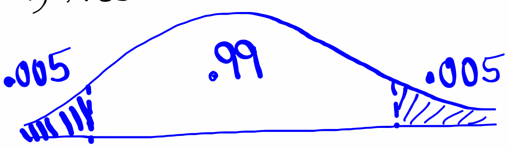
$= 600.25$ $\quad \boxed{n = 601}$

May 9-7:28 PM

Find min. Sample Size needed to construct 99% conf. interval for the pop. proportion and error not to exceed 5%.

$n = \hat{p}\hat{q}\left(\frac{Z_{\alpha/2}}{E}\right)^2$

1) Assume $\hat{p} = .7$



$Z_{\alpha/2} = \text{invNorm}(.995, 0, 1)$

$= (.7)(.3)\left(\frac{2.576}{.05}\right)^2$

$= 557.405\ldots$

$\boxed{n = 558}$

2) Assume $\hat{p}$ & $\hat{q}$ are unknown.

$n = .25\left(\frac{Z_{\alpha/2}}{E}\right)^2 = .25\left(\frac{2.576}{.05}\right)^2 = 663.5776$

$\boxed{n = 664}$

May 9-7:36 PM

2) Population Mean

$E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \xrightarrow[\text{Algebra}]{\text{with Some}} n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2$

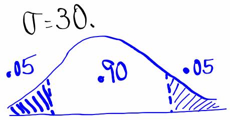
If decimal
$\Rightarrow$ Always Round-up

If σ is unknown, use S instead of σ.

$n = \left(\frac{Z_{\alpha/2} \cdot S}{E}\right)^2$

Find min. Sample Size need to Construct 90% Conf. interval for pop. mean with error not exceed 5 units and $\sigma = 30$.

$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{E}\right)^2$



$Z_{\alpha/2} = \text{invNorm}(.95, 0, 1) =$

$= \left(\frac{1.645 \cdot 30}{5}\right)^2$

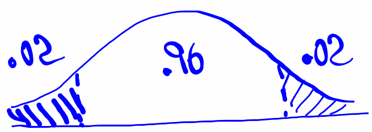
$= 97.4169$

$\boxed{n = 98}$

May 9-7:42 PM

Find min. Sample Size need to construct 96% conf. interval for the population mean with error not to exceed 10 units and $S = 25$.

$$n = \left(\frac{Z_{\alpha/2} \cdot S}{E}\right)^2$$



$Z_{\alpha/2}$ = invNorm (.98, 0, 1)

$$= \left(\frac{2.054 \cdot 25}{10}\right)^2 = 26.368\ldots \quad \boxed{n = 27}$$

SG 22 or SG 23 ✓✓

May 9-7:49 PM

SG 24–27

Testing Claims:

Claim could be about
1) Population Proportion P
2) Population Mean μ
3) Population Standard deviation σ

I claim 10% of all students are left-handed.

I claim the mean age of all college students is below 30 Yrs.

I claim the standard deviation of Scores of all exams is at least 10.5.

May 9-8:05 PM

why do we need to test claims?

It is simply to determine the validity of the claim.

If claim is valid ⇒ we support it.

If claim is invalid ⇒ we reject it.

Possible errors:

when we reject a valid claim

when we support an invalid claim.

May 9-8:09 PM

Testing Methods:

1) Traditional Method
2) P-Value Method

we only use these two methods.

3) Confidence Interval Method

Regardless of the method, final conclusion should be the same.

Final conclusion must be about the claim

Reject the claim OR Fail-to-Reject (Support) the claim

when claim is invalid — when claim is valid

Action \ claim	Valid	Invalid
Reject	Error	Not error
Support	Not error	Error

May 9-8:13 PM

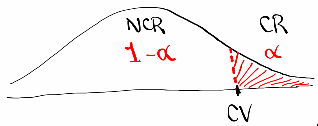
Testing Types:

1) Right-Tail Test (RTT)
2) Left-Tail Test (LTT)
3) Two-Tail Test (TTT)

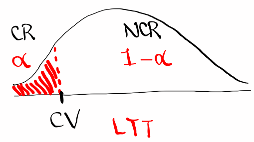
with every testing, we have a Significance level α

$0 < \alpha < 1$

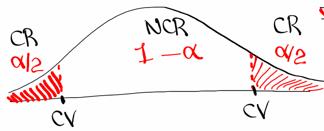
when α not given $\Rightarrow$ use .05



RTT

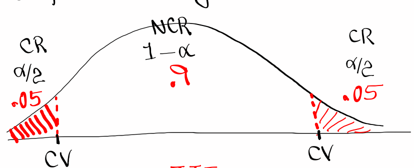


LTT



TTT

Ex: Do complete drawing with $\alpha = .1$ for TTT



TTT

May 9-8:20 PM

Testing Process:

1) Set-up H_0 & H_1

Null Hypothesis (H_0); Alternative Hypothesis (H_1)

2) Find all Critical Values

Drawing, labeling, Shading, and TI Command.

3) Find Computed Test Statistic (CTS) and P-Value (P).

TI Command or formula required.

4) Use Testing Chart to determine the Validity of H_0 & H_1.

H_0 Valid $\Leftrightarrow$ H_1 invalid

H_0 invalid $\Leftrightarrow$ H_1 valid

5) Draw Final Conclusion about the claim

Reject the claim when claim is invalid

OR

FTR the claim (Fail-to-Reject) when claim is Valid.

May 9-8:28 PM

More on H_0 & H_1:

H_0 must contain = Sign ⇒ $=, \geq, \leq$

H_1 cannot contain = Sign ⇒ $\neq, <, >$

Keywords:

H_0: is, equal, Same, at least, at most, not different, ...

H_1: isnot, not equal, different, more than, less than, above, below, exceed, ...

Claim could be H_0 or H_1, but not both at the Same time.

Always identify the claim and type of testing:

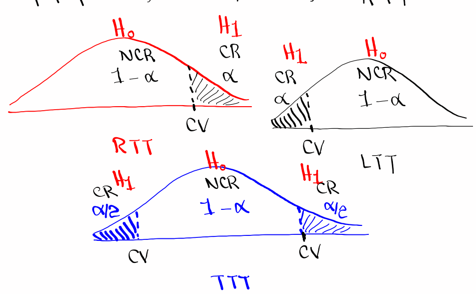
H_1: $\neq$ ⇒ Two-Tail Test

H_1: $>$ ⇒ Right-Tail Test

H_1: $<$ ⇒ Left-Tail Test

May 9-8:36 PM

H_0: $=$	H_0: $\geq$	H_0: $\leq$
H_1: $\neq$	H_1: $<$	H_1: $>$
TTT	LTT	RTT



$$\boxed{\begin{aligned} P(H_0 \text{ valid}) &= 1-\alpha = P(H_1 \text{ invalid}) \\ P(H_0 \text{ invalid}) &= \alpha = P(H_1 \text{ valid}) \end{aligned}}$$

Type I & Type II errors

Four-Possible outcomes for H_0:

Conclusion \ Reality	H_0 valid	H_0 invalid
Support H_0	Correct Decision	Type II error
Reject H_0	Type I error	Correct Decision

May 9-8:43 PM

I claim 10% of all students are left-handed.

$H_0: p = .10$ claim

$H_1: p \neq .10$ TTT

I claim the mean age of all college students is below 30 yrs.

$\mu < 30$

$H_0: \mu \geq 30$

$H_1: \mu < 30$ claim, LTT

I claim the standard deviation of all exams is at most 10.

$\sigma \leq 10$

$H_0: \sigma \leq 10$ claim

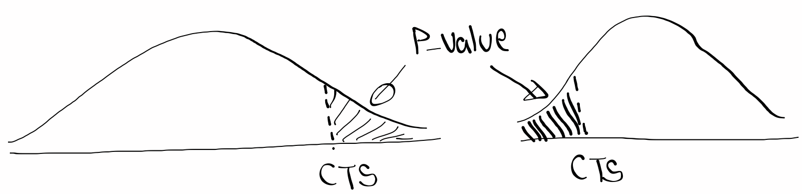
$H_1: \sigma > 10$ RTT

May 9-8:54 PM

What is P-Value?

P-value is the area of the tail of the graph of prob. dist. marked by CTS.

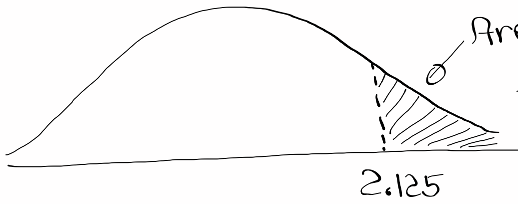
Multiply this area by 2 only when doing TTT.



only for TTT, multiply by 2.

May 9-8:59 PM

CTS $Z=2.125$, RTT, find P-Value

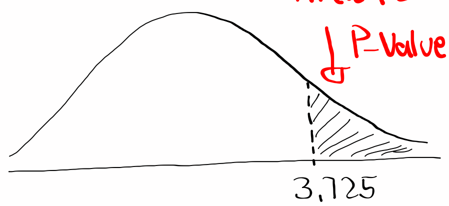


Area = P-Value
$= \text{normalcdf}(2.125, E99, 0, 1)$
$= \boxed{.017}$

CTS $t=3.725$, TTT, $df=19$

Find P-Value

Area * 2 =
P-Value

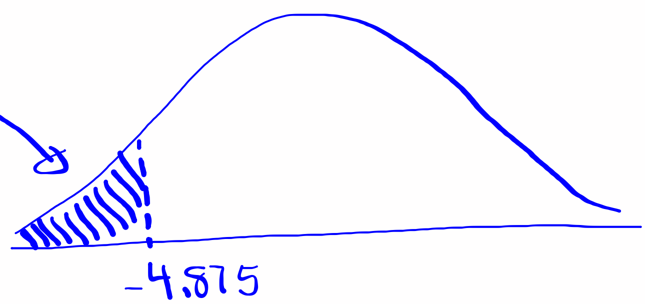


P-Value $= 2 * \text{tcdf}(3.725, E99, 19) = \boxed{.001}$

May 9-9:02 PM

CTS $t=-4.875$, LTT, $df=8$.

Find P-Value.



P-value $= \text{tcdf}(-E99, -4.875, 8) = \boxed{6.161 \times 10^{-4}}$

May 9-9:06 PM

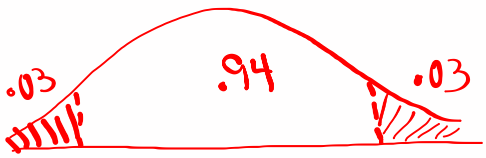
Class QZ 6

Drawing, labeling, shading, TI command required.

1) Find $z_{\alpha/2}$ for 94% C-level.

$z_{\alpha/2}$ = invNorm(.97, 0, 1) = $\boxed{1.881}$

2) Find $t_{\alpha/2}$ for $\alpha = .06$ with df = 19.



$t_{\alpha/2}$ = invT(.97, 19) = $\boxed{2.000}$

May 9-9:09 PM