Statistics

Lecture 14

Feb 19-8:47 AM

How to determine minimum Sample Size when constructing Conf. interval:

1) Population Proportion

$$E = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

with some algebra work, we can solve for n

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

if decimal, always Round-up

if $\hat{p}$ & $\hat{q}$ are both unknown, we use $\hat{p} = .5$ & $\hat{q} = .5$

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

Always round-up when decimal.

Feb 4-4:34 PM

Among 125 randomly selected students, 20% of them were smokers. $\hat{q} = 1 - \hat{p} = 1 - .2 = .8$

$n = 125$, $\hat{p} = .2$ → $x = n\hat{p} = 125(.2) = 25$

if decimal → Round-up

Find 90% Conf. interval for the prop. of all students that are smokers.

C-level : .9

1-PropZInt

x : 25

n : 125

C-level : .9

Calculate

$.14 < p < .26$

we are 90% confident that between 14% & 26% of all students are smokers.

$E = \frac{.26 - .14}{2} = .06$

$\hat{p} = \frac{.26 + .14}{2} = .2$

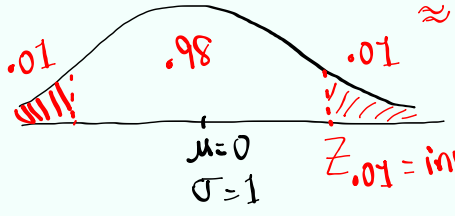
Feb 4-4:39 PM

what if we wish error to be no more than .04 and Conf. level to be 98%, what is the minimum Sample Size?

$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 = (.2)(.8)\left(\frac{2.326}{.04}\right)^2$

≈ 541.0276

≈ 542



$z_{.01} = \text{invNorm}(.99, 0, 1)$

Suppose $\hat{p}$ & $\hat{q}$ were unknown

$n = .25\left(\frac{2.326}{.04}\right)^2 \approx 845.356 \approx 846$

Feb 4-4:45 PM

In a sample of 525 students, 38% of them were in favor of True/false questions.

$n = 525$ $\quad \hat{p} = .38$ $\quad \hat{q} = 1 - \hat{p} = .62$

$x = n\hat{p} = 525(.38) = 199.5 \approx \boxed{200}$

Find **Conf. interval** for the **prop.** of **all** students that are in favor of True/false questions.

↳ no C-level → use .95

1-PropZInt
x: 200
n: 525
C-level: .95

$E = \frac{.42 - .34}{2} = \boxed{.04}$

$\hat{p} = \frac{.42 + .34}{2} = \boxed{.38}$

$.34 < p < .42$

we are 95% confident that between 34% & 42% of all students are in favor of True/false questions.

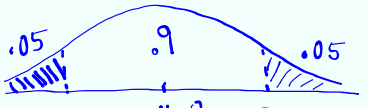
Feb 4-4:53 PM

Find min. sample size needed to construct conf. interval for prop. of all students with error not to exceed 5% and conf. level of 90%.

$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 = (.38)(.62)\left(\frac{1.645}{.05}\right)^2$

≈ 255.016

$\approx \boxed{256}$



$z_{.05} = \text{invNorm}(.95, 0, 1)$

what if $\hat{p}$ & $\hat{q}$ were both unknown,

$n = .25\left(\frac{z_{\alpha/2}}{E}\right)^2 = .25\left(\frac{1.645}{.05}\right)^2$

≈ 277.6025

$\approx \boxed{278}$

Feb 4-5:01 PM

How to determine minimum Sample Size when constructing Conf. interval:

2) Population mean

$$E = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

with some algebra work, Solve for n

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

if decimal, Round-up

If σ is unknown, use S in its place.

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

if decimal, Round-up

Feb 4-4:34 PM

In a Sample of 40 doctors, their mean age was 52.5 yrs.

$n = 40$ $\bar{x} = 52.5$ $\sigma = 7.8$

It is known that standard deviation of ages of all doctors is 7.8 yrs.

Find 99% Conf. interval for the mean age of all doctors.

C-level: .99

$49.3 < \mu < 55.7$

σ Known → Z Interval

inpt: Stats

σ: 7.8

$\bar{x}$ = 52.5 ← 1-dec.

n = 40

C-level: .99

$E = \frac{55.7 - 49.3}{2} = 3.2$

$\bar{x} = \frac{55.7 + 49.3}{2} = 52.5$

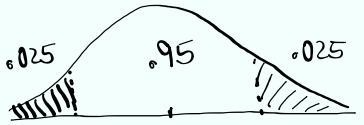
Feb 4-5:14 PM

Find min. Sample Size needed if we wish the error to be within 5 yrs and C-level to be 95%.

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{1.960 \cdot 7.8}{5}\right)^2$$

≈ 9.349

$\approx \boxed{10}$



$\mu = 0$
$\sigma = 1$

$z_{.025} = \text{invNorm}(.975, 0, 1)$

Redo with $E = 2.5$

$$n = \left(\frac{1.960 \cdot 7.8}{2.5}\right)^2 \approx 37.396$$

$\approx \boxed{38}$

Feb 4-5:23 PM

Ages of 10 randomly selected nurses are given below:

38	42	50	28	30
40	55	45	35	48

Store in a list, use 1-Var Stats to find

$\bar{x} = 41.1$
$S = 8.7$
$n = 10$

} Round to 1 dec.

Find 90% Conf. interval for the mean age of all nurses.

σ unknown → TInterval

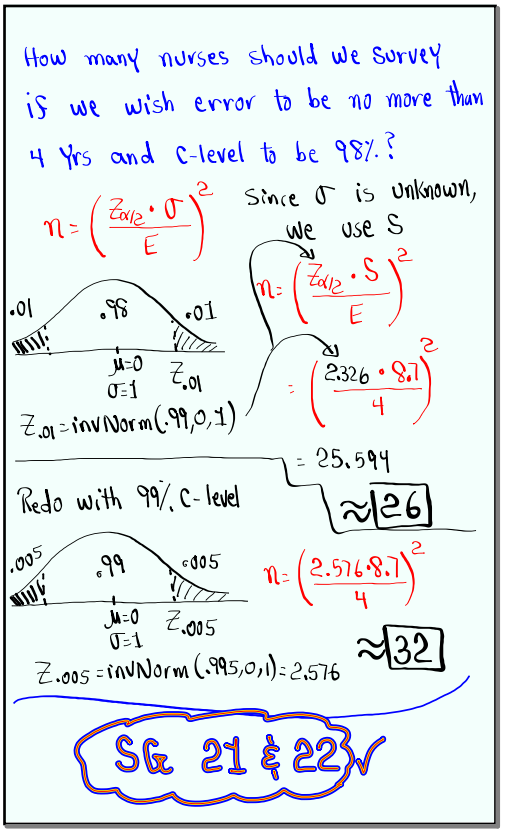
inpt: Stats
$\bar{x} = 41.1$ ← 1 dec.
$S = 8.7$
$n = 10$
C-level: .9

$E = \frac{46.1 - 36.1}{2} = 5$

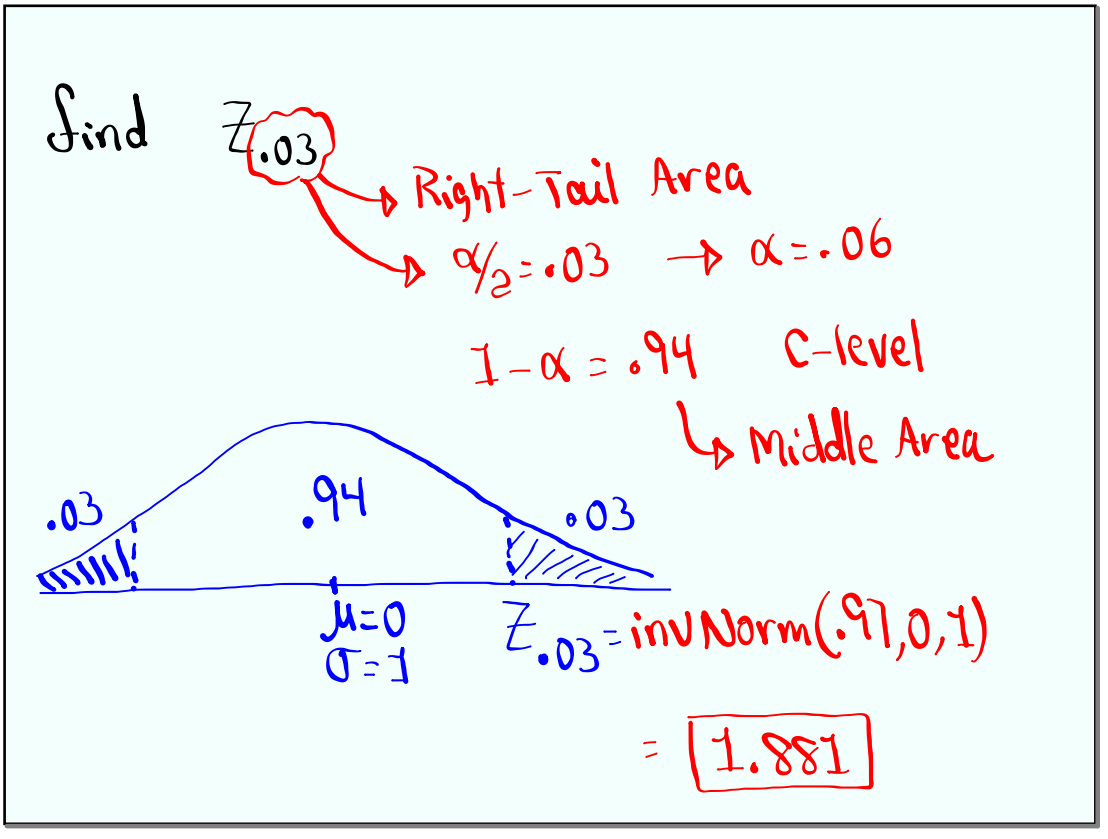
$\bar{x} = \frac{46.1 + 36.1}{2} = 41.1$

$\boxed{36.1 < \mu < 46.1}$

Feb 4-5:30 PM



Feb 4-5:38 PM



Feb 4-6:01 PM

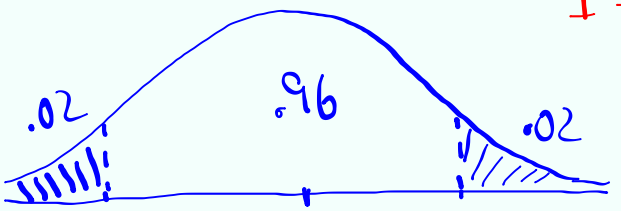
Find $\pm z_{\alpha/2}$ for 96% C-level

Middle area .96

$1-\alpha = .96$

$\alpha = .04$

$\alpha/2 = .02$



$= \text{invNorm}(.98, 0, 1)$

$= 2.054$

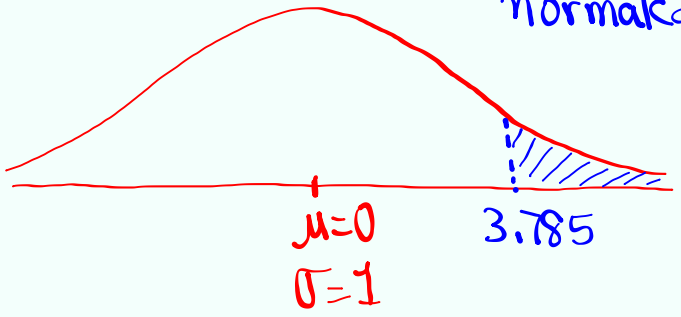
-2.054

Feb 4-6:04 PM

Find the area to the right of $z = 3.785$.

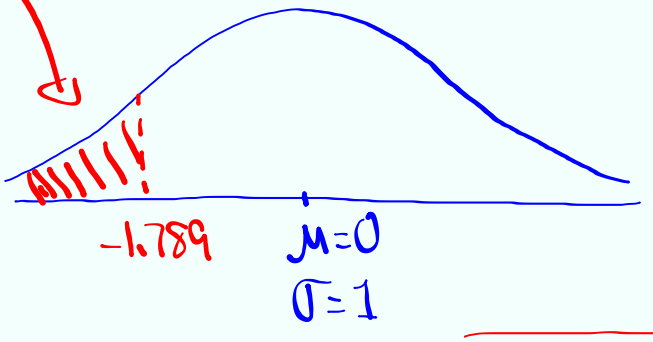
normalcdf(3.785, E99, 0, 1)

$= 7.688 \times 10^{-5}$



Feb 4-6:07 PM

Find twice the area to the left of $Z = -1.789$.



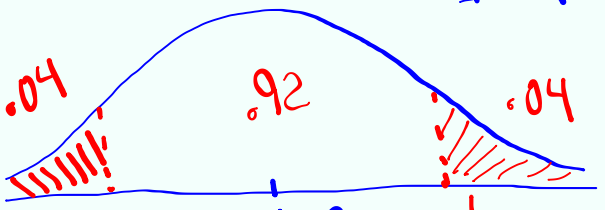
2 • normalcdf(-E99, -1.789, 0, 1) = .074

Feb 4-6:10 PM

Find $t_{\alpha/2}$ for $\alpha = .08$ and $df = 15$.

$\alpha/2 = .04$ Area of each tail

$1 - \alpha = .92$ C-level Middle area



$t_{.04} = \text{invT}(.96, 15)$

≈ 1.878

Feb 4-6:14 PM

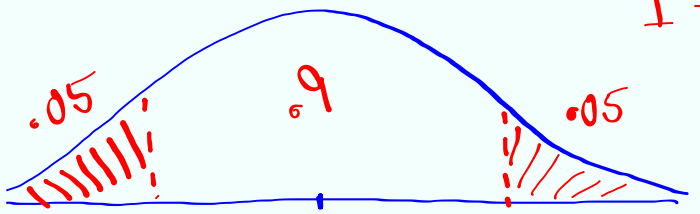
Find $\pm t_{\alpha/2}$ for 90% C-level & df = 24.

Middle Area = .9

$1 - \alpha = .9$

$\alpha = .1$

$\alpha/2 = .05$



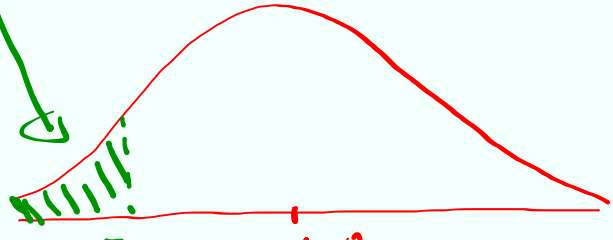
$-t_{.05}$ $\mu = 0$ $t_{.05}$ = invT(.95, 24)

$\boxed{-1.711}$ σ unknown df = 24 $= \boxed{1.711}$

Feb 4-6:17 PM

Find the area to the left of $t = -5.678$ with df = 12.

use tcdf(L, U, df)



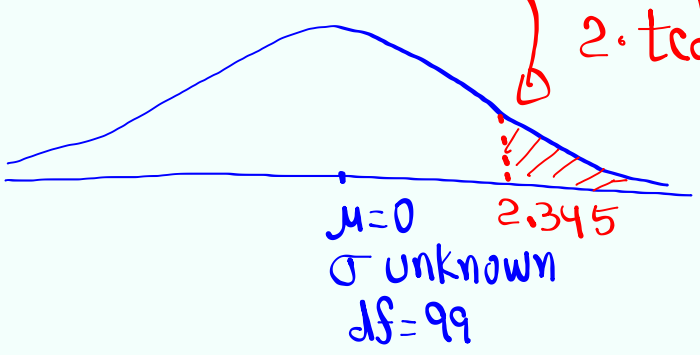
σ unknown df = 12

tcdf(-E99, -5.678, 12)

$= \boxed{5.132 \times 10^{-5}}$

Feb 4-6:20 PM

Find twice the area to the right of $t = 2.345$ with $df = 99$.



$2 \cdot tcdf(2.345, E99, 99) = .021$

Feb 4-6:24 PM

F - Dist:

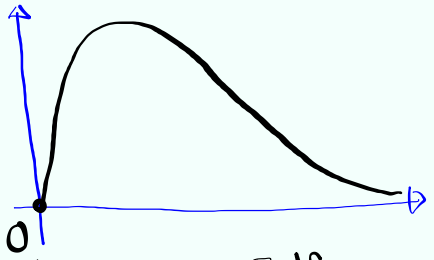
1) Graph begins at 0 and is Positively Skewed.

2) Not symmetric but total area is 1.

3) It comes with two degrees of freedom

Numerator df → Ndf

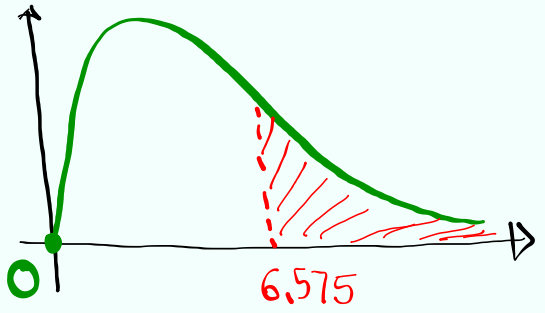
Denominator df → Ddf



to find prob. → we use Fcdf

Feb 4-6:27 PM

Find the area to the right of $F = 6.575$ with $Ndf = 4$ & $Ddf = 25$.

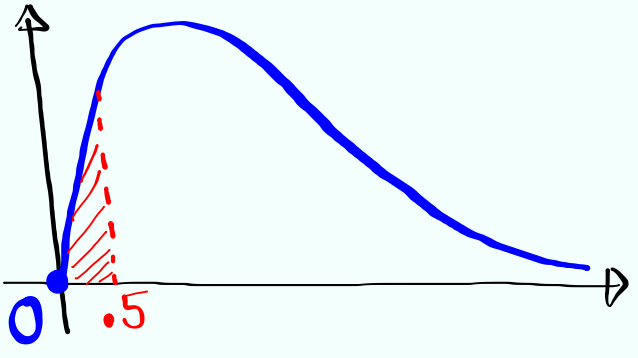


$fcdf(L, U, Ndf, Ddf)$

$fcdf(6.575, E99, 4, 25) = \boxed{9.279 \times 10^{-4}}$

Feb 4-6:31 PM

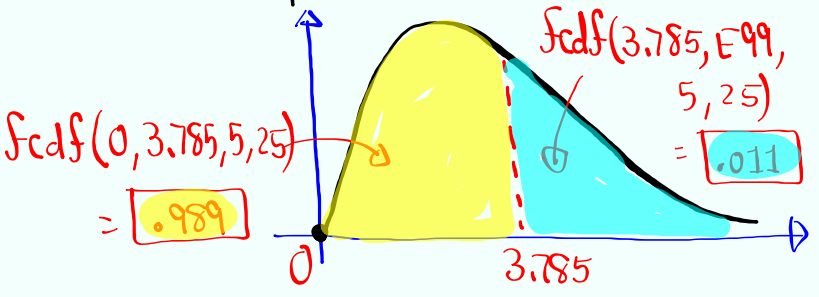
Find the area to the left of $F = .5$ with $Ndf = 3$ & $Ddf = 30$.



$fcdf(0, .5, 3, 30) = \boxed{.315}$

Feb 4-6:34 PM

Find the area on each side of $F = 3.785$ with $Ndf = 5$ & $Ddf = 25$, then multiply the Smaller area 2.



$2 \cdot$ Smaller area $= 2(.011)$

$= .022$

Feb 4-6:37 PM

Introduction to testing: (Sec 22)

We test claims to determine their validity.

A claim is made.

We test the claim to see if it is valid or invalid.

If claim is valid → we Support it.
we fail-to-Reject it.

If claim is invalid → we reject it.

Claim valid ⟺ Fail-to-Reject

Claim invalid ⟺ Reject

Feb 4-6:43 PM

Possible errors:

Claim is valid but we reject it.

Claim is invalid but we fail-to-Reject it.

Testing Methods:

1) Traditional

2) P-Value

3) Confidence Interval

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

Reject the claim (claim is invalid)

Fail-to-Reject the claim (claim is valid)

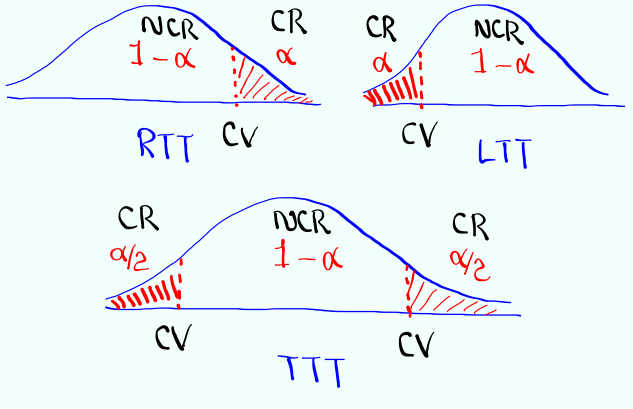
Feb 4-6:48 PM

Testing types:

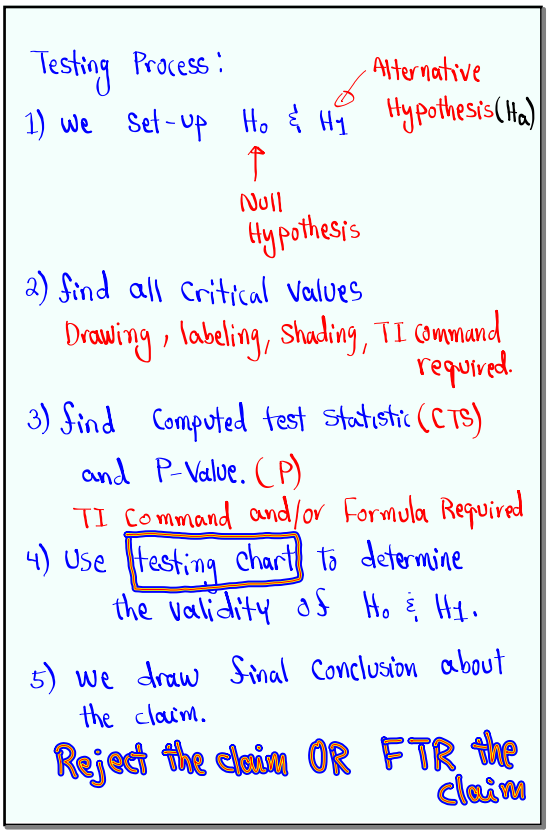
1) Right-Tail Test (RTT)

2) Left-Tail Test (LTT)

3) Two-Tail Test (TTT)



Feb 4-6:54 PM


Testing Process:
1) We set-up H_0 & H_1
Alternative Hypothesis (H_a)
Null Hypothesis
2) find all Critical Values
Drawing, labeling, Shading, TI command required.
3) find Computed test Statistic (CTS)
and P-Value. (P)
TI Command and/or Formula Required
4) Use Testing Chart to determine the validity of H_0 & H_1.
5) We draw Final Conclusion about the claim.
Reject the claim OR FTR the claim

Feb 4-7:00 PM