Statistics

Lecture 11

Feb 19-8:47 AM

SG21

find $\pm z_{\alpha/2}$ for 96% Conf. level.

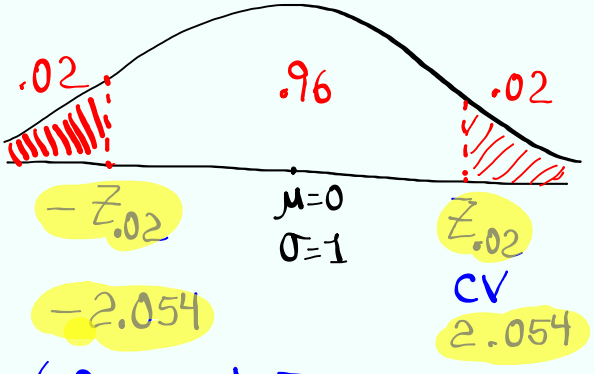
Middle Area is .96

$z_{\alpha/2}$ ↑ Critical Value

$1 - .96 = .04$

$.04 \div 2 = .02$

$\frac{\alpha}{2} = .02 \rightarrow \alpha = .04$



$z_{.02} = \text{invNorm}(.98, 0, 1) = 2.054$

May 9-8:03 AM

Estimating Population Proportion p

$\hat{p} - E < p < \hat{p} + E$

$\hat{p}$: Sample Proportion

$\hat{p} = \frac{x}{n}$

Point-estimate

E: Margin of error

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

C.V. for $(1-\alpha)\cdot 100\%$ C-level

$\hat{q} = 1 - \hat{p}$

n Sample Size

May 9-8:08 AM

Given $n = 280$, $x = 70$,

Find Conf. interval for Population Prop.

$\hat{p} = \frac{x}{n} = \frac{70}{280} = \frac{1}{4} = .25$

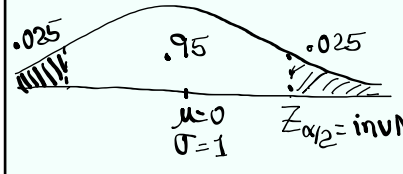
$\hat{q} = 1 - \hat{p} = 1 - .25 = .75$

No C-level → use 95%

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

$= 1.960 \cdot \sqrt{\frac{(.25)(.75)}{280}}$

$\approx .05$



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

1-PropZInt
x=70
n=280
C-level: .95

.199 < p < .301

.20 < p < .30

$\hat{p} - E < p < \hat{p} + E$

$.25 - .05 < p < .25 + .05$

$.2 < p < .3$

May 9-8:11 AM

In a Survey of 275 students, 73% of them were driving alone to campus.

$n=275$ $\quad \hat{p}=\frac{x}{n}$ $\quad x=n\hat{p}=275(.73)=200.75$ → $x=201$

$\hat{p}=.73$ $\quad$ if decimal → Round-up

Find 99% Conf. interval for the prop. of all students that drive to campus alone.

C-level: .99

1-PropZInt

$x=201$

$n=275$

C-level: .99

$E=\frac{.80-.66}{2}=.07$

$\hat{p}=\frac{.80+.66}{2}=.73$

$.66<p<.80$

May 9-8:19 AM

How to determine minimum Sample Size needed to construct Conf. interval for pop. prop.:

$E=z_{\alpha/2}\cdot\sqrt{\frac{\hat{p}\hat{q}}{n}}$ ⟹ $n=\hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$

If decimal ⇒ Round-up

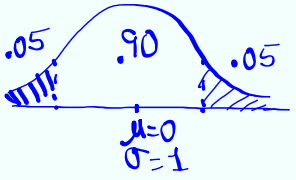
If $\hat{p}$ & $\hat{q}$ are both unknown ⇒ use .5 for each

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

May 9-8:28 AM

Suppose we wish to construct 90% Conf. interval with margin of error not to exceed 5%. (from last example)

$\hat{P}=.73$, $\hat{q}=.27$



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

$z_{\alpha/2}=\text{invNorm}(.95,0,1)=1.645$

$=(.73)(.27)\left(\frac{1.645}{.05}\right)^2=213.34$

$n=214$

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

$n=.25\left(\frac{1.645}{.05}\right)^2=270.6025$

$n=271$

May 9-8:32 AM

I Surveyed 320 students and 78% of them were using TikTok.

$n=320$

$\hat{p}=.78$

$x=n\hat{p}$

$=320(.78)$

$=249.6$

$x=250$

find 98% Conf. interval for the prop. of all students using TikTok.

C-level: .98

1-PropZInt

$.73<p<.84$

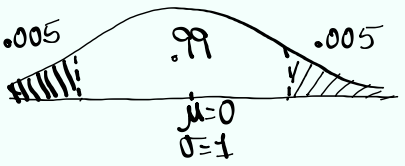
$E=\frac{.84-.73}{2}=.055$

$\hat{p}=\frac{.84+.73}{2}=.785$

May 9-8:41 AM

Find min. Sample Size needed to have 99% C.I. with error not to exceed 4%.

$$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2 = (.78)(.22)\left(\frac{2.576}{.04}\right)^2$$



$= 711.68\ldots$

$\boxed{n = 712}$

$z_{.005} = \text{invNorm}(.995, 0, 1) = \boxed{2.576}$

If we did not know $\hat{p}$ & $\hat{q}$,

$$n = .25\left(\frac{2.576}{.04}\right)^2 = 1036.84$$

$$\boxed{n = 1037}$$

May 9-8:49 AM

Estimating Population Mean μ

$$\bar{x} - E < \mu < \bar{x} + E$$

$\bar{x}$: Sample Mean, Point-estimate

E: Margin of error

Case I: σ Known

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

STAT | TESTS | ZInterval

inpt: Stats

$$E = \frac{-}{2}, \quad \bar{x} = \frac{+}{2}$$

May 9-9:11 AM

Given $n=25$, $\bar{x}=80$, $\sigma=10$ C-level: .9

Find Conf. interval for population mean.

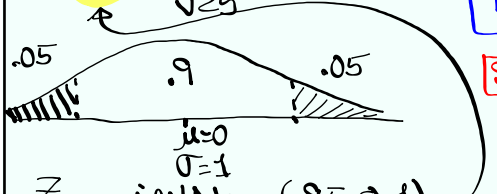
$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$

$= 1.645 \cdot \frac{10}{\sqrt{25}} = 3.29 \approx 3$

$\bar{x} - E < \mu < \bar{x} + E$

$80 - 3 < \mu < 80 + 3$

$77 < \mu < 83$



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

$77 < \mu < 83$

$E = \frac{83-77}{2} = 3$, $\bar{x} = \frac{83+77}{2} = 80$

STAT
TESTS
Z Interval
inpt: Stats
$\sigma = 10$
$\bar{x} = 80$
$n = 25$
C-level: .9

May 9-9:16 AM

I randomly selected 40 students, their mean age was 32.5 yrs. $n=40$, $\bar{x}=32.5$

$\sigma = 12.8$

It is known that standard deviation of ages of all students is 12.8 yrs.

No C-level ⇒ use .95

Find Conf. interval for the mean age of all students.

σ known
⇒ Z Interval
inpt: Stats
σ: 12.8
$\bar{x}$=32.5
n=40
C-level: .95

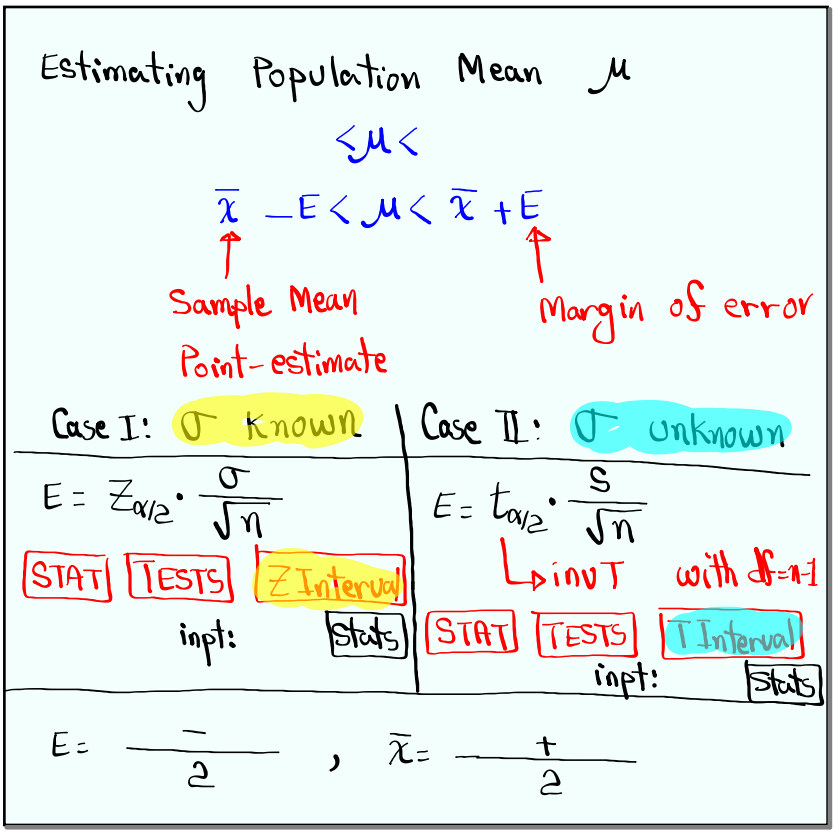
1-dec.

$28.5 < \mu < 36.5$

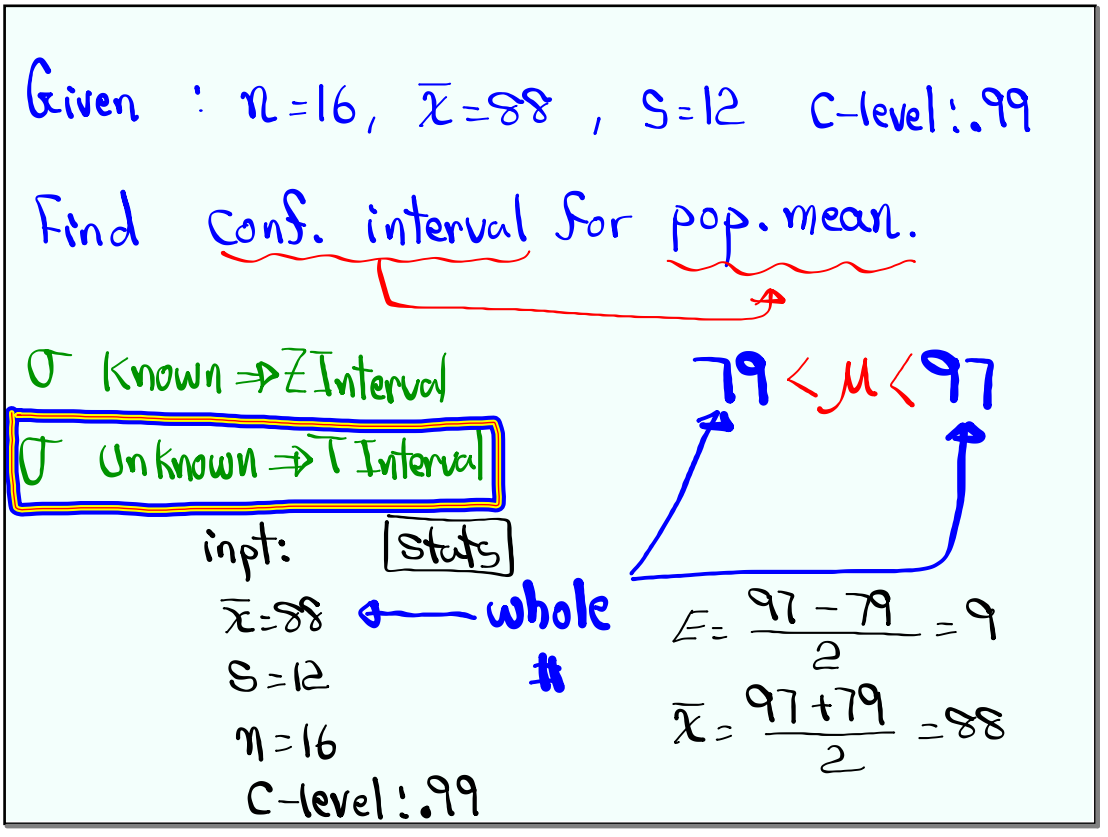
$E = \frac{36.5 - 28.5}{2} = 4$

$\bar{x} = \frac{36.5 + 28.5}{2} = 32.5$

May 9-9:24 AM



May 9-9:11 AM



May 9-9:37 AM

I randomly selected 12 students, their mean age was 31.8 yrs with standard dev. of 8.5 yrs. $n=12,\ \bar{x}=31.8,\ S=8.5$

Find 98% Conf. interval for the mean age of all students. C-level: .98

σ Known $\Rightarrow$ ZInterval

σ unknown $\Rightarrow$ TInterval

inpt: Stats

$\bar{x}=31.8$ ← 1-dec.

$S=8.5$

$n=12$

C-level: .98

$25.1 < \mu < 38.5$

$E = \frac{38.5-25.1}{2} = \square$

$\bar{x} = \frac{38.5+25.1}{2} = \square$

May 9-9:42 AM

I randomly selected 10 exams and here are the scores.

83	77	100	65
90	95	55	80
70	98		

No C-level → .95

Find

1) $\bar{x} \approx 81$ } Round to whole

2) $S \approx 15$ }

3) $S^2 = \frac{6667}{30}$ } Reduced fraction

4) Find Conf. interval for the mean of all exams.

$70 < \mu < 92$

σ Known $\Rightarrow$ ZInterval

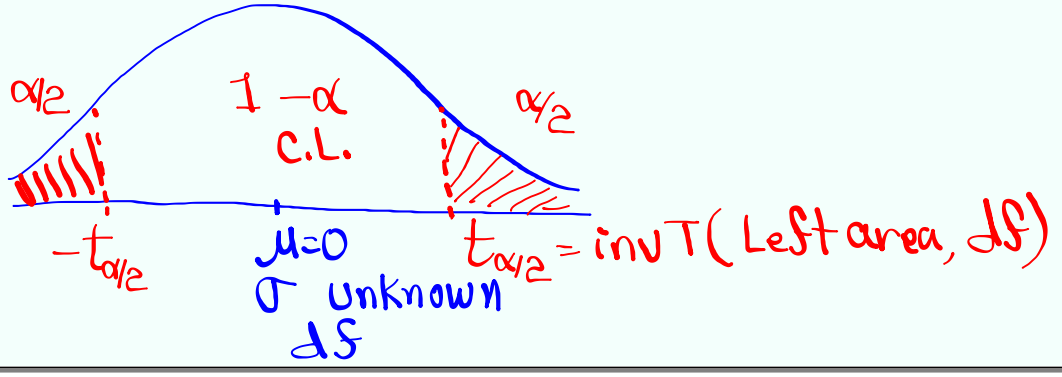
σ unknown $\Rightarrow$ TInterval

May 9-9:50 AM

what is t-Dist.?

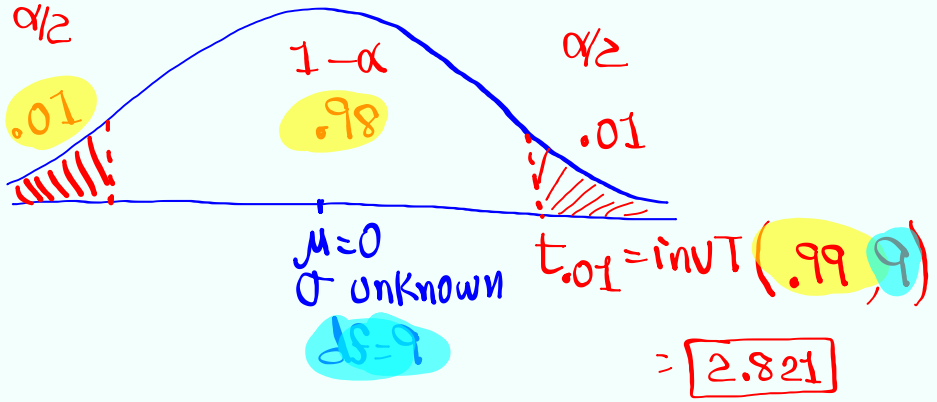
It is similar to Z-Dist

except σ is unknown but it comes with degrees of freedom.



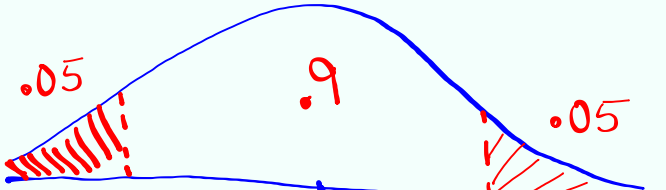
May 9-9:59 AM

Find $t_{\alpha/2}$ for $\alpha = .02$ and $df = 9$.



May 9-10:02 AM

Find $t_{\alpha/2}$ for 90% C-level & df = 49.



$\mu = 0$
σ unknown
df = 49

$t_{.05} = \text{invT}(.95, 49)$
$= \boxed{1.677}$

As df gets bigger
$t_{\alpha/2} \approx Z_{\alpha/2}$

May 9-10:06 AM

What is degrees of freedom?
You find df by topic.

17 students
I bring 17 donuts.

Max → 17 choices
George → 16 "
Alexis → 15 "
⋮
Kaytie → NO choice 1 Donut left

$df = 17 - 1 = \boxed{16}$

Daniel has 7 clean shirts.
He wears one clean shirt per day.

Monday → 7 choices
Tuesday → 6 "
Wednesday → 5 "
⋮
Sunday → 0 choices only 1 clean shirt

$df = 7 - 1 = \boxed{6}$

May 9-10:10 AM

How to determine minimum Sample Size needed when working with pop. mean:

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

if decimal ⇒ Round-up

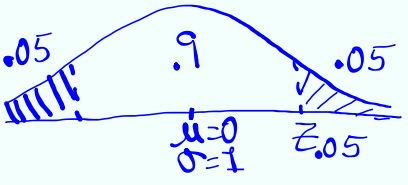
If σ is unknown, use s in its place.

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

May 9-10:30 AM

Find min. Sample Size needed to construct 90% conf. interval for pop. mean if with error not to exceed 5 pts and σ is 18.

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



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

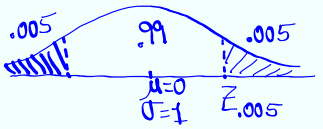
$$= \left(\frac{1.645 \cdot 18}{5} \right)^2$$

$= 35.07\ldots$

Round-up → $n = 36$

May 9-10:33 AM

Find min. sample size needed for 99% Conf. interval for pop. mean with error not to exceed 2.5 pts and $s = 20$.



$z_{.005} = \text{invNorm}(.995, 0, 1) =$

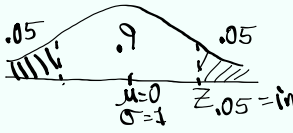
$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2 = \left(\frac{2.576 \cdot 20}{2.5}\right)^2 \approx 425$$

Redo with $E = 5$

$$n = \left(\frac{2.576 \cdot 20}{5}\right)^2 \approx 107$$

Redo with 90% clevel and $E = 5$

$$n = \left(\frac{1.645 \cdot 20}{5}\right)^2 \approx \boxed{44}$$



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

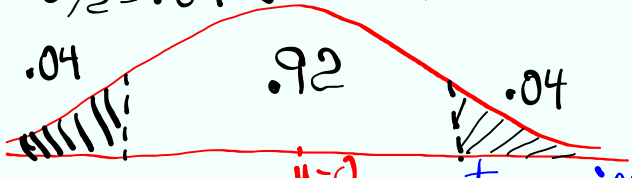
May 9-10:38 AM

Given $n = 15$, C-level: .92, σ unknown

Find $t_{\alpha/2}$ for constructing Con. level for Pop. mean.

$1 - .92 = .08 \leftarrow \alpha$

$.08/2 = .04 \leftarrow \alpha/2$



$t_{.04} = \text{invT}(.96, 14) = \boxed{1.887}$

May 9-10:47 AM