Statistics

Lecture 11

Feb 19-8:47 AM

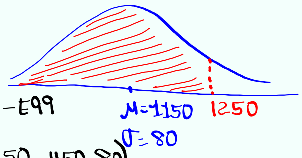
SG 19&20

SAT Scores are normally dist. with mean of 1150 and standard deviation of 80.

$N(1150, 80)$

If we randomly Select one SAT exam, find the prob. that $\overset{x}{\text{Score}}$ is $\overset{<1250}{\text{below 1250}}$.

$P(x<1250)$



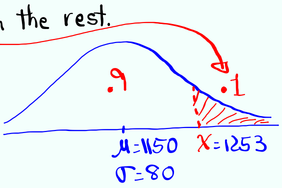
$= \text{normalcdf}(-E99, 1250, 1150, 80)$

$= \boxed{.894} \approx 89.4\%$

find $\overset{x}{\text{the Score}}$, round to whole number, that separates the $\underset{\text{Right Area}}{\text{top 10\%}}$ from the rest.

$x = \text{invNorm}(.9, 1150, 80)$

$\approx 1252.524 \approx \boxed{1253}$



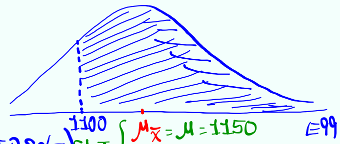
Jan 29-4:34 PM

If we randomly select 5 SAT exams, find
$n=5$

the prob. that their $\bar{x}$ mean score is above
>1100

1100.

$P(\bar{x} > 1100)$

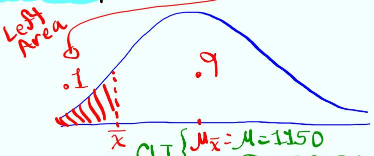


$= \text{normalcdf}(1100, E99, 1150, 80/\sqrt{5})$

CLT $\begin{cases} \mu_{\bar{x}} = \mu = 1150 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{80}{\sqrt{5}} \end{cases}$

$= \boxed{.919} \approx 91.9\%$ whole % 92%

Find the mean that separates the bottom 10% from the rest for randomly selected group of 4 SAT exams. Round up to a whole



CLT $\begin{cases} \mu_{\bar{x}} = \mu = 1150 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{80}{\sqrt{4}} = \frac{80}{2} = 40 \end{cases}$

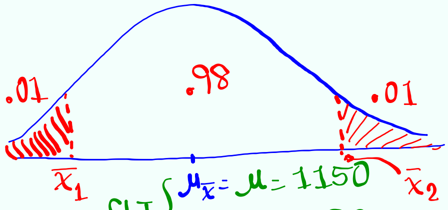
$\bar{x} = \text{invNorm}(.1, 1150, 40)$

$= 1098.738 \approx \boxed{1099}$

Jan 29-4:43 PM

Find two mean scores, round to whole numbers, that separate the middle 98% from the rest for randomly selected group of 6 SAT exams.
$n=6$

$\frac{1-.98}{2} = .01$



CLT $\begin{cases} \mu_{\bar{x}} = \mu = 1150 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{80}{\sqrt{6}} \end{cases}$

$\bar{x}_1 = \text{invNorm}(.01, 1150, 80/\sqrt{6})$

$\approx \boxed{1074}$

$\bar{x}_2 = \text{invNorm}(.99, 1150, 80/\sqrt{6}) \approx \boxed{1226}$

SG 19 & 20

Jan 29-4:54 PM

$Z_{\alpha/2}$ → Right area for our bell-shape with $\mu=0$ & $\sigma=1$

Alpha

$0<\alpha<1$

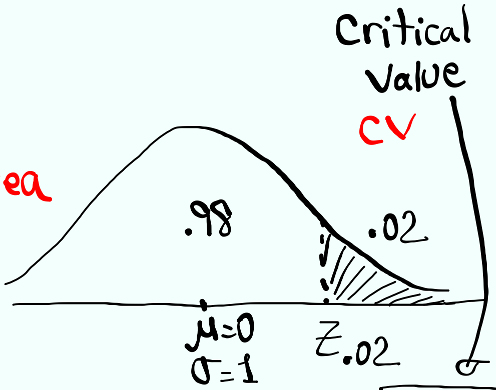
Significance level

Round answers to 3-decimal place.

Find $Z_{.02}$ → Right Area

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

$\alpha=.04$



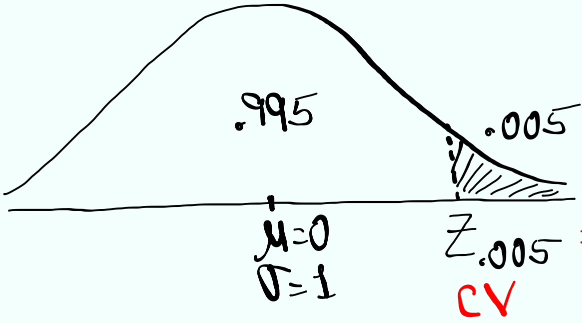
$Z_{.02}$ = invNorm(.98,0,1) = $\boxed{2.054}$

Jan 29-5:04 PM

Find $Z_{\alpha/2}$ for $\alpha=.01$.

$\alpha/2 = \frac{.01}{2} = .005$

Find $Z_{.005}$ ← Right Area



= invNorm(.995, 0, 1)

$\approx \boxed{2.576}$

Jan 29-5:09 PM

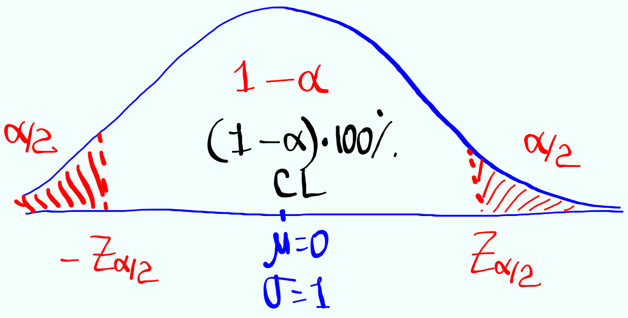
α Alpha

Significance level

$0<\alpha<1$

$(1-\alpha)\cdot 100\%$ Confidence level **(CL)**

$1-\alpha$ is the area in the middle of the graph of prob. dist. curve.



Jan 29-5:12 PM

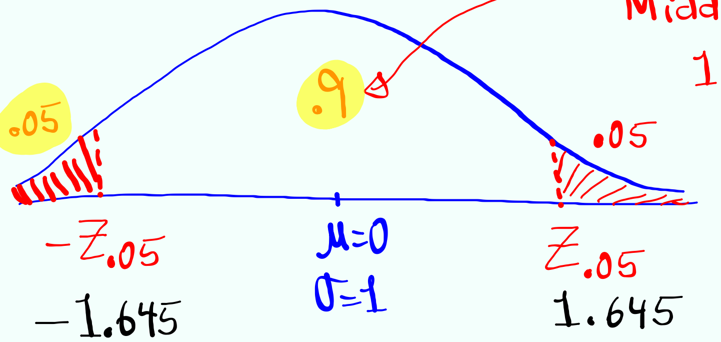
Find $\pm Z_{\alpha/2}$ for $\underbrace{90\%\text{ Confidence level.}}$

Middle Area .9

$1-\alpha=.9$

$\alpha=.1$

$\alpha/2=.05$



$Z_{.05}=\text{invNorm}(.95,0,1)=\boxed{1.645}$

Jan 29-5:16 PM

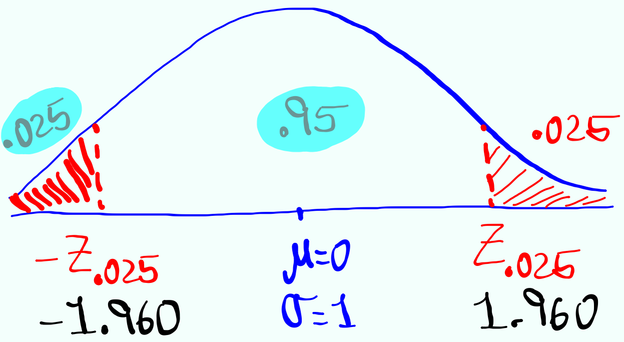
Find ± $Z_{\alpha/2}$ for 95% Confidence level.

Middle Area .95

$1-\alpha=.95$

$\alpha=.05$

$\alpha/2=.025$



$Z_{.025}=\text{invNorm}(.975, 0, 1)\approx\boxed{1.960}$

Jan 29-5:20 PM

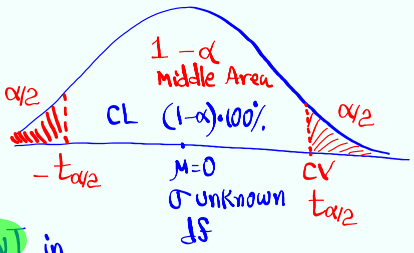
---

t - Dist.

1) Graph is symmetric, bell-shape, and has total area = 1.

2) $\mu=0$, σ unknown

3) It comes with degrees of freedom (df)



we use tcdf and invT in this topic

2nd VARS ↓ ↓

If You don't have them, download G.Calculator for Your Smart device.

Jan 29-5:25 PM

find $t_{.01}$ with df = 9.

Right Area

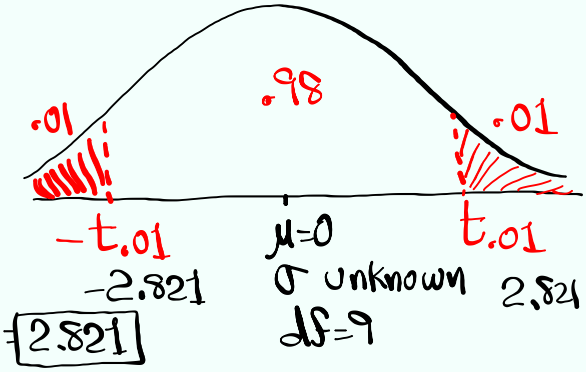
$\alpha/2 = .01$

$\alpha = .02$

left Area $1-\alpha = .98$

df

$t_{.01} = \text{invT}(.99, 9) = \boxed{2.821}$



Jan 29-5:31 PM

find $\pm t_{\alpha/2}$ for $\alpha=.04$ with df = 15.

$\alpha/2 = \frac{.04}{2} = .02$ ← Right Area

← Left Area

$1-\alpha = .96$ ← Middle Area

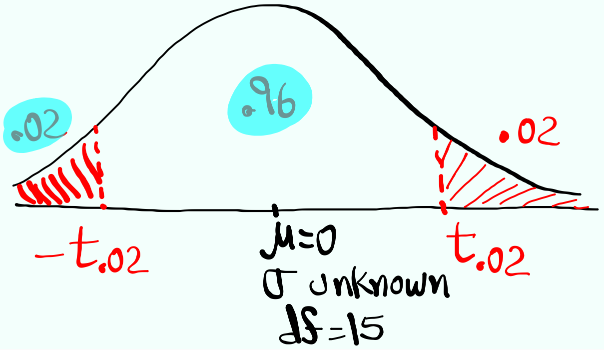
96% ← Confidence level

$t_{.02} = \text{invT}(.98, 15)$

Left Area

df

$= \boxed{2.249}$



Jan 29-5:35 PM

Find $\pm t_{\alpha/2}$ for 88% C-level with $df = 99$.

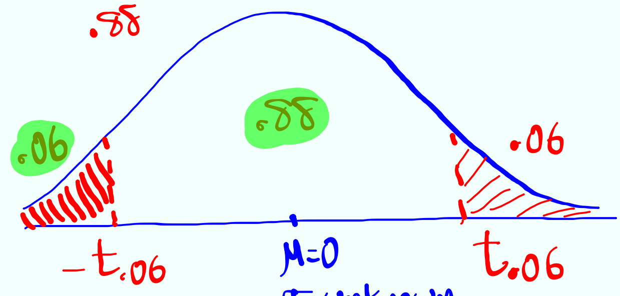
Middle Area

.88

$1-\alpha = .88$

$\alpha = .12$

$\alpha/2 = .06$



σ unknown

$df = 99$

$t_{.06} = \text{invT}(.94, 99) = \boxed{1.568}$

↑ Left Area

↖ df

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

Jan 29-5:39 PM

what is degrees of freedom?

degrees of freedom is determined by topics.

Non-Statistical way:

13 students

I bring 13 donuts, one for each.

Tyler	has	13	choices.
Larissa	"	12	"
William	"	11	"
⋮	|	|	|
Ean	has	no choice (1 donut left)	

$\boxed{df = 12}$

Jan 29-5:43 PM

Erick has 7 clean shirts.

He only wears clean shirt daily.

Monday	7	choices
Tuesday	6	"
Wednesday	5	"
⋮		
Sunday	No choice (1 clean shirt left)	

df = 6

Jan 29-5:47 PM

SG 21

Estimating Parameters

Sample ↔ Statistic

Population ↔ Parameter

To estimate parameter, we work with similar statistic

To estimate

Population Proportion p

Population Mean μ

we use

Sample Proportion $\hat{p}$ p-hat

Sample Mean $\bar{x}$

Point-estimates

For p, we use $\hat{p}$

For μ, we use $\bar{x}$

Jan 29-6:02 PM

Estimation of a parameter is a range of Values

$\underbrace{\text{range of Values}}_{\text{Confidence Interval}}$

Confidence Interval is a range of values in order to estimate a Parameter.

Every Confidence Interval comes with Confidence level.

$\underbrace{\text{Confidence level.}}_{\text{Middle Area}}$

$(1-\alpha)\cdot 100\%$

If C-level is not given, we use 95% C-level.

If α not given, we use .05.

Jan 29-6:07 PM

Confidence Interval for Population proportion

$\underbrace{\text{Population proportion}}_{p}$

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

Sample Proportion

Point-Estimate

Margin of error

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

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

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

x ← # of favorable responses

n ← Sample Size

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

Jan 29-6:13 PM

I surveyed 100 College students and 80 of them were in favor of face-to-face classes.

Find 98% Confidence interval for the prop. of all students in favor of in-Person classes.

$n = 100$

$x = 80$

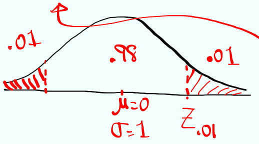
$\hat{p} = \frac{x}{n} = \frac{80}{100} = .8$

$\hat{q} = 1 - \hat{p} = .2$

C-level: .98

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

$= 2.326\sqrt{\frac{(.8)(.2)}{100}} \approx \boxed{.09}$



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

$= \boxed{2.326}$

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

$.8 - E < p < .8 + E$

$.8 - .09 < p < .8 + .09$

$\boxed{.71 < p < .89}$

Between 71% to 89% of all students are in favor of in-Person classes.

I am 98% Confident.

using TI:

STAT → TESTS ↓

1 - PropZInt

$x = 80$

$n = 100$

C-level: .98

Calculate

$\boxed{.71 < P < .89}$

Jan 29-6:18 PM

I Surveyed 250 students and 40 were Smokers.

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

$n = 250$

$x = 40$

C-level: .9

STAT

TESTS

1-PropZInt

$\boxed{.12 < P < .20}$

I am 90% Confident that between 12% & 20% of all students are Smokers.

$E = \frac{.20 - .12}{2} = \boxed{.04}$

$\hat{p} = \frac{.20 + .12}{2} = \boxed{.16}$

Jan 29-6:32 PM

I Surveyed 125 College Students and 8% of them were on diet to lose Weight.

Find 99% Conf. interval for the prop. of all Students that are on diet to lose weight.

$<P<$

$n=125$

$\hat{p}=.08$

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

$x=n\hat{p}$

if decimal → Round-up

$x=n\hat{p}$

$=125(.08)=10$

C-level: .99

$\boxed{\text{STAT}}$

$\boxed{\text{TESTS}}$

$\boxed{\text{1-PropZInt}}$

$x=10$

$n=125$

C-level: .99

$\boxed{\text{Calculate}}$

$E=\frac{.14-.02}{2}=\boxed{.06}$

$\hat{p}=\frac{.14+.02}{2}=\boxed{.08}$

$\boxed{.02<P<.14}$

Jan 29-6:39 PM

CNN Surveyed 240 Voters and 72% of them trusted the result of election.

$n=240$

$\hat{p}=.72$

$x=n\hat{p}=240(.72)\Rightarrow\boxed{x=173}$

if decimal → Round-up

Find Conf. interval for the prop. of all Voters that trust the result of election.

NO C-level

⇒ Use .95

$x=173$

$n=240$

C-level: .95

1-PropZInt

$\boxed{.66<P<.78}$

$E=\frac{.78-.66}{2}=\boxed{.06}$

$\hat{p}=\frac{.78+.66}{2}=\boxed{.72}$

we are 95% Confident that between 66% & 78% of all voters trust the election.

Jan 29-6:47 PM