Statistics

Lecture 12

Feb 19-8:47 AM

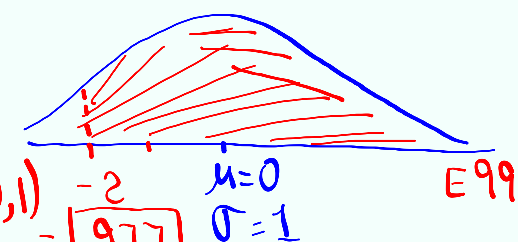
Class QZ 7

Drawing, labeling, shading, and full TI command required.

1) Find $P(Z > -2)$

$= \text{normalcdf}(-2, E99, 0, 1)$

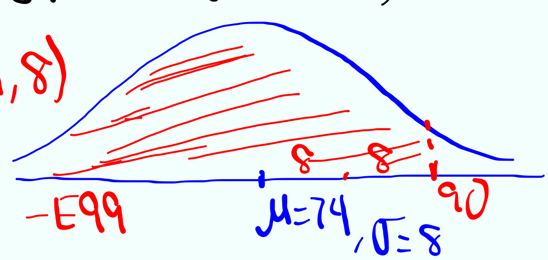
$= \boxed{.977}$



2) Given $N(74, 8)$, find $P(x < 90)$

$= \text{normalcdf}(-E99, 90, 74, 8)$

$= \boxed{.977}$



Jan 29-4:04 PM

A normal prob. dist. has a mean of 125 and standard dev. of 30.

If we take samples of size 16, find ($n = 16$) CLT

1) $\mu_{\bar{x}} = \mu = \boxed{125}$

2) $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{16}} = \frac{30}{4} = \boxed{7.5}$

Jan 29-4:45 PM

Ages of certain breed of dog has a normal prob. dist. with the mean of 12 years and standard dev. of 3 yrs.

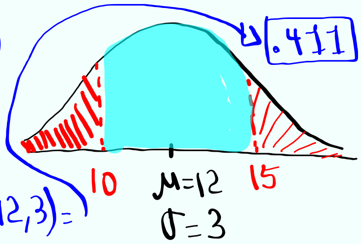
$N(12, 3)$

If we randomly select one of this breed of dogs, find the prob. that his/her age is below 10 yrs or above 15 yrs.

$P(x<10 \text{ OR } x>15)$

$= 1 - P(10 < x < 15)$

$= 1 - \text{normalcdf}(10, 15, 12, 3) = \boxed{.411}$



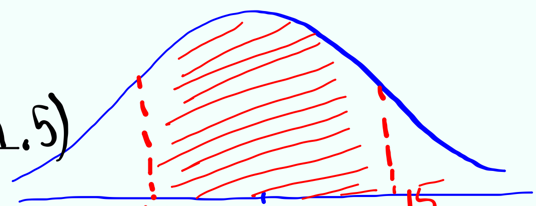
Jan 29-4:48 PM

If we randomly select 4 of these dogs, find the prob. that their mean age $\bar{x}$ will be between 10 & 15 yrs.

$P(10<\bar{x}<15)$

$= \text{normalcdf}(10, 15, 12, 1.5)$

$= \boxed{.886}$



CLT $\begin{cases} \mu_{\bar{x}} = \mu = 12 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{4}} = \frac{3}{2} = 1.5 \end{cases}$

Jan 29-4:55 PM

Ages of teacher in PUSD has a normal dist with mean of 48.5 yrs and standard dev. of 7.5 yrs.

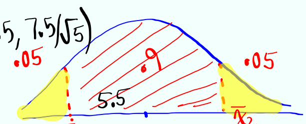
$N(48.5, 7.5)$

For randomly selected group of 5 teachers ($n=5$) from the district, find two mean ages ($\bar{x}$) that separate the middle 90% from the rest. Round to 1-dec.

$\bar{x}_1 = \text{invNorm}(.05, 48.5, 7.5/\sqrt{5})$

$= 42.983$

$\approx \boxed{43.0}$



CLT $\begin{cases} \mu_{\bar{x}} = \mu = 48.5 \\ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{7.5}{\sqrt{5}} \end{cases}$

$\bar{x}_2 = \text{invNorm}(.95, 48.5, 7.5/\sqrt{5})$

$= 54.017 \approx \boxed{54.0}$

SG 18, 19, 20 ✓

Jan 29-4:59 PM

Critical Value $z_{\alpha/2}$

α Alpha

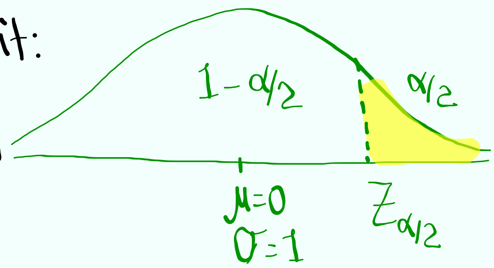
α Significance level

$0 < \alpha < 1$

$z_{\alpha/2}$ Separate the right tail with area $\alpha/2$ from the rest.

How to find it:

Use $\boxed{\text{invNorm}}$



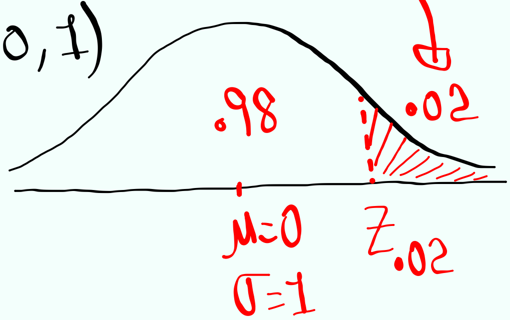
Jan 29-5:10 PM

Find $z_{.02}$ → Right area

→ $\alpha/2 = .02$ $\boxed{\alpha = .04}$

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

$= \boxed{2.054}$



Jan 29-5:14 PM

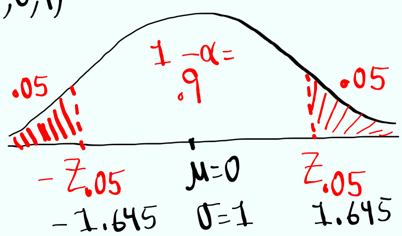
Find $\pm Z_{\alpha/2}$ for $\alpha = .1$

$\alpha/2 = .05$

$\pm Z_{.05}$

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

$= \boxed{1.645}$



$\alpha \rightarrow$ Significance level

$\alpha/2 \rightarrow$ Area of right-tail

$0 < \alpha < 1$

$1 - \alpha$ is the middle area

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

Jan 29-5:18 PM

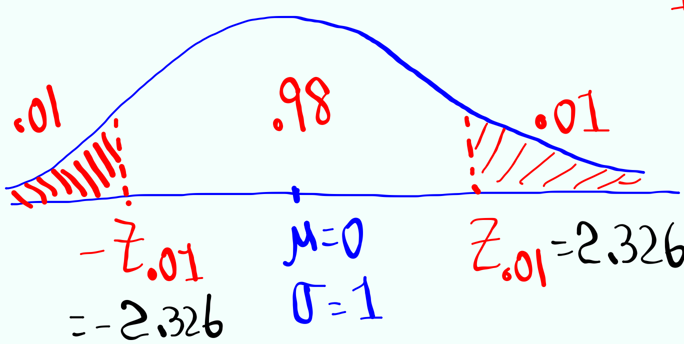
Find $\pm Z_{\alpha/2}$ for 98% confidence level

Middle Area .98

$1 - \alpha = .98$

$\alpha = .02$

$\alpha/2 = .01$



$Z_{.01} = \text{invNorm}(.99, 0, 1) = \boxed{2.326}$

Jan 29-5:24 PM

Find $\pm z_{\alpha/2}$ for 90% C-level.

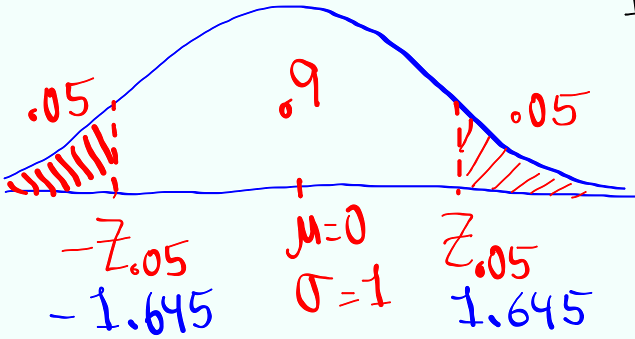
Confidence level

Middle Area $= .9$

$1 - \alpha = .9$

$\alpha = .1$

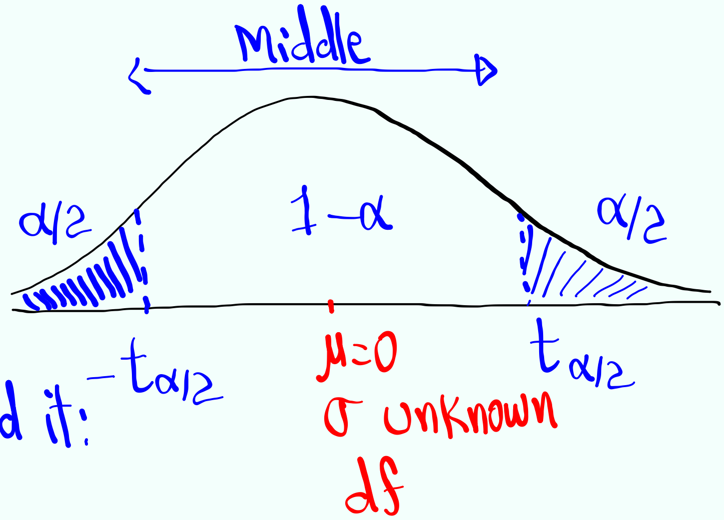
$\alpha/2 = .05$ area of each tail



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

Jan 29-5:29 PM

Critical Value $t_{\alpha/2}$:



How to find it:

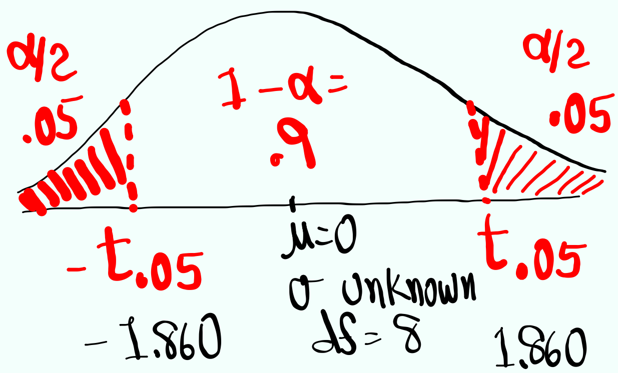
$\boxed{\text{invT}}$

Jan 29-5:34 PM

find $t_{.05}$ with $df = 8$.

$\alpha/2 = .05$ $\alpha = .1$

Right-Tail area



$\text{invT}(.95, 8) = \boxed{1.860}$

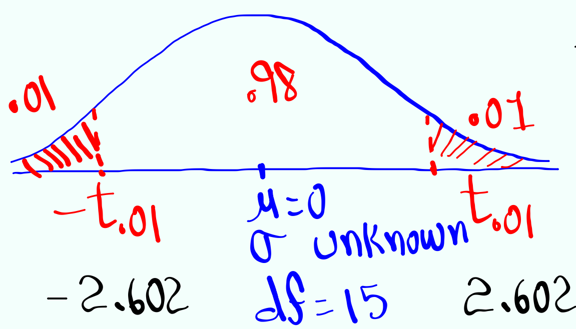
Jan 29-5:37 PM

find $\pm t_{\alpha/2}$ for $\alpha = .02$ with $df = 15$.

$\alpha/2 = .01$ Area of each tail

$1 - \alpha = .98$ Middle Area

$.98 = 98\%$ C-level



$t_{.01} = \text{invT}(.99, 15)$

$= \boxed{2.602}$

Jan 29-5:41 PM

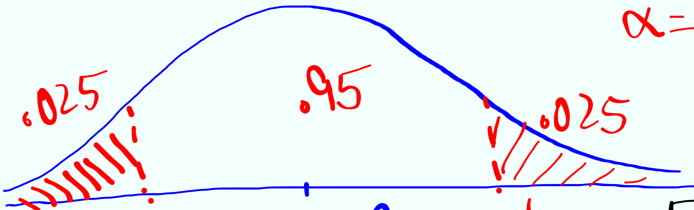
Find $\pm t_{\alpha/2}$ for $\underbrace{95\%\text{ C-level}}$ with $df = 24$.

Middle Area .95

$1 - \alpha = .95$

$\alpha = .05$

$\alpha/2 = .025$



$-t_{.025}$ $\quad \mu = 0$ $\quad t_{.025} = \boxed{2.064}$

σ unknown

-2.064 $\quad df = 24$ $\quad t_{.025} = \text{invT}(.975, 24) =$

Jan 29-5:47 PM

Estimating Parameters:

Population ⟷ Parameter

Sample ⟷ Statistic

We use Statistic to estimate Parameters.

We use	to estimate
Sample Proportion $\hat{P}$	Pop. Prop. P
Sample Mean $\bar{x}$	Pop. Mean μ
Sample Standard dev. S	Pop. standard dev. σ

Point-estimate is our best guess

Jan 29-6:11 PM

Our estimation of a parameter will be $\underbrace{\text{a range of values.}}_{\text{Confidence Interval}}$

Every confidence interval comes with confidence level $(1-\alpha)\cdot 100\%$

α ← Significance level.

Popular C-level are

90% , 95% , 98% , 99%

$\boxed{\text{If C-level is not given, we use 95\% C-level.}}$

Jan 29-6:19 PM

Estimating Population Proportion P:

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

$\hat{p}$: Sample Proportion

E: Margin of error

$\hat{p} = \dfrac{x}{n}$ ← x: # of favorable Responces; n: Sample Size

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

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

$z_{\alpha/2}$: Critical Value for $(1-\alpha)\cdot 100\%$ C-level.

Jan 29-6:23 PM

I Surveyed 100 Students, and 20 of them were Smokers.

$\hat{p} = \frac{x}{n} = \frac{20}{100} = .2 \qquad \hat{q} = 1 - \hat{p} = .8$

I want to construct 90% Conf. interval for the prop. of all Students that Smoke.

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

$.2 - .07 < p < .2 + .07$

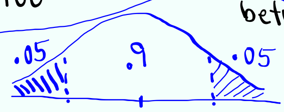
$E = Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$

$\boxed{.13 < p < .27}$

$= 1.645 \cdot \sqrt{\frac{(.2)(.8)}{100}} \approx \boxed{.07}$

we are 90% confident that between 13% and 27% are Smokers

C-level: .9



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

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

Jan 29-6:29 PM

I Surveyed 250 students, and 100 of them in favor of online classes.

$\hat{p} = \frac{x}{n} = \frac{100}{250} = \boxed{.4} \qquad \hat{q} = 1 - \hat{p} = \boxed{.6}$

C-level: .98

Let's construct $\boxed{\text{98\% Conf. interval}}$ for the prop. of all students in favor of online classes.

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

$.4 - .07 < p < .4 + .07$

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

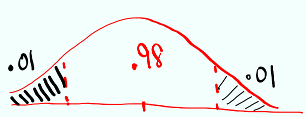
$\boxed{.33 < p < .47}$

$= 2.326 \cdot \sqrt{\frac{(.4)(.6)}{250}}$

$\approx \boxed{.07}$

we are 98% confident that between 33% & 47% of all students are in favor of online classes.

C-level: .98



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

$Z_{.01} = \text{invNorm}(.99, 0, 1) = \boxed{2.326}$

Jan 29-6:37 PM

How to use TI:

$\boxed{STAT} \rightarrow \boxed{TESTS} \downarrow \boxed{1-PropZInt}$

$.327 < p < .472$

$\boxed{.33 < p < .47}$

$E = \frac{.47 - .33}{2} = \boxed{.07}$

$\hat{p} = \frac{.47 + .33}{2} = \boxed{.4}$

x: 100

n: 250

C-level: .98

$\boxed{\text{Calculate}}$

Jan 29-6:46 PM

I randomly selected 400 voters, 32% of them were in favor of ICE operation.

$\hat{p} = .32$

$\hat{q} = 1 - \hat{p} = .68$

How many of them were in favor of ICE operation? $\hat{p} = \frac{x}{n} \rightarrow x = n\hat{p}$

if decimal, Always Round up

$x = 400(.32) = 128$

Construct conf. interval for the Prop. of all voters in favor of ICE operation

No C-level $\rightarrow$ use .95

1-PropZInt $\boxed{.27 < p < .37}$

x: 128

n: 400

C-level: .95

$E = \frac{.37 - .27}{2} = \boxed{.05}$

$\hat{p} = \frac{.37 + .27}{2} = \boxed{.32}$

Jan 29-6:52 PM