Statistics

Lecture 14

Feb 19-8:47 AM

Class QZ 7

Consider a binomial Prob. dist with $n = 100$ and $p = .5$. x is # of Successes

1) $P(x \le 60) = \text{binomcdf}(\overset{n}{100}, \overset{p}{.5}, \overset{x}{60}) = \boxed{.982}$

at most

2) $P(x = 45) = \text{binompdf}(\overset{n}{100}, \overset{p}{.5}, \overset{x}{45}) = \boxed{.048}$

Exactly

$\mu = np = 100(.5) = \boxed{50}$

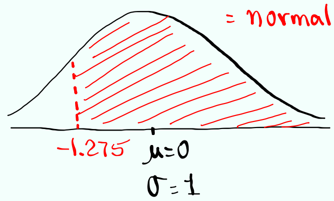
$\sigma^2 = npq = 100(.5)(.5) = \boxed{25}$

$\sigma = \sqrt{\sigma^2} = \sqrt{25} = \boxed{5}$

Jul 22-6:58 PM

find $P(Z > -1.275)$

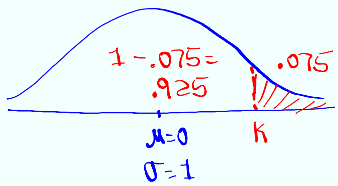
$= \text{normalcdf}(-1.275, E99, 0, 1)$



$= \boxed{.899}$

find K such that $P(Z > K) = .075$

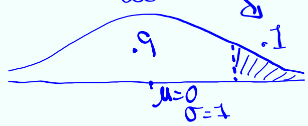
Right ↑ Area



$K = \text{invNorm}(.925, 0, 1)$

$= \boxed{1.440}$

find $Z_{.1}$



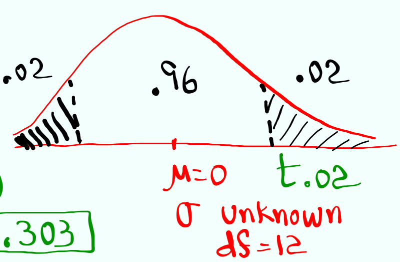
$Z_{.1} = \text{invNorm}(.9, 0, 1)$

$= \boxed{1.282}$

Jul 23-4:35 PM

Use t-dist with df=12

1) find $t_{\alpha/2}$ for 96% C-level.



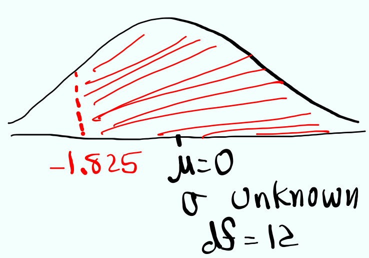
$t_{.02} = \text{invT}(.98, 12)$

$= \boxed{2.303}$

find $p(t > -1.825)$

2nd VARS

tcdf(L: -1.825, U: E99, df: 12)



$= \boxed{.954}$

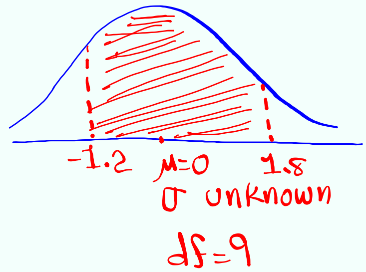
Jul 23-4:41 PM

Consider t-dist with $df=9$, find

$P(-1.2<t<1.8)$

$=\text{tcdf}(-1.2,\ 1.8,\ 9)$

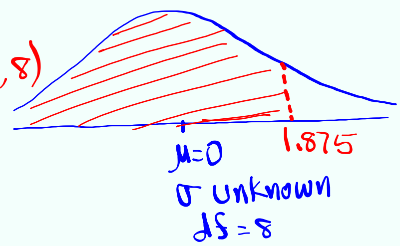
$=\boxed{.817}$



find $P(t<1.875)$ with $df=8$.

$=\text{tcdf}(-E99,1.875,8)$

$=\boxed{.951}$



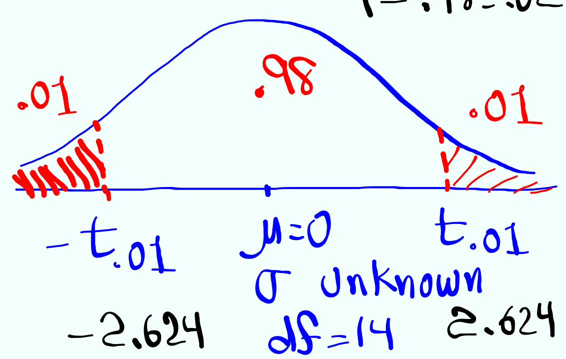
Jul 23-4:47 PM

find $\pm t_{\alpha/2}$ for $\underbrace{98\%\text{ C-level}}_{\text{Middle Area}}$ & $df=14$.

$t_{.01}=\text{invT}(.99,\ 14)$

$=\boxed{2.624}$

$1-.98=.02$



Jul 23-4:53 PM

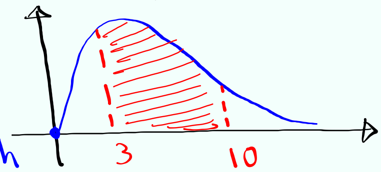
F - Dist

1) Graph begins at 0, and it is skewed to the right.

2) Not symmetric, Total Area is 1.

3) It comes with 2 df. Ndf & Ddf

4) use Fcdf (L , U, Ndf, Ddf)

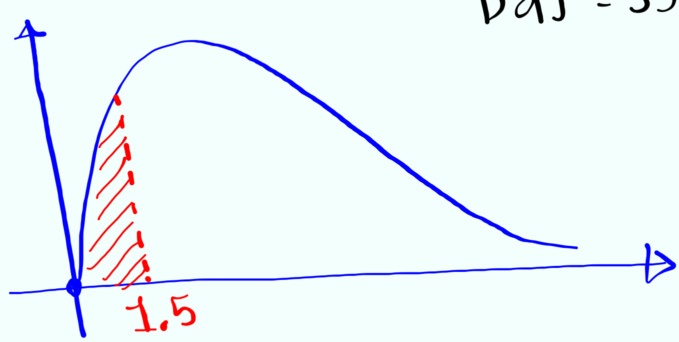


Find $P(3 < F < 10)$ with
Ndf = 4
Ddf = 25

Fcdf(3, 10, 4, 25) = $\boxed{.038}$

Jul 23-4:57 PM

find $P(F < 1.5)$ with Ndf = 4
Ddf = 35



Fcdf(0 , 1.5 , 4 , 35) = $\boxed{.777}$

Jul 23-5:03 PM

SG 24

Testing claim:

claim is a statement about one or more Parameters.

I claim 35% of all voters are undecided.

Pop. Proportion

$P = .35$

I claim the mean score of all exams in a math class is at least 80.

$\mu \geq 80$

College claims that standard deviation of ages of all students is less than 10

$\sigma < 10$

Jul 23-5:53 PM

why do we need to test a claim?

we want to know if claim is valid or invalid.

If claim is valid, we ~~Support~~ Fail-to-Reject it.

If claim is invalid, we reject it.

Possible errors:

claim is valid but we reject it.

claim is invalid but we fail-to-reject it.

---

when testing Process is complete, our final conclusion should be

Reject the claim OR FTR the claim

(claim is invalid) (claim is valid)

Jul 23-5:57 PM

Testing Methods:

1) Traditional Method

2) P-Value Method

3) Confidence Interval Method

Regardless of method used, final Conclusion must be the Same.

Possible Errors

Claim \ Action	Reject	Fail-to-Reject
Valid	Error	Not Error
Invalid	Not error	Error

Jul 23-6:03 PM

Testing Types:

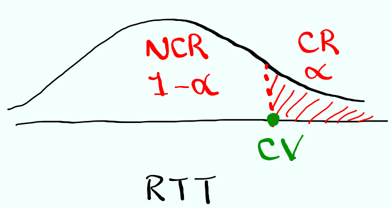
1) Right-Tail Test RTT

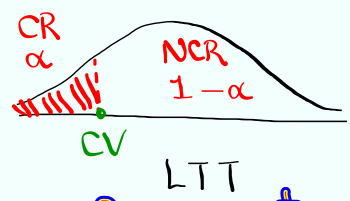
2) Left-Tail Test LTT

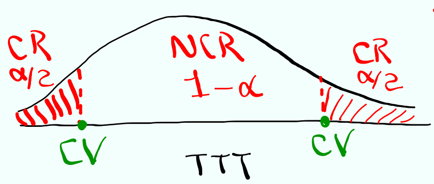
3) Two-Tail Test TTT

With every testing, we have a Significance level α, $0<\alpha<1$

α is the area of the Critical Region





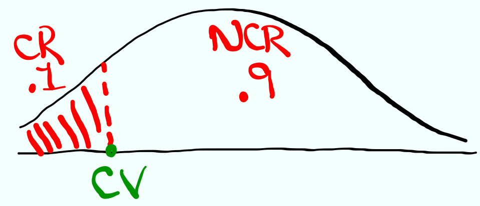


If α not given $\Rightarrow$ use .05

Jul 23-6:09 PM

$\alpha = .01$, RTT

$\alpha = .1$, LTT



$\alpha = .05$, TTT

Jul 23-6:16 PM

Testing Process:

1) Set-up H_0 & H_1 ← Alternative Hypothesis

↑ Null Hypothesis

2) Find all Critical Values
Drawing, labeling, shading, TI Command required.

3) Find Computed Test statistic (CTS) and P-Value (P).
Formula or TI command required.

4) Use Testing chart to determine the validity of H_0 & H_1.

5) Final conclusion about the claim.
claim could be H_0.
claim " " H_1.
H_0 valid $\Leftrightarrow$ H_1 invalid
H_0 invalid $\Leftrightarrow$ H_1 Valid

Jul 23-6:19 PM

More on H_0 & H_1:

H_0 must contain = sign. $=, \geq, \leq$

H_1 Cannot have = Sign. $\neq, <, >$

Key words for H_0:

is, equal, at least, at most, ...

Keywords for H_1:

Not equal, not, different, more than, less than, below, above, exceed, ...

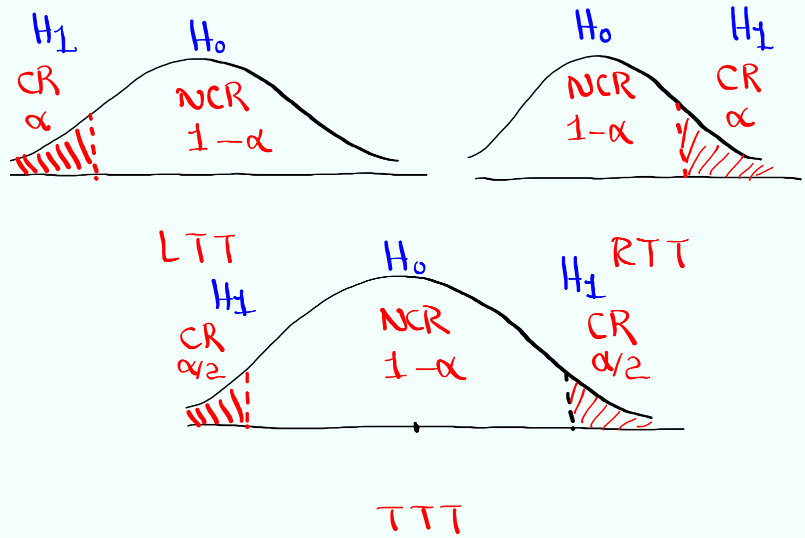
Jul 23-6:26 PM

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

H_1 tells us what type of testing we are doing.

$\boxed{\begin{array}{l} H_0 \text{ is valid in NCR} \\ H_1 \text{ is valid in CR.} \end{array}}$

Jul 23-6:30 PM



$$P(H_0 \text{ valid}) = 1-\alpha = P(H_1 \text{ invalid})$$

$$P(H_0 \text{ invalid}) = \alpha = P(H_1 \text{ valid})$$

Jul 23-6:33 PM

Possible outcomes for H_0:

Action \ H_0	Valid	Invalid
Reject	Type I Error	Good Decission
FTR	Good Decission	Type II Error

$\boxed{\text{Type I error is when we reject a valid } H_0}$

$\boxed{\text{Type II error is when we fail to reject an invalid } H_0.}$

Jul 23-6:38 PM

CNN claims that 30% of all voters are undecided.

$H_0: p = .3$ claim

$H_1: p \neq .3$ TTT

College claims that the mean age of all students is below 32.5 Yrs.

$\mu < 32.5$

$H_0: \mu \geq 32.5$

$H_1: \mu < 32.5$ claim, LTT

Jul 23-6:44 PM

LA Times has an article and it says standard deviation of salaries of all nurses in LA County is more than \$500.

$\sigma > 500$

$H_0: \sigma \leq 500$

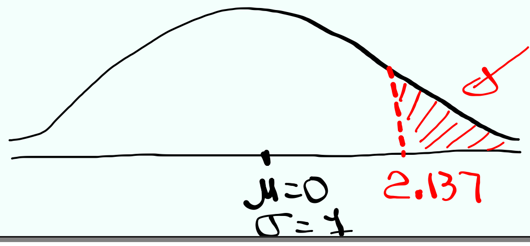
$H_1: \sigma > 500$ claim, RTT

Jul 23-6:47 PM

P-Value is the area of the tail marked by CTS (Computed Test Statistic)

Multiply by 2 only for TTT.

CTS $Z = 2.137$, RTT, find p-value.

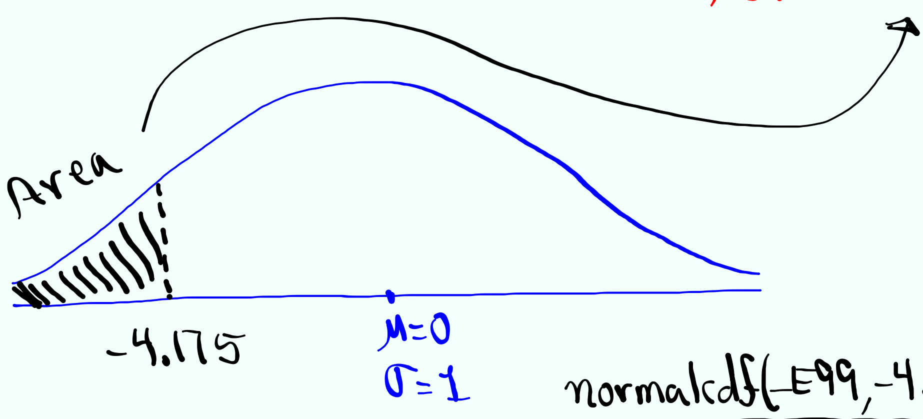


normalcdf(2.137, E99, 0, 1)

$= \boxed{.016}$

Jul 23-6:51 PM

CTS $Z = -4.175$, LTT, find P-value.



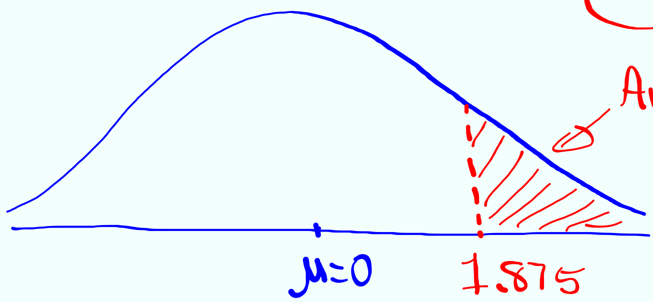
normalcdf(-E99, -4.175, 0, 1)

$= \boxed{1.491 \times 10^{-5}}$

Jul 23-6:54 PM

CTS $Z = 1.875$ TTT find p-value

Area $\cdot 2$



$\sigma = 1$

$2 \cdot \text{normalcdf}(1.875, E99, 0, 1)$

$= \boxed{.061}$

Jul 23-6:57 PM

Class QZ 8

Consider a geometric Prob. dist with

$P = .4$

TI Command required.

1) $P(x < 4)$

$= P(x \le 3) = \text{geometcdf}(.4, 3) = \boxed{.784}$

2) $P(x = 3) = \text{geometpdf}(.4, 3) = \boxed{.144}$

Jul 23-6:59 PM