Statistics

Lecture 19

Feb 19-8:47 AM

Testing Claims:

SG 23

A claim is made, our task is to determine whether claim is valid or not.

If claim is valid, then we support it.

If claim is invalid, then we reject it.

Possible errors:

Claim is valid but we reject it.

Claim is invalid but we support it.

Apr 22-6:58 PM

Claim could be made about any parameter.

It could be about

1) Population Proportion p
2) Population Mean μ
3) Population Standard deviation σ
4) ....

with any testing, there is a level of Significance α

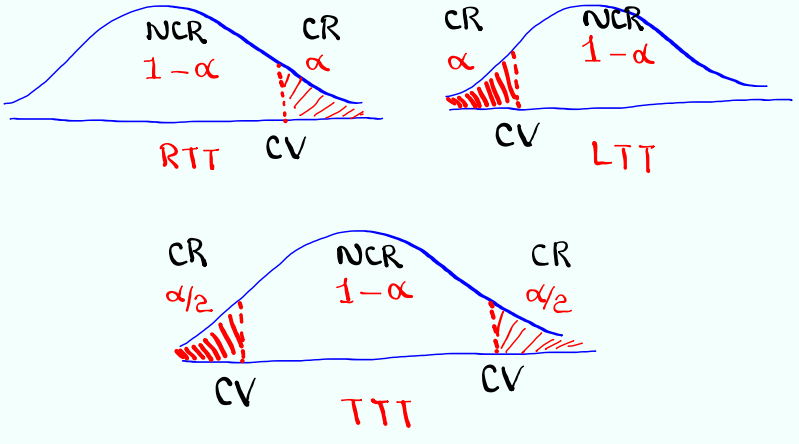
$0 < \alpha < 1$

When α not given, we use .05

Apr 22-7:02 PM

Testing Types:

1) Right-Tail Test RTT
2) Left-Tail Test LTT
3) Two-Tail Test TTT



Apr 22-7:07 PM

Testing Methods:

1) Traditional Method
2) P-Value Method
3) Confidence Interval Method

Regardless of the method, the conclusion must be the Same.

when claim is valid → we Support it.

when claim is invalid → we reject it.

Final Conclusion: Support

Reject the claim OR Fail-to-Reject the claim

Apr 22-7:14 PM

Testing Process

1) Set-up H_0 and H_1. ← Alternative Hypothesis H_a
   ↑ Null Hypothesis

2) find Critical Value(s).
   Drawing, labeling, shading, and full 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 and H_1.
   H_0 valid ⟷ H_1 invalid
   H_0 invalid ⟷ H_1 valid

5) Draw final Conclusion about the claim.
   claim could be H_0 or H_1 but not both at the same time.

Apr 22-7:21 PM

More on H_0 & H_1:

H_0 must contain the = Sign. $=, \geq, \leq$

H_1 Cannot contain the = Sign. $\neq, <, >$

Keywords for H_0:

is, equal, same, atleast, at most, ...

Keywords for H_1:

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

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

Apr 22-7:31 PM

CNN claims that 70% of all students have iPhone

$p = .7$ claim

H_0

H_0: $p = .7$ claim

H_1: $p \neq .7$ TTT

LA Times claims that the mean age of all nurses in So. Cal. is below 45 Yrs.

$\mu < 45$ claim

H_1

H_0: $\mu \geq 45$

H_1: $\mu < 45$ claim, LTT

Apr 22-7:38 PM

The dept claims that Standard deviation of all exam Scores is at most 10.

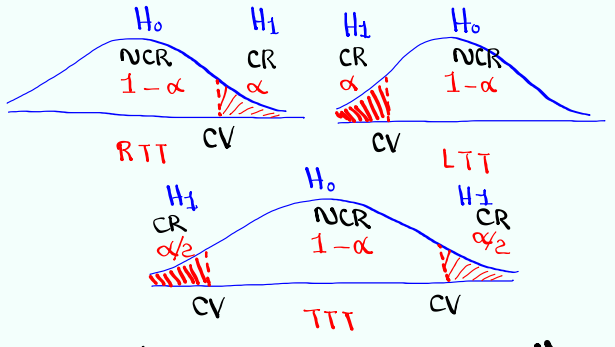
$\sigma \leq 10$ claim

H_0

$H_0: \sigma \leq 10$ claim

$H_1: \sigma > 10$ RTT

Apr 22-7:44 PM



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

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

Four-Possible outcomes For H_0:

Conclusion \ Reality	H_0 Valid	H_0 invalid
Support H_0	Good Decision	Type II error
Reject H_0	Type I error	Good Decision

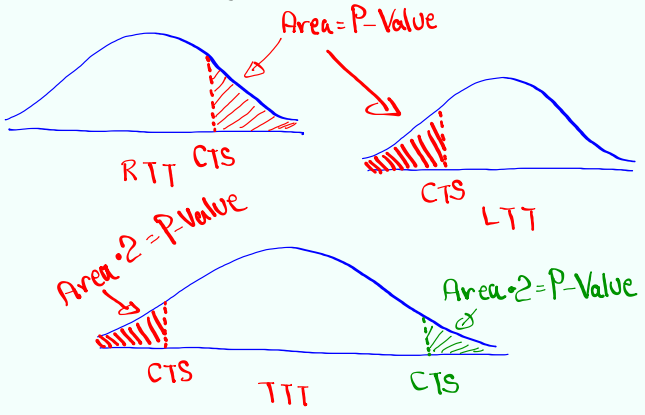
Apr 22-7:48 PM

what is P-Value?

P-Value is the area of the tail of graph of the dist. marked by CTS.

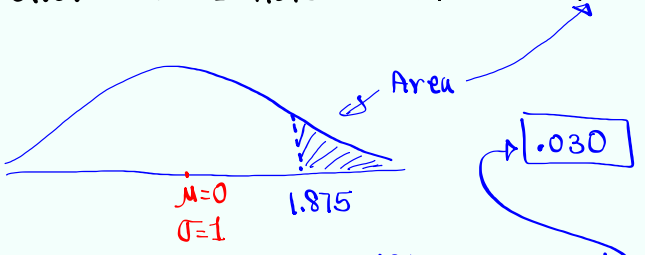
CTS → Computed Test Statistic

Multiply that area by 2 when Performing only TTT.



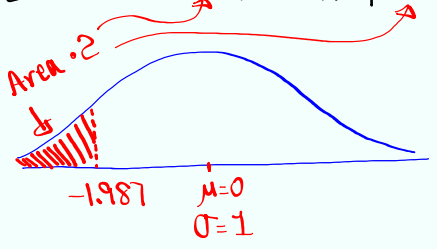
Apr 22-8:00 PM

Given CTS $Z = 1.875$ RTT Find P-Value.



P-Value = normalcdf(1.875, E99, 0, 1) =

Given CTS $Z = -1.987$ TTT Find p-value.



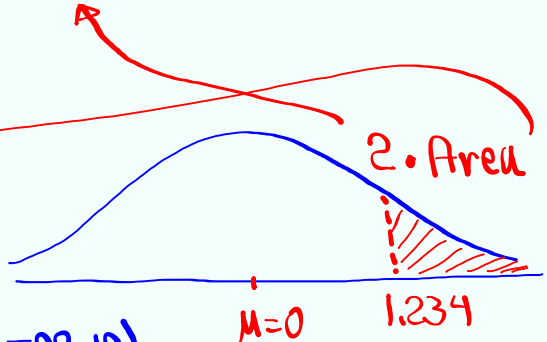
P-Value = 2 · normalcdf(−E99, −1.987, 0, 1)

= .047

Apr 22-8:06 PM

CTS $t = 1.234$ TTT $df = 19$

find P-value.



$P\text{-Value} = 2 \cdot tcdf(1.234, E99, 19)$

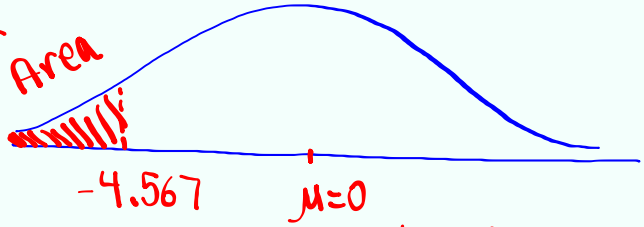
$= \boxed{.232}$

$\mu = 0$
σ unknown
$df = 19$

Apr 22-8:12 PM

CTS $t = -4.567$ LTT $df = 24$

find p-value.



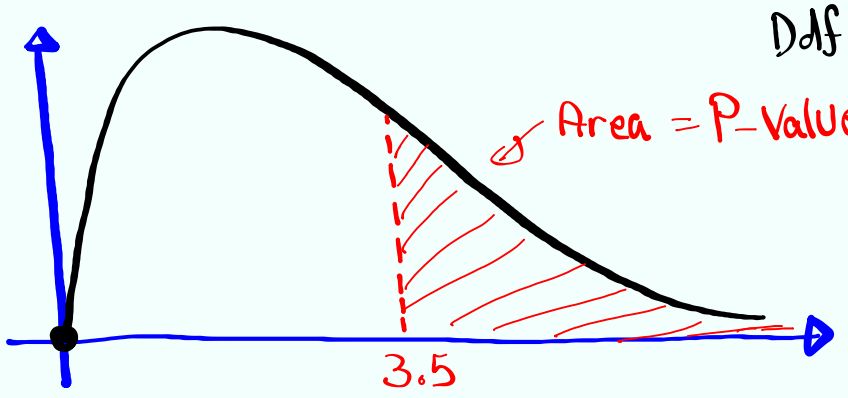
$P\text{-Value} = tcdf(-E99, -4.567, 24)$

$= \boxed{6.2 \times 10^{-5}}$

$\mu = 0$
σ unknown
$df = 24$

Apr 22-8:15 PM

Given CTS F = 3.5 RTT Ndf = 4
Ddf = 10

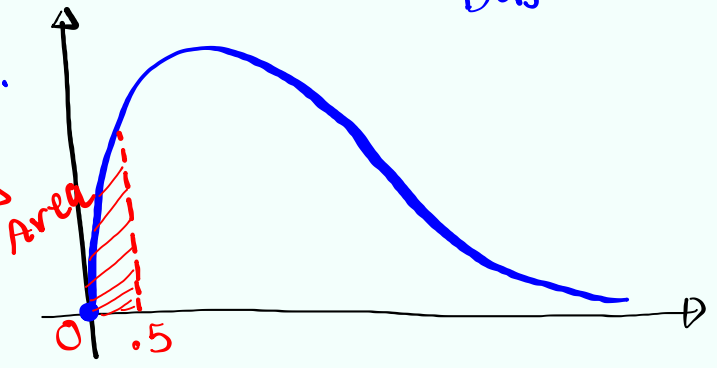


P-Value = fcdf(3.5, E99, 4, 10) = .049
L U Ndf Ddf

Apr 22-8:19 PM

Given CTS F = .5, LTT, Ndf = 6
Ddf = 6

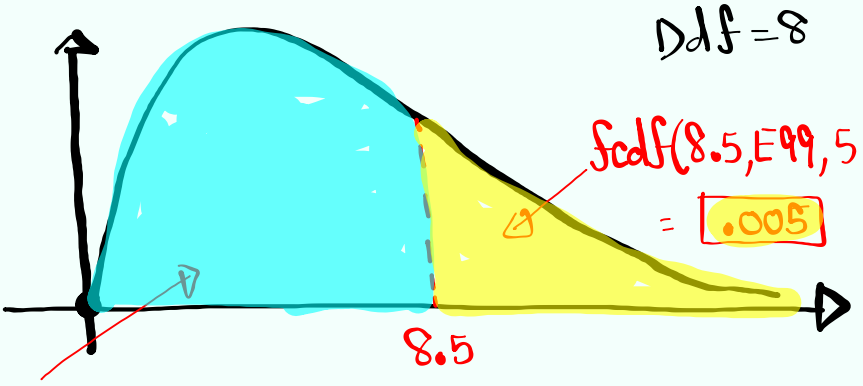
find p-value.



P-Value = fcdf(0, .5, 6, 6) = .210

Apr 22-8:22 PM

Given CTS F=8.5 TTT ndf=5 Ddf=8



fcdf(0,8.5,5,8) = .995

P-Value = 2 · Smaller area

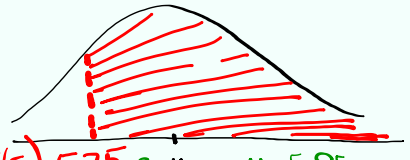
= 2(.005) = .010

Apr 22-8:25 PM

Gas Prices are normally dist. with the mean of $5.85 and standard deviation of $.45. N(5.85, .45)

If we randomly select n=3 gas stations, find the prob. that their mean $\bar{x}$ gas price is more that $5.75.

$P(\bar{x} > 5.75)$



= normalcdf(5.75, E99, 5.85, .45/√3)

5.75

CLT $\mu_{\bar{x}} = \mu = 5.85$, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{.45}{\sqrt{3}}$

≈ .650

Apr 22-8:30 PM

I flipped a fair coin 1600 times.

Success is to land tails.

$n=1600$

$p=.5$

$q=.5$

Find usual Range for # of tails.

95% Range $\rightarrow \mu \pm 2\sigma$

$800 \pm 2(20)$

$\mu = np$

$\sigma^2 = npq$

$\sigma = \sqrt{\sigma^2}$

$\mu = 1600(.5) = 800$

$\sigma^2 = 1600(.5)(.5) = 400$

$\sigma = \sqrt{\sigma^2} = \sqrt{400} = 20$

760 to 840

P(# tails is between 760 & 840, inclusive)

$P(760 \le x \le 840) = \text{binomcdf}(1600, .5, 840) -$

$\text{binomcdf}(160, .5, 759) =$ .957

Apr 22-8:36 PM

x	y
5	82
6	85
7	80
7	90
9	95
10	90

L1 L2

Find

1) $a = 71.346 \approx 71$

2) $b = 2.135 \approx 2$

$\rightarrow y \approx 71 + 2x$

3) $r^2 \approx 49\%$

LinReg(a+bx)

4) $r = .703$

Predict y if $x = 8$ assume

1) r is Significant

$y = 71 + 2x$

$= 71 + 2(8)$

$= 71 + 16$

$= 87$

2) r is not Sig.

use $\bar{y}$

$\bar{y}$ 87

Apr 22-8:43 PM