Statistics

Lecture 28

Feb 19-8:47 AM

Use the chart below to test the claim that $\sigma_1 > \sigma_2$. SG 29

No $\alpha \rightarrow .05$

Group 1	Group 2
$n_1 = 6$	$n_2 = 10$
$S_1 = 5$	$S_2 = 4$

1) $S_1 > S_2$ ✓

2) CTS $F = \frac{S_1^2}{S_2^2} = \frac{5^2}{4^2} = 1.5625$

3) $Ndf = n_1 - 1 = 5$, $Ddf = n_2 - 1 = 9$

$H_0: \sigma_1 \le \sigma_2$

$H_1: \sigma_1 > \sigma_2$ claim, RTT

2-SampF Test

$\boxed{\text{CTS } F = 1.5625 \quad \text{P-Value } P = .264}$ ✓

P-Value $> \alpha$

$.264 > .05$

H_0 Valid & H_1 invalid

Invalid claim $\rightarrow$ Reject the claim

May 27-6:50 PM

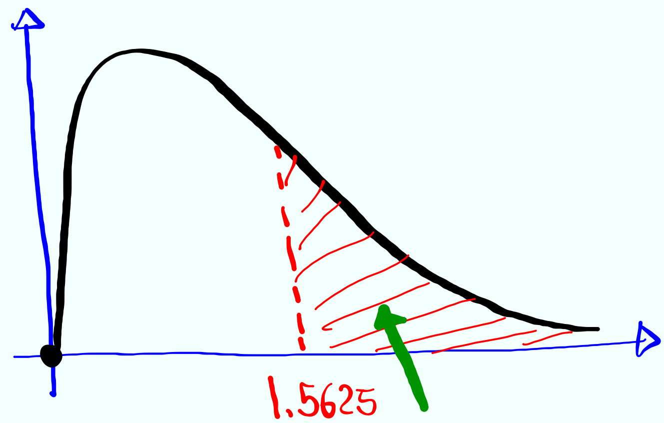
CTS $F = 1.5625$

RTT

$Ndf = 5$

$Ddf = 9$

Find P-Value → Area



$$P\text{-Value} = Fcdf(1.5625, E99, 5, 9)$$

$$\approx \boxed{.264}$$

May 27-6:59 PM

SG 30

Consider the matched-Pair dependent Sample below:

Before	After	L3
20	24	−4
15	18	−3
30	28	2
35	32	3
25	25	0

Before → L1

After → L2

→ ↑ L3

2nd 1 − 2nd 2 Enter

Use L3 with 1-Var Stats to find

1) $\bar{d} = \bar{x} = -.4$

2) $S_d = S_x = 3.050$

3) $S_d^2 = 9.3 = \frac{93}{10}$

4) Find 98% Conf. interval for the mean of all differences

$\boxed{-5.5 < \mu_d < 4.7}$

T Interval

$$E = \frac{4.7 - (-5.5)}{2} = \frac{10.2}{2} = \boxed{5.1}$$

May 27-7:02 PM

5) Test the claim that the mean of all differences is not zero.

$\mu_d \neq 0$

$H_0: \mu_d = 0$

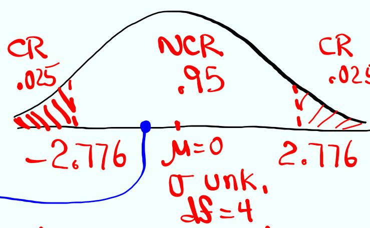
$H_1: \mu_d \neq 0$ claim, TTT

CV t TTT $\alpha = .05$

$df = n - 1 = 4$

$\boxed{\begin{array}{l} \text{CTS } t = -.293 \\ \text{P-Value } P = .784 \checkmark \end{array}}$

T-Test



$t = \text{invT}(.975, 4)$

CTS is in NCR

P-Value $> \alpha$

H_0 valid, H_1 invalid

Invalid claim → Reject the claim

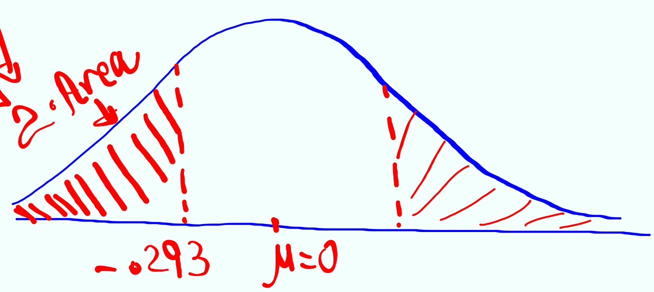
May 27-7:12 PM

CTS $t = -.293$

TTT

$df = 4$

Find P-Value



$2 \cdot \text{tcdf}(-E99, -.293, 4) = \boxed{.784}$

May 27-7:21 PM

Consider the table below: (SG 31)

x	y
2	8
3	10
4	12
4	13
5	15

$x \to L1$, $y \to L2$

Use $\boxed{\text{2-Var Stats}}$ to find

$\sum x = 18$ $\quad \sum y = 58$ $\quad \sum xy = 221$

$\sum x^2 = 70$ $\quad \sum y^2 = 702$ $\quad n = 5$

Use LinReg$(a+bx)$ to find

$a = 3.154 \approx 3$

$b = 2.346 \approx 2$ $\longrightarrow$ $\boxed{y \approx 3 + 2x}$

$r^2 = .980 = 98\%$ 98% of y-values are explained by x-values.

$r = .990$

r is close to 1

Linear Correlation is Significant

May 27-7:25 PM

Use $\alpha = .02$ to determine whether Linear Correlation is Significant or not.

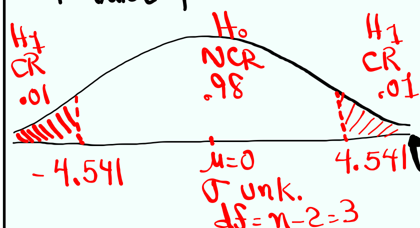
$H_0: \rho = 0 \to$ Linear Correlation is not Sign.

$\boxed{H_1: \rho \neq 0 \to \text{ " " " Significant.}}$

Use $\boxed{\text{LinReg TTest}}$ to find

CTS $t = 12.2$

P-Value $P = .001$



P-Value $\leq \alpha$
$.001 \quad .02$

H_0 invalid & H_1 Valid

Linear Correlation is Significant.

$t = \text{invT}(.99, 3)$

May 27-7:33 PM

Finding CTS t using Formula:

$$t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$$

$$= .990 \cdot \sqrt{\frac{5-2}{1-.980}}$$

$$= .990\sqrt{\frac{3}{.02}} \approx \boxed{12.125}$$

Predict y for $x = 3$

Since r is significant → use $\hat{y} \approx a + bx$

$y \approx 3 + 2x$

$\approx 3 + 2(3)$

$\approx \boxed{9}$

If r was not significand → use $\bar{y}$

$\bar{y} = 11.6$

May 27-7:42 PM

Given: $n = 6$, $y = 25.8 - 2.5x$, $r = -.6$,

$\boxed{\bar{y} = 18}$

Testing R

$H_0: \rho = 0$ Not

$H_1: \rho \neq 0$ Is

H_0 Valid

Predict y when $x = 4$.

CTS $t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$

$= -.6 \cdot \sqrt{\frac{6-2}{1-(-.6)^2}}$

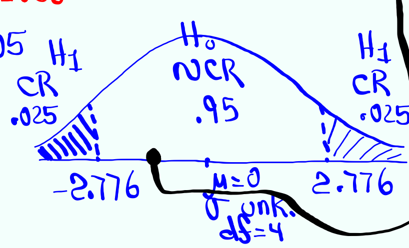
not significant

$= -.6 \cdot \sqrt{\frac{4}{1-.36}} = -.6 \cdot \sqrt{\frac{4}{.64}} \approx \boxed{-1.5}$

CV t TTT $\alpha = .05$

$df = n - 2 = 6 - 2 = 4$

$t = \text{invT}(.975, 4)$



May 27-7:50 PM

Comparing at least 3 Pop. means. SG 33

$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_K$

H_1: At least one mean is different. RTT

Method: ANOVA (Analysis of Variance)

CTS F $\quad$ Ndf $= K - 1$ $\quad$ $K \to$ # of groups

P-Value P $\quad$ Ddf $= n - K$ $\quad$ $n \to$ Total Sample Size

Use P-Value Method to proceed.

May 27-8:01 PM

L1 Morning			L2 Afternoon		L3 Evening		
75	92	88	78	85	72	88	93
80	100		96	99	80	100	75
			70				

$K = 3$ $\quad$ $n = 5 + 5 + 6 = 16$

ndf $= K - 1 = \boxed{2}$ $\quad$ Ddf $= n - K = \boxed{13}$

Use $\alpha = .1$ to test the claim that all means are the same.

$H_0: \mu_1 = \mu_2 = \mu_3$ claim

H_1: At least one mean is different, RTT

$\boxed{\text{STAT}}\ \boxed{\text{TESTS}}\ \boxed{\text{ANOVA(}}$ L1, L2, L3 $\boxed{\text{Enter}}$

CTS $F = .062$ $\quad$ H_0 Valid $\to$ Valid claim

P-Value $P = .940 > \alpha$ $\quad$ H_1 invalid

FTR the claim

May 27-8:07 PM

I randomly selected students from 4 different schools. Here are their ages.

L1 ELAC	L2 Mt. SAC	L3 Glendale	L4 UCLA
21 29 25 18 30	26 19 25 32 28	18 32 35 25 20	25 33 40 28 50

$K=4$ $n=5+5+5+5=20$ $NDF=K-1=3$

no $\alpha \to .05$ $DDF=n-K=16$

Test the claim that not all means are equal.

$H_0: \mu_1=\mu_2=\mu_3=\mu_4$

H_1: At least one mean is different, RTT claim

STAT
TESTS
ANOVA(L1,L2,L3,L4)
$\boxed{\text{Enter}}$

CTS $F=2.333$

P-Value $P=.113$

$\underset{.113}{\text{P-Value}} > \underset{.05}{\alpha}$

H_0 valid

H_1 invalid

Invalid claim

Reject

to reverse the conclusion

P-Value $\leq \alpha$

$.113 \leq \alpha$

choose $.12, .13, .14, .15, \ldots$

May 27-8:17 PM