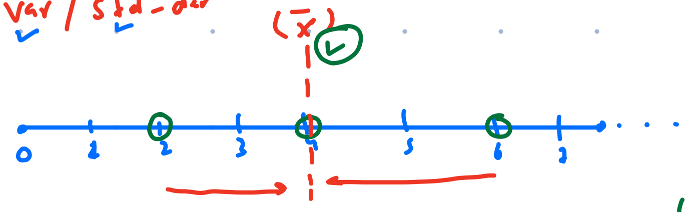
(logic) — (DS) = (Desc p) (Stats)

{ ① MCT - Mean / median / mode }
{ ② var / std-dev }

(mean) => (graph)

(problem)



(mean) info

(Mean) — $\bar{x}$

-4   0 → 4 (s)

-6   0 → 6 (sl)

-8   0 → 8 (L)

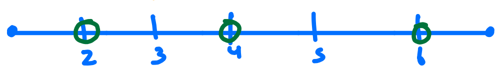
(spread)

(5 3 7 - 9, 10)

(Analyze)

(Variance) = $\boxed{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

Sample



2 ← [2, 4, 6] = $\boxed{\bar{x} = 4}$

$$\frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{2} = \frac{(-2)^2 + 0 + 2^2}{2} = \frac{4+4}{2} = \frac{8}{2} = 4$$

(love) = $\frac{Q}{a}$

(Quantification)

I love you 3000
" " 8000
" 19000

Cm

Punishment

[2, 3, 9] = 3.6

[1, 2, 19] → (1, 2, 19)

V = 8.3

bias

① [1 7 9 10]

② [2, 7, 15, 20]

(why) ?

V lower

m | Std

CI

O

L

S

1 3 GB
2 3 GB } Compress

$$Var = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$Std = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \rightarrow Same$$

$v_1 = 10,000 \rightarrow v_1 = 100$

$v_2 = 100 \rightarrow v_2 = 10$

$\boxed{v_1 > v_2}$ $\boxed{v_1 > v_2}$

1 (Pop) = diff sizes ?

(pop) large / small

Yes

1.4 B

250

(Avg) (100) (150) Exactly

153 $\bar{x}$

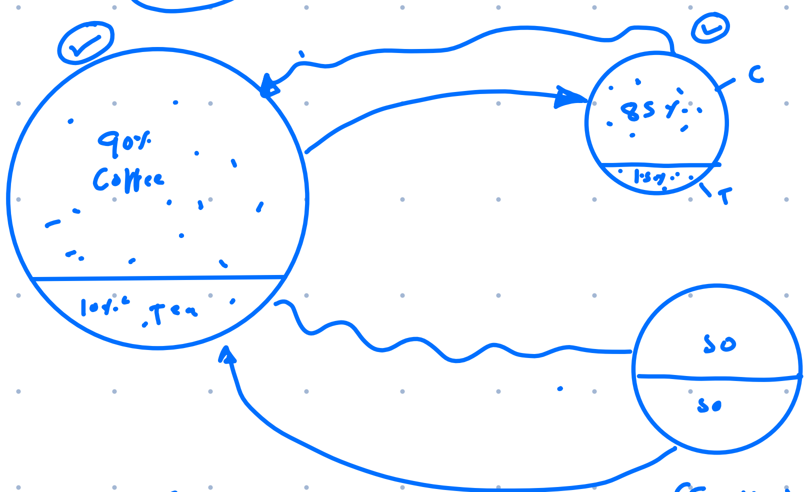
Wash

(pop_mean)

(pop_std)

Sample But, good / bad

representive Samp



biased Sample

(Survey)

(are you happy with our Cars)

OV

Chance IS

PaF

BS =

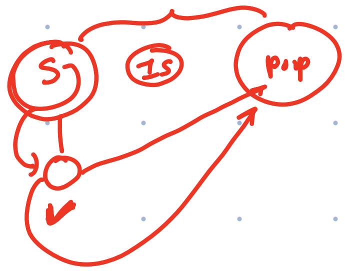
While descriptive statistics describes the data, inferential statistics is used to draw conclusions about the population based on statistical findings on sample analysis.

Ex

result

on pop

Idea



Mean | V
n | Sr

CI

(3.9 #) = Int

(Bell's curve / Gauss / Nor Dist)

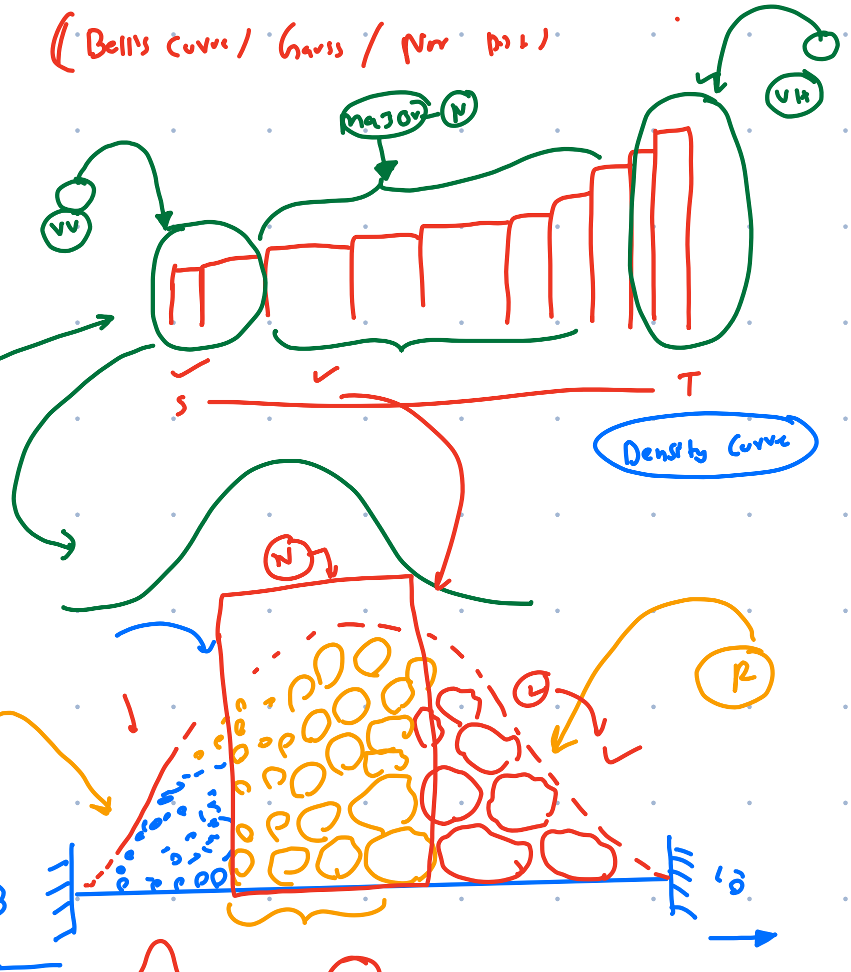
(Village) (200)

(height)

Shot

Pattern

(height)
(wei)
(wan)
n
ES

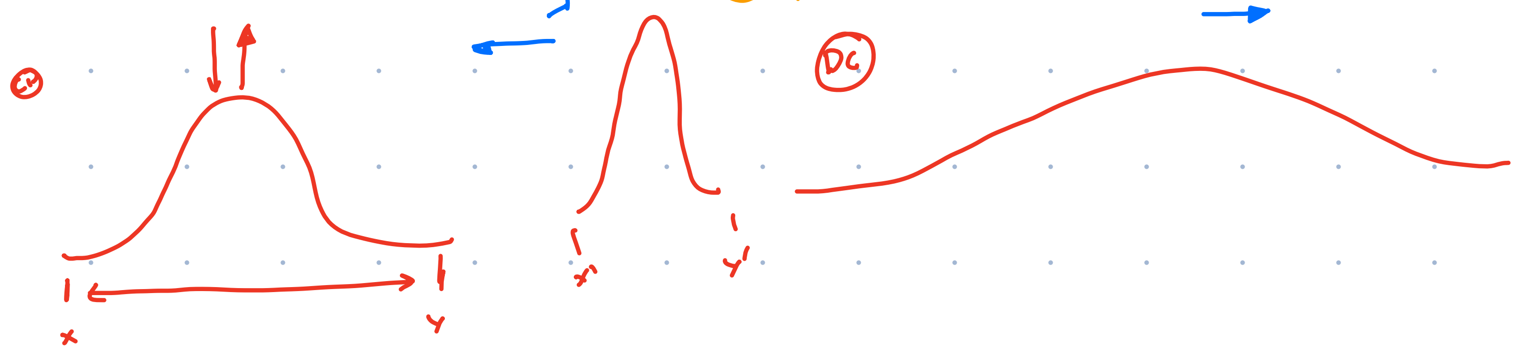


SDC

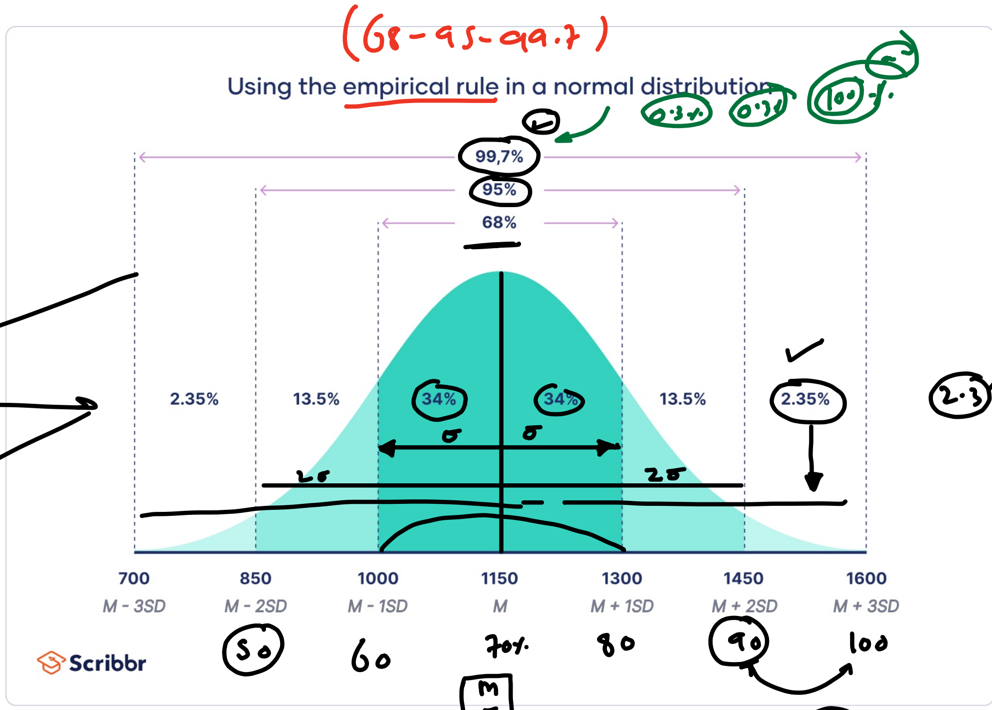
n

Priyadarsh pani

BH



(68 - 95 - 99.7)



0.3% 0.3% 100%

100

$\boxed{\bar{x} = 70\%}$

$\sigma = 10$

2.3

%

50 60 70% 80 90 100

68 = (60 - 80) %

60 M = M = M m+5 70+10 80

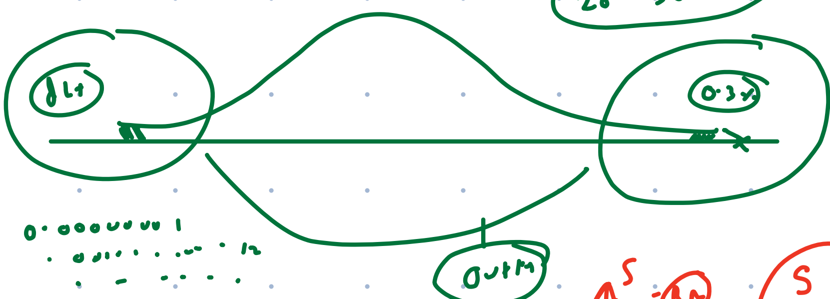
95

(BQ)

(99.7% why) — 100%

(Asymptotic 91)

25 35 45 100



0.3%

0.00000000 1

Outm

S - Bn

$\frac{S}{\sqrt{n}}$

Pan

$\{ CI = \bar{x} \pm Z \cdot \frac{S}{\sqrt{n}} \}$

$\boxed{CI = \text{don't Calculate} / \text{assume}}$

$\bar{x}$ = Sample mean

Z = Z_Critical Value / Z_Score

S = Std_Dev

n = Sample Size

$CI = \bar{x} \pm Z \cdot \frac{S}{\sqrt{n}}$

Margin of Error

Clear

Upper Bound

lower Bound

$\boxed{UB = \bar{x} + Z \cdot \frac{S}{\sqrt{n}}}$

$\boxed{LB = \bar{x} - Z \cdot \frac{S}{\sqrt{n}}}$

Pat

Ex — Def

1.4 Bharatis (Avg Cons of protein by Indians) = 1.4 — 5,00,000/-

5,00,000/- — guess

A random sample of 55 males were taken and they show that their average protein consumption is 122 g of protein with a standard

deviation of 14 g at 95% confidence interval, can you calculate the range?

Z — $\boxed{95\%}$ UB LB

A random sample of 55 males were taken and they show that their average protein consumption is 122 g of protein with a standard deviation of 14 g at 95% confidence interval, can you calculate the range?

99%

(Yes)

YES

$\bar{x} = 122$ grm

$n = 55$

$S = 14$

$Z = $ Z-table (1.96) 2.57

Clear

Z

SS

$UB = \bar{x} + Z \cdot \frac{s}{\sqrt{n}}$

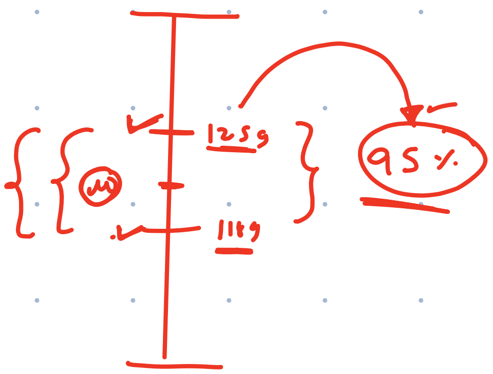
$= 122 + 1.96\left(\frac{14}{\sqrt{55}}\right)$

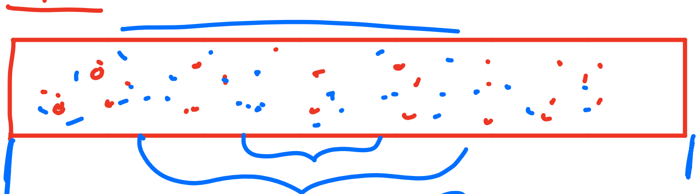
$\boxed{UB = 125\text{ g}}$

$LB = \bar{x} - Z \cdot \frac{s}{\sqrt{n}}$

$= 122 - 1.96\left(\frac{14}{\sqrt{55}}\right)$

$\boxed{LB = 118\text{ g}}$

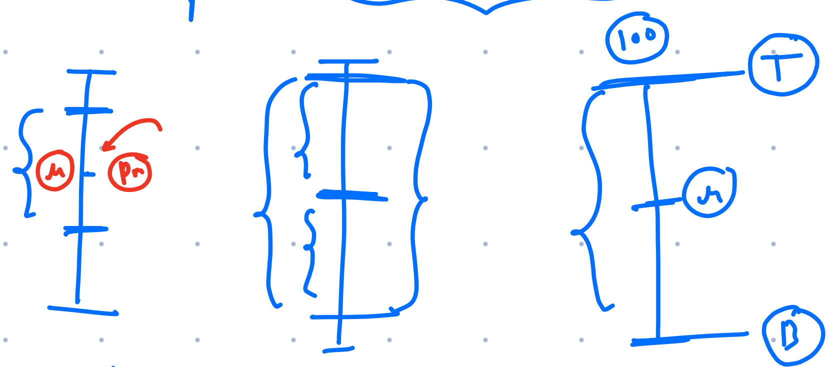




less

95 := c

5% := v



95 CI

(VOIP) = WebRTC

CI — Statistical range..

A confidence ~~central wall~~ is a statistical range with a specific probability that a population parameter will fall within. It represents the uncertainty in the calculation of an estimate and is expressed as a percentage usually given in the range of 95, 92,99, 93, 9468, it could be any range.

95 %

5 ✓

μ

σ

125

Range

118

95

95

5 %

mom

500 × (0.30) = 150

(CI) :

(Hyp test)

(assumption)

Mani

Nand

G

IS

USA

Nanda

Mani – Same (H0)

mani – Not Same

Two

(Ha)

(Taking assumpt)

① H0 : Null

② Ha : Alternate

(H0) — Null ⟶ (No change)

(Ha/H1) = Alternate ⟶ (Change)

IS

P

Ho

Ha

Ex : Survey 2019, Avg Sweet 23g

2025, 55 ⟶ 48g

H0 : Avg(μ) = $\boxed{23g}$

Ha : Avg(μ) ≠ $\boxed{23g}$

(>, <)

(Stand)

NO

T

Ex

(Choc) — Bar = 100grams

50,000/- → 600 → 99.25 grams

$H_0 : (\mu) = 100g$

$H_a : (\mu) \neq 100g$

(>, <)

H_0	H_a
H_a	H_0

=> H_0

H_a — H_1

H_0
H_1 / H_a

HT

P → H_0, H_a } apply Test → H_0 | H_a

H_0 → Accepted / Rejected

2

H_0 – T	H_0 – F
H_a – F	H_a – T

H_0 → H_a

P → H_0, H_a → T 2 F

H_0 – Accepted	H_0 – Rejected
H_a – Rejected	H_a – Accepted

None

Namaste folks,

lets start in 2-4 minutes

(Errors) in (HT) — (wrong) → (Error)

(Ground truth) => (actual reality)

Pred / Calc

Assump

IS → 100%

deepak (R)

kind
nice

(Ankit)

not kind
rude

Sarah — Pd

kind
nice

(R) (P)

✓ ⟷ ✗

✓ — ✓

Pat

RT Pl

+
–

–
+

+
+

–
–

ide

id

Cr ✓

Clo

Pat : +
Doc : +
(True positive)

(Pat : −)
(Doc : +) → 14 alone

[False positive) Error] → [Type-1 / α / FPE]

MI (FNF) D = MI

dah =

T I / II

Pat : +
Doc : − = 100

[False Negative error]

(Type-2 / β / FNE)

Pat : −
Doc : −
(True Negative)

Don

jdm

Cn

Guy : + Judge : + TP	Guy : − Judg : +
Guy : + Judge : −	Guy : − Judg : − TN

1.1

PS ⇒

[2019 − IT − 32,000/-
2019 → 45,000]

H_0 / H_a → (H_a)

(H_0)

1 One

PS < H_0, H_a → T/I/β/α → H_0 | H_a

PS

(Pred)

	Null Hypothesis is TRUE	~~Null Hypothesis is FALSE~~ H_a − True
Reject null hypothesis	Type I Error (False positive)	Correct Outcome! (True positive)
~~Fail to reject null~~ Accept hypothesis	Correct Outcome! (True negative)	Type II Error (False negative)

Real
Ho - T
Ho - T

Pred
Rej - Ho
Ho - T

Real
Ha - T

Ha ✓

(Reality)

PS

idea

(Pred)	Ho - True	Ha - True
Accept Ho	✓	× - TS
Reject Ho (Ac Ha)	× I	✓

CM = FP / FN

Survey (mistk)

DA mistake

MD

T I
T II

Error

Clear

(T-test)

CI =

{T-test is a parametric test that compares the means of the two samples. group
~~Ideally, a sample for t test should have less than 30 values. There are a few other assumptions that are taken before we can conduct a t-test.~~

T-test

1.4B

Pop

30

30

S1 S2

$\bar{S}_1$ $\bar{S}_2$

$\bar{x}_1$ 5.7 5.8 $\bar{x}_2$

5.8 > 5.7

$\bar{S}_1 > \bar{S}_2$
$\bar{S}_2 > \bar{S}_1$
$\bar{S}_1 \approx \bar{S}_2$
$\bar{S}_1 = \bar{S}_2$
$\bar{S}_1 \neq \bar{S}_2$

T - OST
IST
PTT

freq = 1,00,000 -

C MS

Cm

$\bar{S_1} \approx \bar{S_2}$ (Consistent)

Vaccine n-1

V1 → S1 → 1m

V2 → S2 → 28

(Parametric)

(Data) = (ND) → T ← PT

NPT

T = ND → PT ✗ } Plot

└ ✗ − NPT

T (−test)

One Sample T-test ✓	Ind-Sample T-test	Paired T-test

Drug (fever) → (15mg - Parach) → Standard

FMCG

50,000 → 30 → 14.98 ≠ 15 mg (Std)

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$\bar{x}$ = S mean

μ = Pop mean

s = Std

n = S1

P1 ↔ Ind ↔ P2

S1 ✗ S2

Ct

idea

USA ↔ India

30   30

−ve

30

0 − 100

(+) (+) (−)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

(+) (−)

$\bar{x}_1 > \bar{x}_2$ =

$\bar{x} > \bar{x}_2$ =

H0   H1

Man

T = ✗ (!)

@/α → t

S 2r

B A

dp = ✓

	Before	After
P1	97	92
P2	82	81
P3	83	82
P4	91	87

$$t = \frac{\bar{x}_d}{\frac{s}{\sqrt{n}}}$$

B	A	xd
92	81	11
81	72	9

$\frac{11+9}{2} = 10$ → $\bar{x}_d$

C

T = 2

X > 2   Clear

Clear

Pop (Pam) = (Ce)

P I => 2

$R_{TJ \cdot 2}$: ✓ [OST | IST | PTT]

done:

(The average salary of an Indian IT worker is approximately hundred dollars and this survey has been conducted in 2019.) If I take a new survey and 30 participants in 2025, we find out that their average salary is $140 with a standard deviation of $20. Can you calculate whether the salaries have significantly changed or not and can you do this at a 95% confidence interval, state alternate and null hypothesis for the given question?

100   90

(Data)

Pop Mean (μ) = 100 | Sample mean ($\bar{x}$) = 140 | CI = 95%

n = 30 | Std_dev = 20

(Sample Size) | (Standard dev)

EP → Clean

One

① Hyp

$\cancel{H_0 : Avg(\mu) = 100}$

H_a : Avg(μ) $\neq$ 100 (>, <) (Two-tailed T-test)

Is

Note :

① 2-tail test

$\alpha \longrightarrow \alpha/2 =$

$\longrightarrow \alpha/2 =$

② $\alpha - \alpha$

logic

② Significance level (α)

α = 100% − CI (%)

= 100 − 95

$\boxed{\alpha = 5\%}$ or $\boxed{\alpha = 0.05}$

③ degrees of freedom $\boxed{df = n-1}$

= 30 − 1

$\boxed{df = 29}$

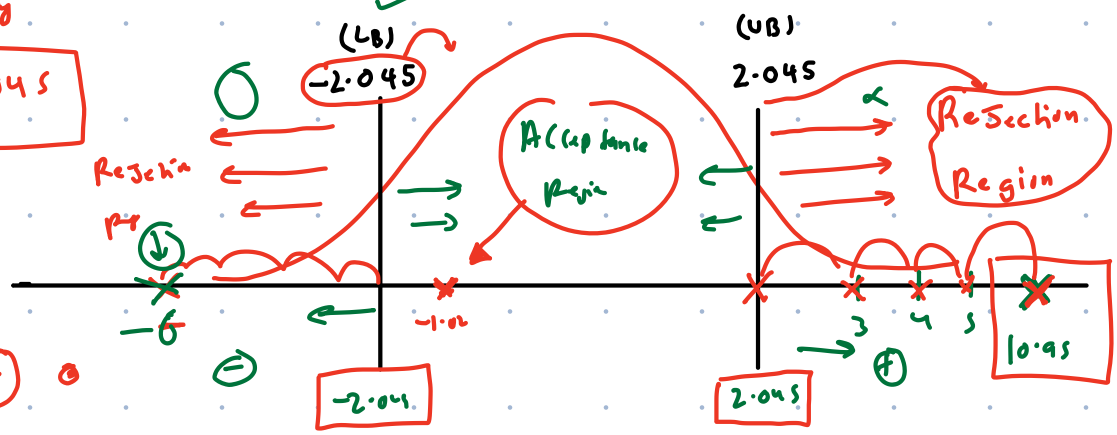
Pyth

log = T-table

idea

④ Decision Boundary

$\boxed{t\_critical\_value = \pm 2.045}$



Reject H_0

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551

α

=> P

S

S t_ Statistic

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{140 - 100}{\frac{20}{\sqrt{30}}} = +\boxed{10.95}$$

= − $\boxed{-1.02}$

−

−6