DAV-3

HYPOTHESIS TESTING

(Class starts @ 9:05 PM)

Lecture 2: Z-Test

#Agenda

1. Types of Error ✓
2. Tailed Test & HT Framework ✓
3. Recap of CLT ✓
4. One Sample Z-test ✓
5. Critical Value ✓

Recap

① Assumption : H_0 & H_a → Alternate Hypo

→ Null Hyp / Default Assump

② Evidence / Gathering data ?

③ p-value → $\boxed{P(\text{Data} \mid H_0 \rightarrow \text{True})}$ ⇒ $\dfrac{P(\text{Data} \cap H_0)}{P(H_0)}$

→ Should be quantifyable

pvalue ↓ → Data & H_0 (not in sync)

pvalue ↑ → Data & H_0 (in sync)

P-value < α → reject null

P-value > α → fail to reject H_0

4 Threshold (Signifinace level / appha / Margin of error)

($\alpha = 5\%$) $\rightarrow$ FP $\rightarrow$ significance level

CL (Confidence level) = $1 - \alpha$ (Sig level)

$= \boxed{1-\alpha}$

5 Type of Errors

- Type I error ✓ FP
- Type II error ✓ FN

p-value < α

Reject H_0

else

(Accept H_0 / fail to reject H_0)

Types of Error

Person

Judge says	I	G	
I	No Error	Type 2 error	(β)
G	Type I error	No Error	

(α) (Sig level)

"Confusion Matrix"

The Type of error
to focus
depends on the problem statement

Person

Doctor says	NC	C
NC	✓	FN Type2
C	FP	✓

Type I

Person (Actual)

Predicted (Judge Says)	I 0	G ✓ / 1 ✓
I 0	TN (0,0)	Type II (FN) (1,0) → Type II
G 1 ✓	(0,1) Type I, FP, α	TP (1,1)

Confusion Matrix

TN → True Neg

TP → True Post

Type I / FP → False Pos

Type II / FN → False Neg

H_0: Person is Innco

H_a: Person is guilty

Rejecting H_0 (calling them guitley) → "Positive" 1

Failing to Reject H_0 (calling them Innocent) → "Negative" 0

Actual Person

(Predicted) Judge says	I 0	G 1
I 0	TN (0,0)	FN (1,0)
G 1	FP (0,1)	TP (1,1)

Person is guilty → +ve 1 → 0

Person is innocent → $\boxed{\text{-ve}}$ 0 → 1

Actual

Predicted	I	G
I	TP (1,1)	FP (0,1)
G	FN (1,0)	TN (0,0)

H_0:- Person is innocent

H_a:- Person is guilty

Intruder Alert System

		Actual	
		NI	I
Alert Says	NI	TN	FN
	I	FP (Type I)	TP

H_0 : No Intruder

H_a : Intruder

Pos $\rightarrow$ Intruder detection (H_a)

Neg $\rightarrow$ No Intruder detection (H_0)

Intruder Detection

Actual

Predicted	I	NI
I	TP (1,1)	FP (0,1)
NI	FN (1,0)	TN (0,0)

H_0 :- No Intruder

H_a :- Intruder

Intruder → +ve 1

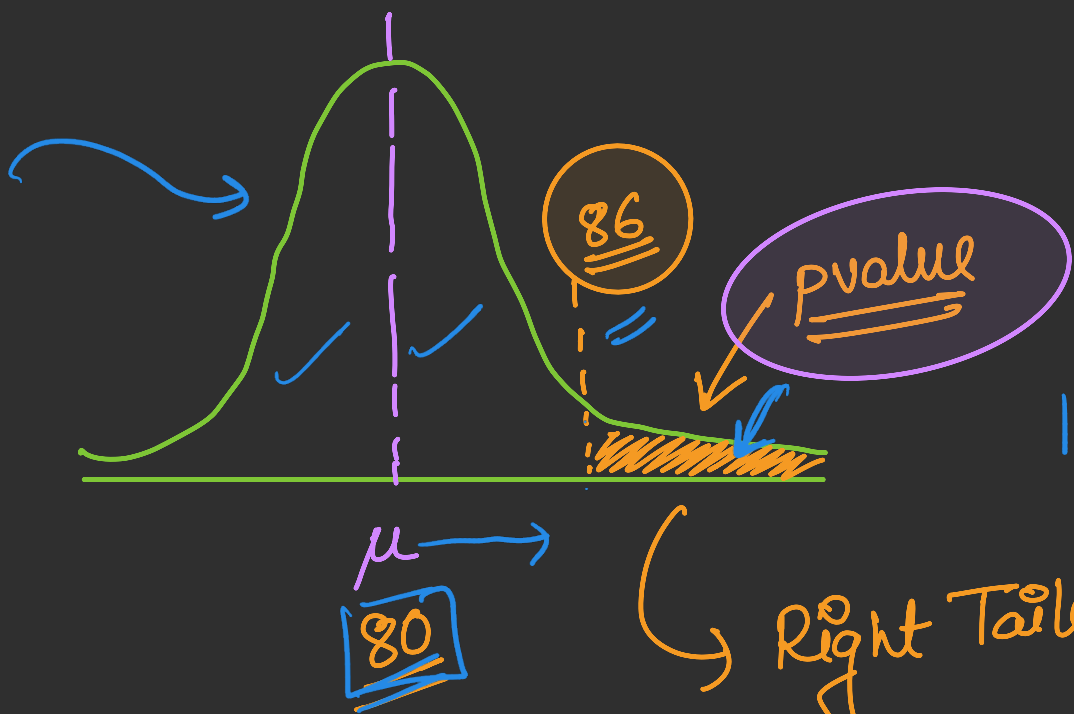
No Intruder → -ve 0

Tailed Test

There are Two Types of Tailed Test

- One Tailed Test
  - Left Tailed Test ✓
  - Right Tailed Test ✓
- Two Tailed Test

For Example : Burger Example



$H_0 : (\mu = 80g)$

$H_a : \mu > 80g$

1 - norm.cdf

Right Tailed Test

Mc donald's says burger weight is 80g

FSB says that weight is > 80 (86 g)

Model deployment Example

ML_1 Vs ML_2

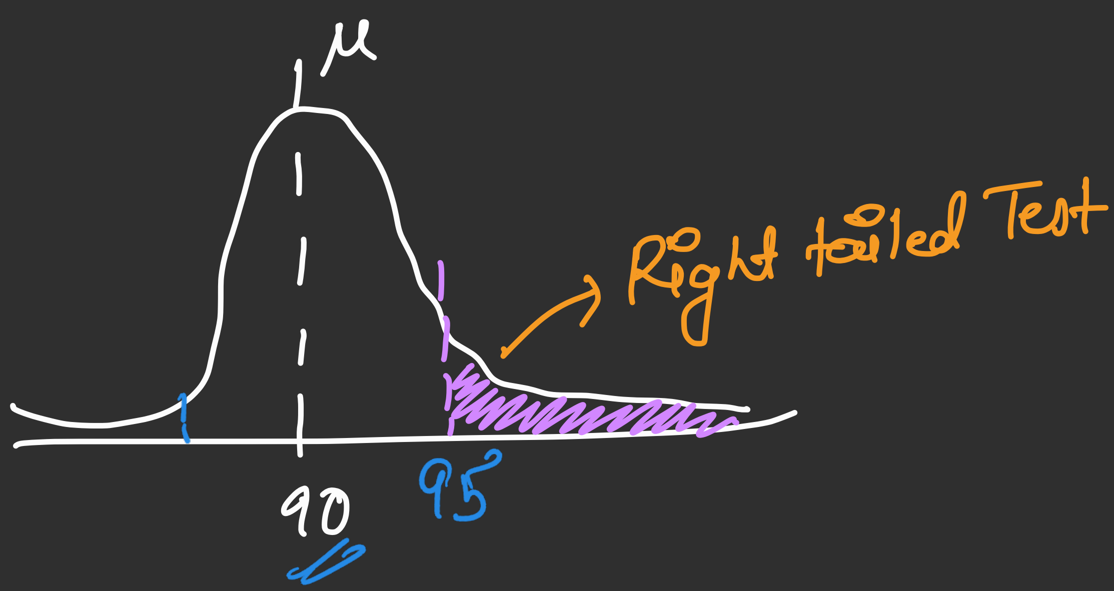
$H_0 \Rightarrow$ Average Accuracy is 90% [$\mu = 90\%$]

$H_a \Rightarrow$ Average Accuracy greater than 90% [$\mu > 90\%$]

($\mu < 90\%$)

Left Tailed Tailed

$\mu < 90\%$



1) $H_0: \mu = 160$

$H_a: \boxed{\mu > 160}$

Right tailed

2) $H_0: \mu = 183$

$H_a: \mu < 183$

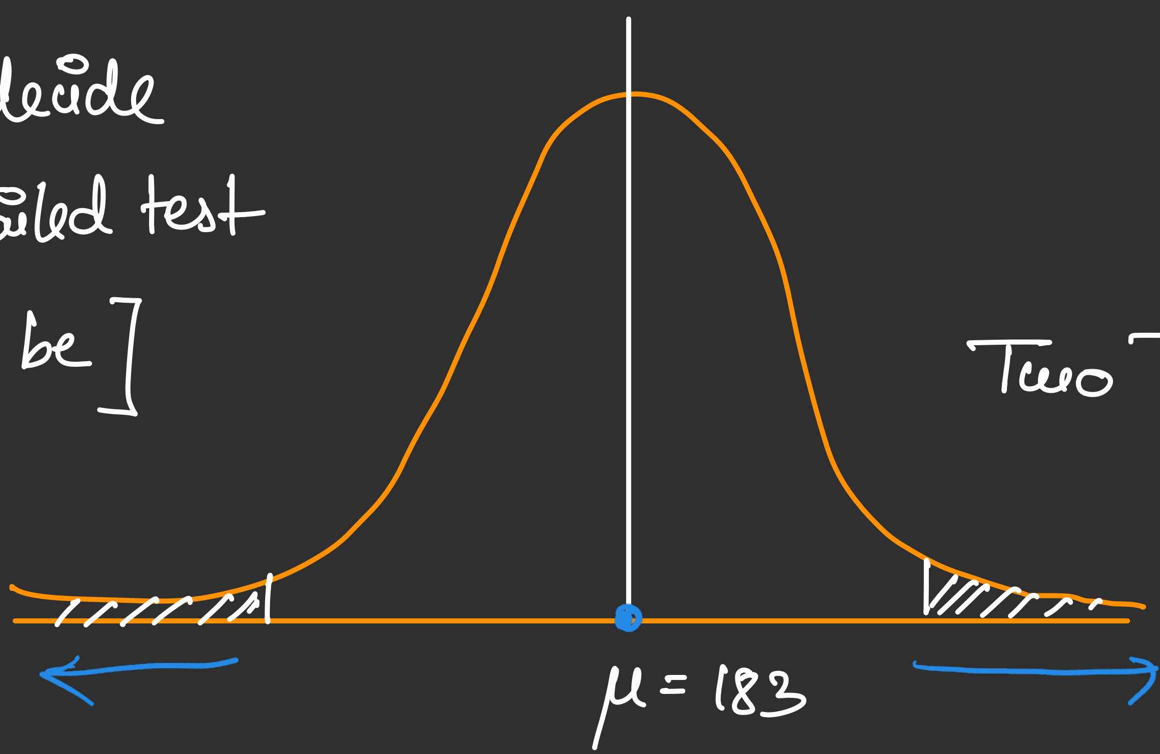
Left Tailed

3) $H_0: \mu = 183$

$H_a: \mu \neq 183$

Two Tailed Test

[Ha helps us decide what type of tailed test it is going to be]



Two Tailed Test

$H_0 \Rightarrow \mu = 64$

$H_a \Rightarrow \mu > 64$

Right Tailed Test

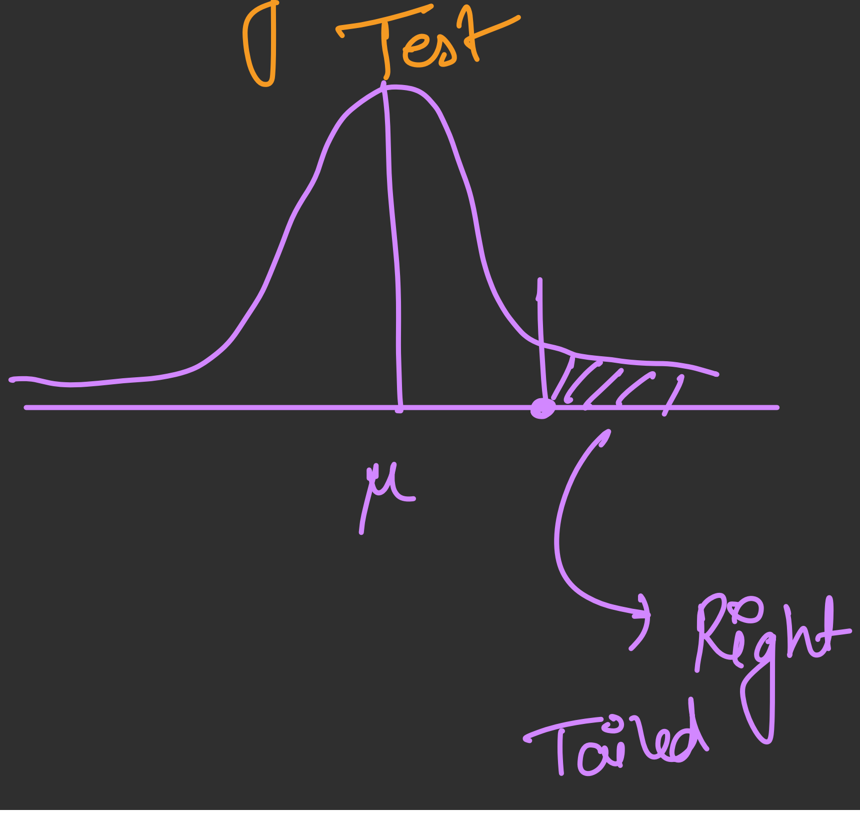
$\mu = 100$

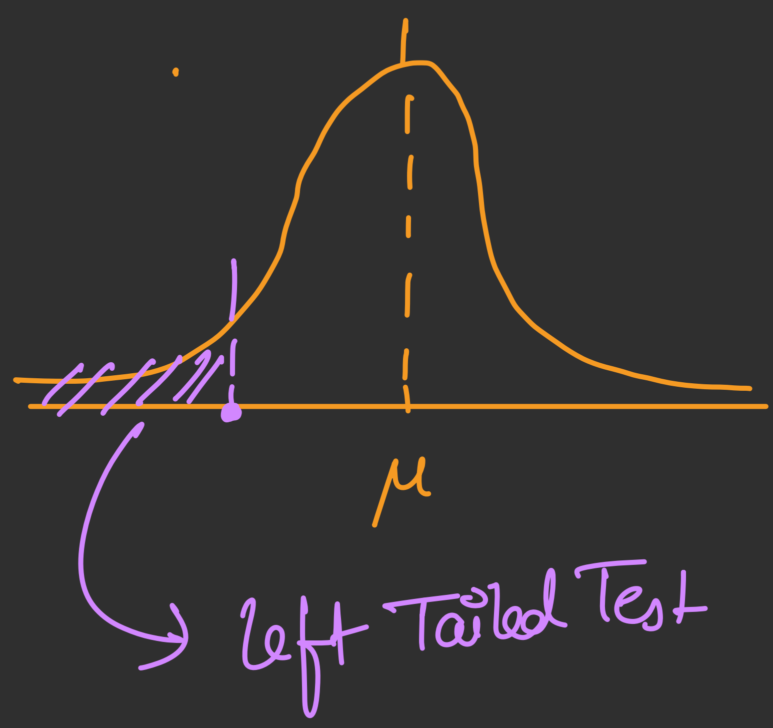
$\mu \neq 100$

$H_0 \Rightarrow \mu = 86$

$H_a \Rightarrow \mu < 86$

Left Tailed Test





Height of people in India

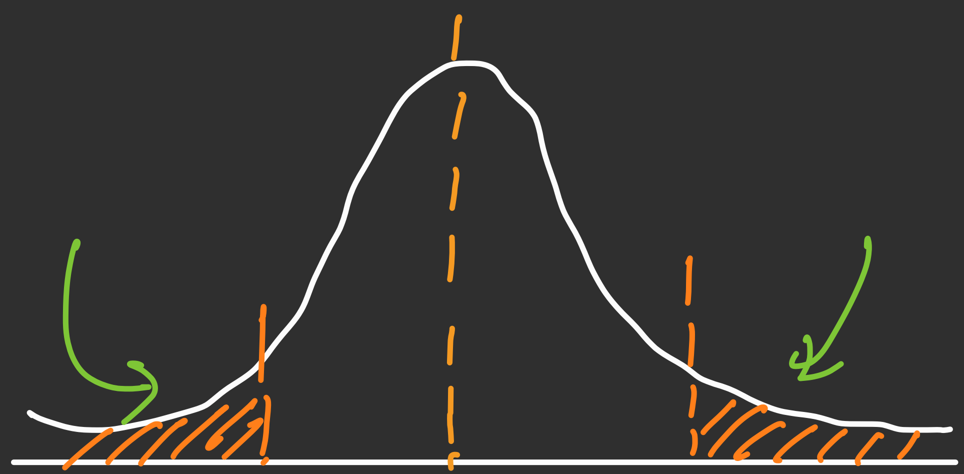
$$\begin{cases} H_0 : \mu = 65 \\ H_a : \mu \neq 65 \end{cases}$$

"Two Tailed Test"

$\mu = 65$

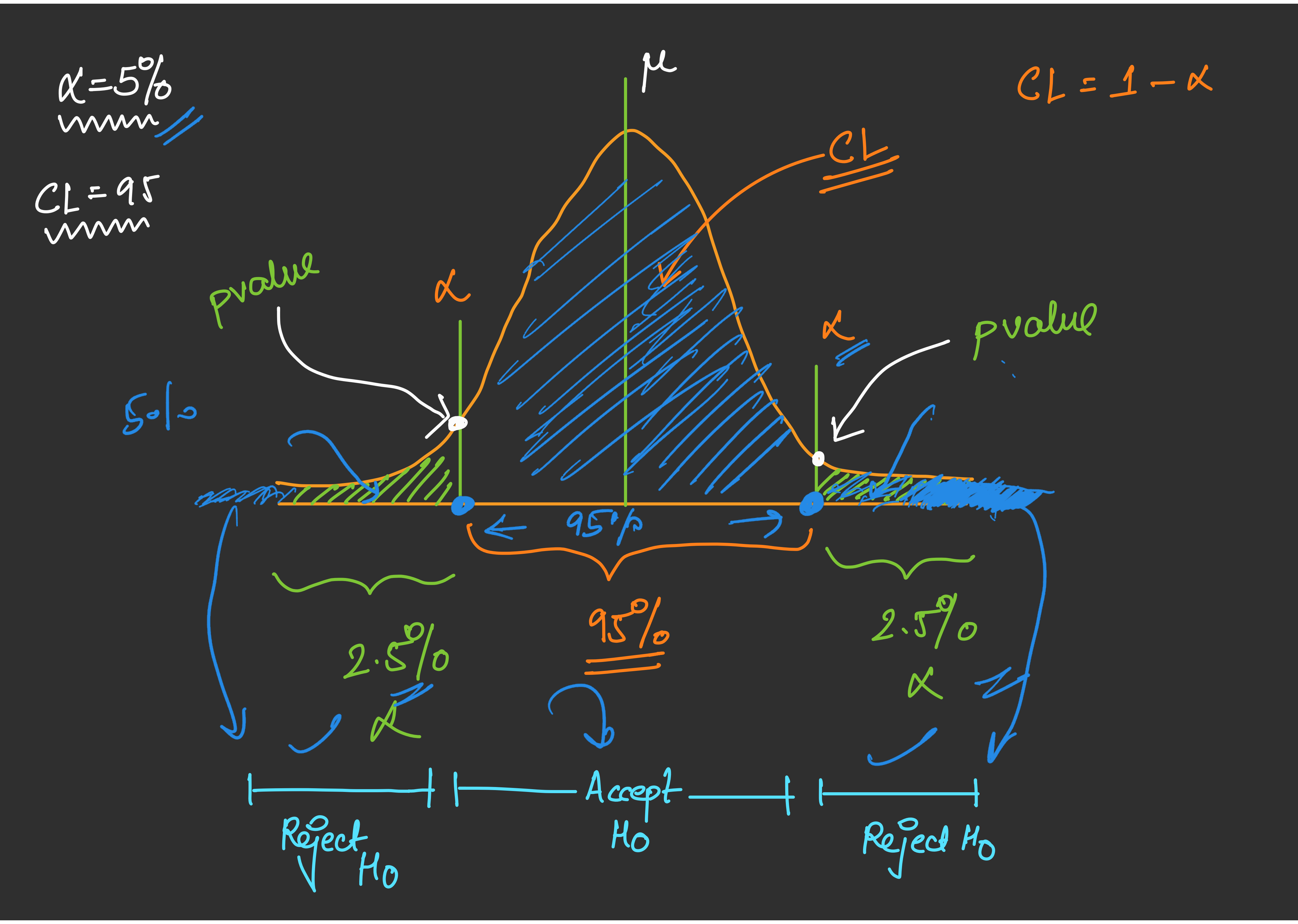
$65 < \mu > 65$

$(\mu \neq 65)$

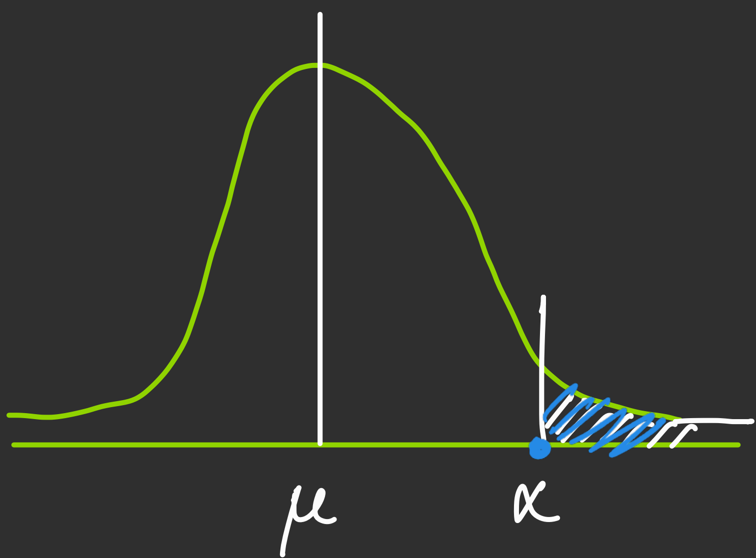


$(\mu = 65)$

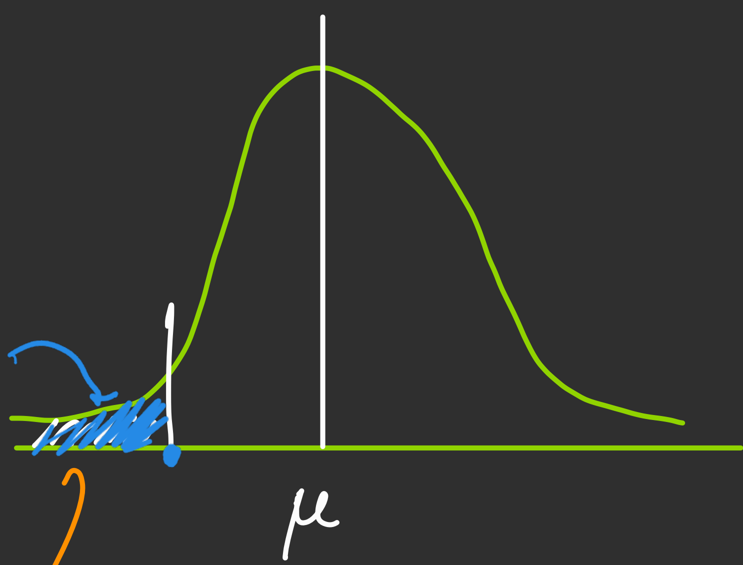
(Two Tailed Test)



α=5%
CL=95
μ
CL=1-α
CL
pvalue
α
α
pvalue
5%
95%
95%
2.5%
2.5%
α
α
Reject H0
Accept H0
Reject H0



$$\underbrace{1-\text{norm.cdf}(\ )}_{\text{pvalue}}$$



$$\underbrace{\text{norm.cdf}(\ )}_{\text{pvalue}}$$

← left to right

Coin Toss Example

1. H_0 : Coin is fair
   H_a : Coin is biased

   Assump

   $P(H) = 0.5$

   H_0 : Coin is fair
   H_a : Coin is biased towards Head

   $P(H) > 0.5$

2. Distribution → (H/T) [Binomial distribution]

3. Type of Tailed Test (LT/RT/TT)

   (Right Tailed Test)

H_0: Coin is fair $(P(H) = 0.5)$

H_a: Coin is biased $(P(H) \neq 0.5)$

1 Two tailed Test

2 Right Tailed Test

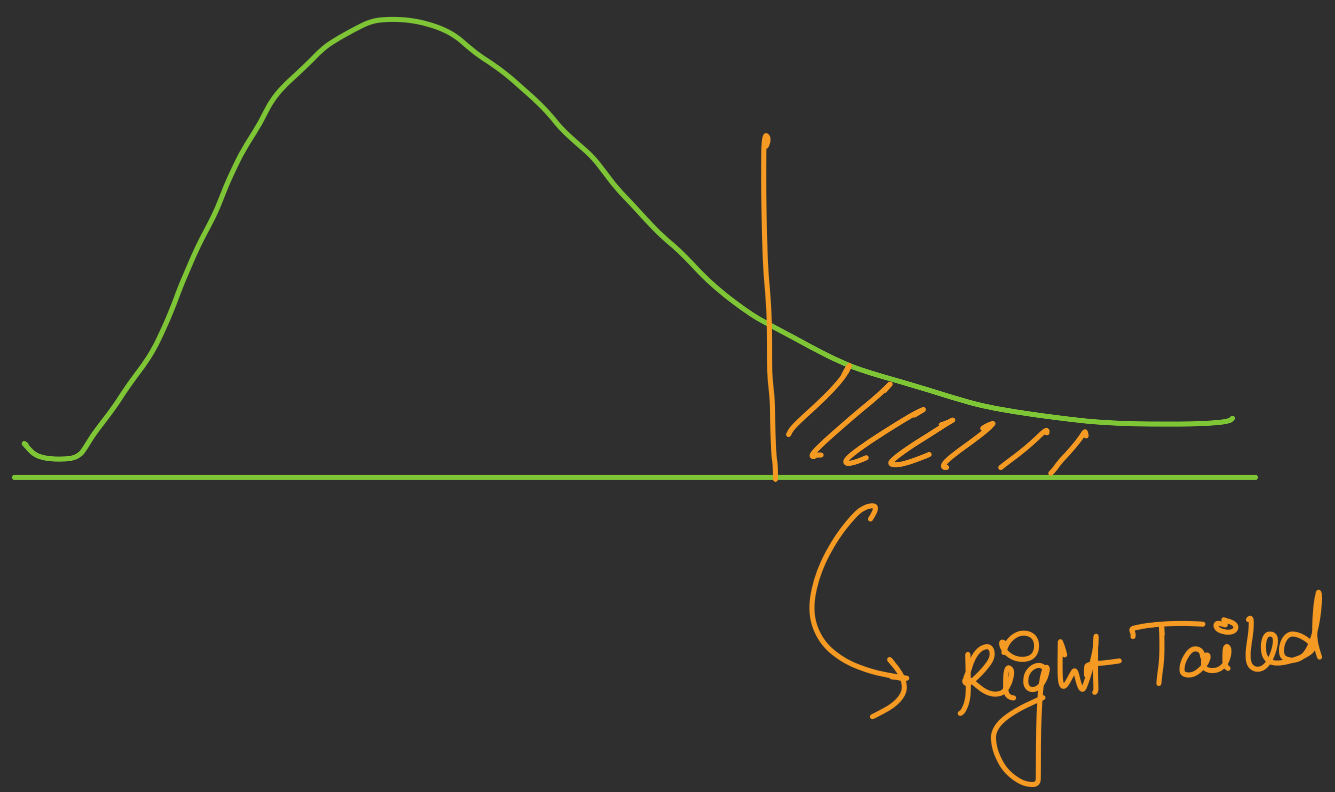
H_0: Coin is fair $(P(H) = 0.5)$

H_a: Coin is biased towards heads $(P(H) > 0.5)$

H_0 : Coin is fair

H_a : Coin is biased Towards Head $(P(H) > 0.5)$

Q) How To decide on the Type of tailed Test?



p-value

$P(\text{No of heads} \rightarrow 7/8/9/10 \mid \underbrace{\text{Fair Coin}}_{H_0})$

10 times → Coin flipped
7 times → Head

$$\boxed{\text{p-value} = 1 - \text{binom.cdf}(k=6,\ n=10,\ p=0.5)}$$

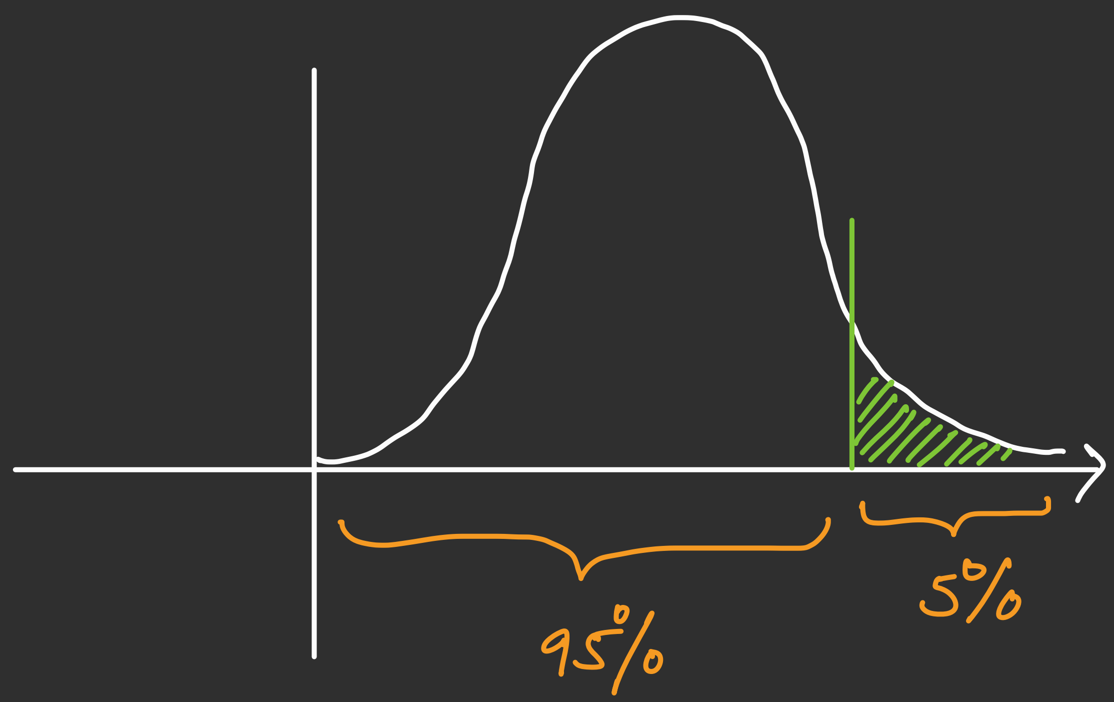
4 p value comparison with threshold

Sig Level / alpha / Margin of Error

($\alpha = 5\%$ or 0.05)

p-value < α

Reject H_0



HT framework

We start any Hypothesis-testing problem with 2 things:

- Assumption
- Data

There exists a framework to compute a quantifiable metric (pvalue) that will help us decide if we should accept or reject our null hypothesis.

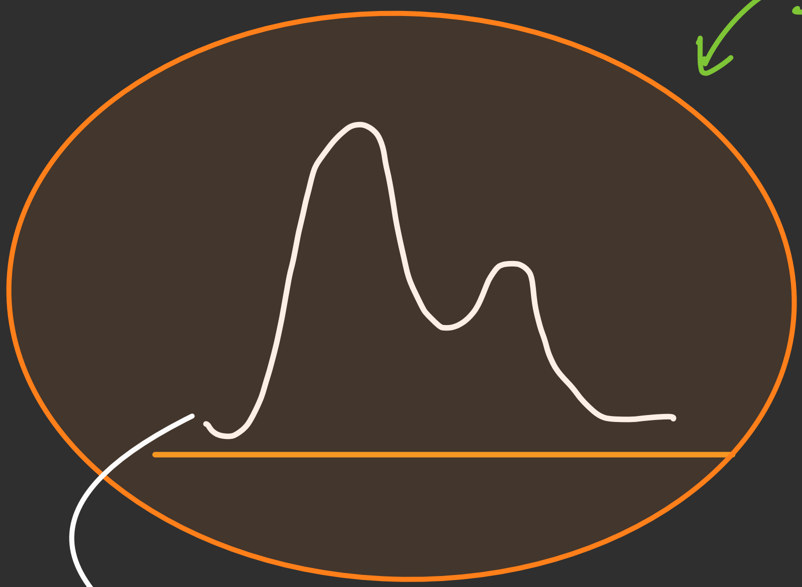
Let's summarise it into steps:-

1. Setup Null and Alternate Hypothesis
2. Choose the **test statistic** → Distribution
3. Select the Left vs Right vs Two-Tailed test, as per the hypothesis
4. Compute the P-Value
5. Compare the P-Value to the Significance Level (α) and Fail to reject/reject the Null Hypothesis accordingly.
    - Another term closely related to Significance Level is **Confidence Level**
    - If we're using $\alpha = 0.05$, this means that 5% significance

$(CL = 1 - \alpha)$

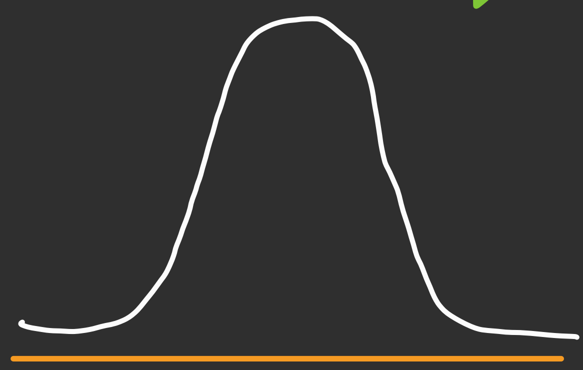
Recap of Central limit theorem

sample of this dist



Normal dist ? X

means plot



Normal

[means of sample mean $\simeq$ pop^n mean]

Conditions for CLT

1. Sample size ≥ 30 (Emperical proven)

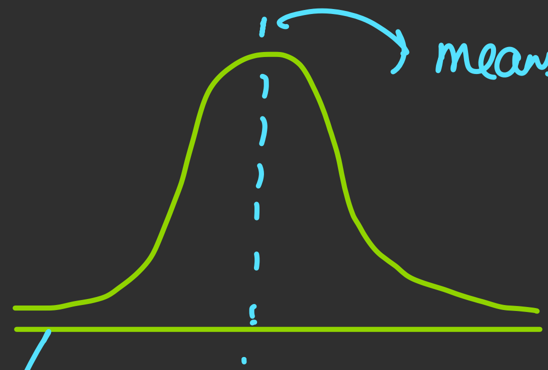
2. $\underbrace{\text{Population parameters}}_{(\mu, \sigma)}$ needs to be finite

Quiz 1

$$\begin{cases} \mu = 65 \\ \sigma = 2.5 \end{cases}$$

50 people

sample



means of the sample mean

($\simeq$ pop mea)

(~ 65)

sampling dist (m)

(Normal / Gaussian)

Expected value (mean)

sample mean $\Rightarrow$ $\underline{\underline{65}}$

$$\left(\frac{2.5}{\sqrt{50}}\right)$$

CLT $\longrightarrow$ $N \sim \left(\mu, \sigma/\sqrt{n}\right)$

$\sigma/\sqrt{n}$ $\rightarrow$ Standard Error

$\sigma \rightarrow$ pop^n std

$n \rightarrow$ sample siz

One Sample Z-test

"Marketing Case Study" → Business Owner → 2000 Retail Stores

→ Shampoo

CEO

Weekly avg sales

$\mu = 1800$

$\sigma = 100$

→ No of bottle sold

Goal: To increase the revenue

To do a marketing campaign

Hire a marketing Agency

2000 Stores

Firm A

$\boxed{\text{50 stores}}$

$\mu = 1850$

Firm B

5 stores

$\mu = 1900$

Firm A

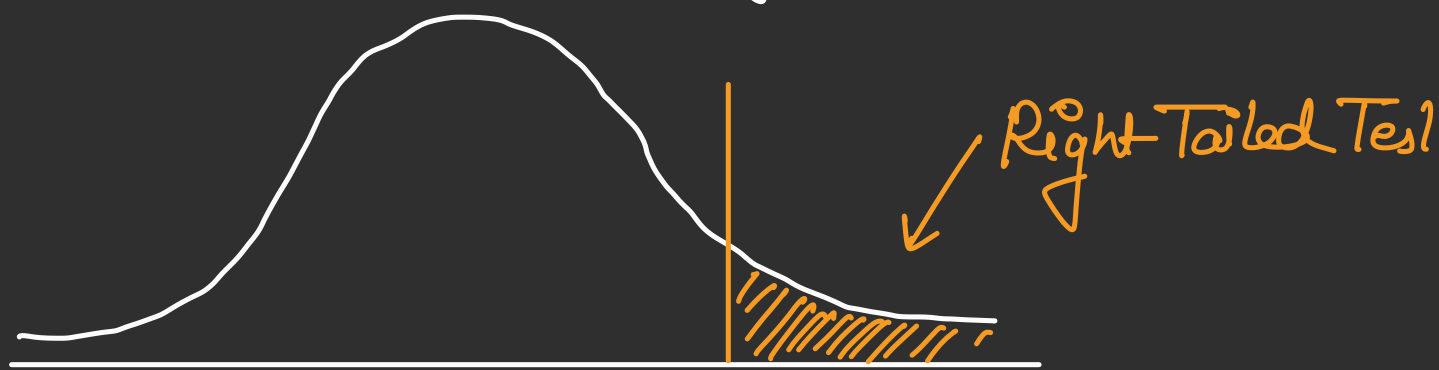
Sample = 50

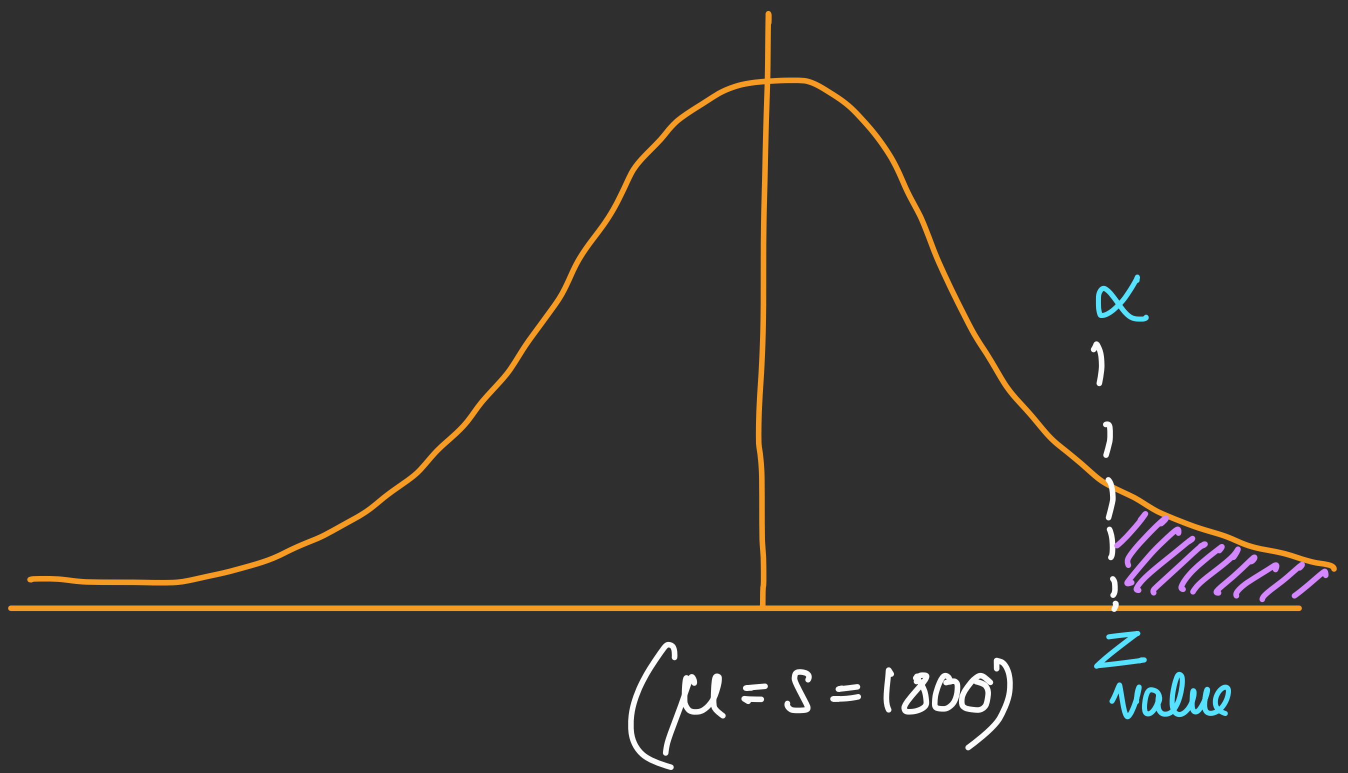
① $H_0 \Rightarrow \mu = 1800$ (No impact of Marketing Camp)

$H_a \Rightarrow \mu > 1800$ (Marketing camp worked)

② Distribution ⟶ "Normal distribution"

③ Tailed Test ⟶ "Right Tailed Test"





p-value ⇒ 1 − normcdf(z)

$$\underset{\sim\sim\sim}{\text{z score}} = \frac{x-\mu}{\left(\sigma/\sqrt{n}\right)} \Rightarrow \frac{1850-1800}{\left(\frac{100}{\sqrt{50}}\right)} =$$

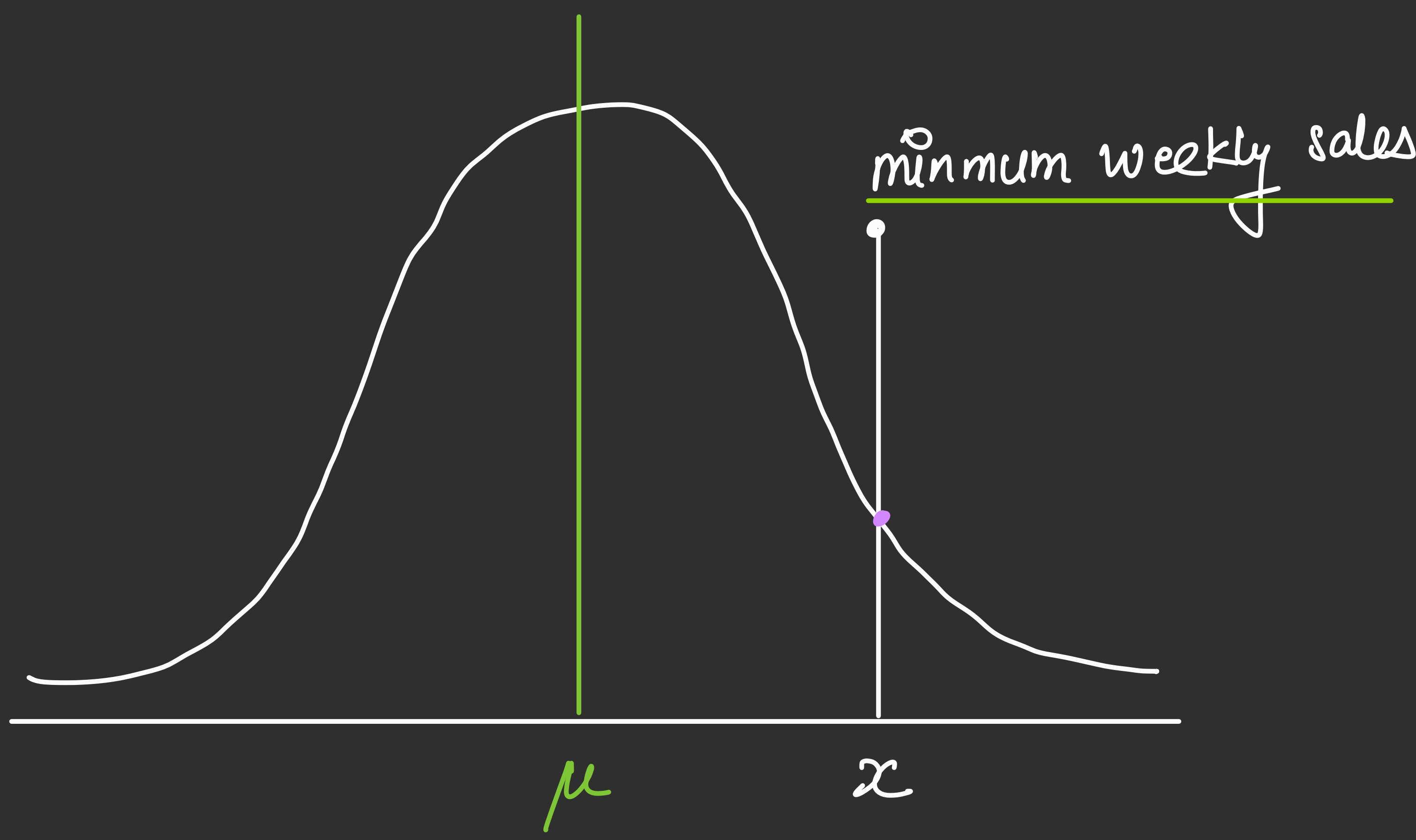
$$Z_{score} = \frac{x-\mu}{\sigma}$$

Sampling Stats

$$Z_{score} = \frac{x-\mu}{\left(\sigma/\sqrt{n}\right)}$$

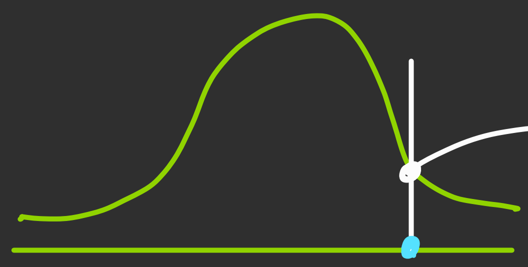
SE

Q) What is the minimum weekly average sales below which I can say the MA is not working?



p-value = 1 − norm.cdf(z)

$\alpha = 0.01$



(pval = α)

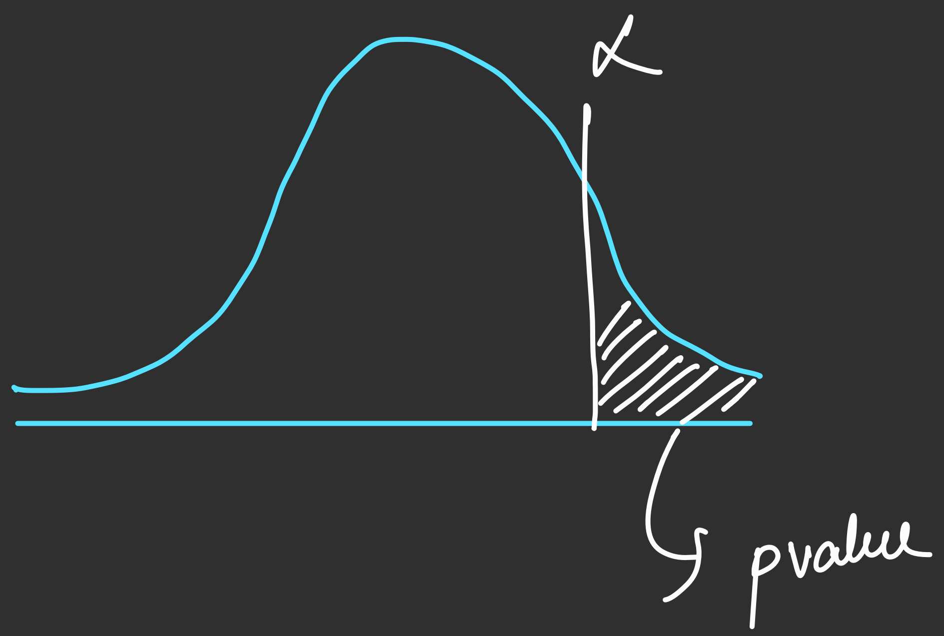
(p-value = α) at that min pt

→ 1 − norm.cdf(z) = 0.01

→ (z = norm.ppf(0.99))

→ z critical

$$z_{crit} = \frac{x - \mu}{(\sigma/\sqrt{n})}$$



$(pval = \alpha)$

if $pval < \alpha$

else $pval > \alpha$

1. p value $< \alpha$ (Reject / Accept H_0)
2. Critical value ($x_{critical}$) $\longrightarrow$ (Reject / Accept H_0)

1) HT Framework

2) Formulating H_0 & H_a