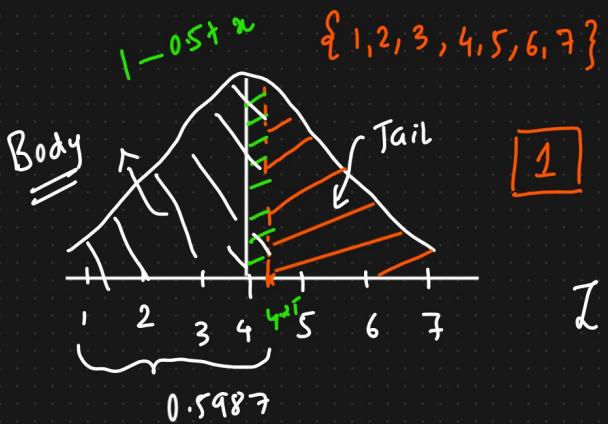


## Day 2 - Stats

$$\textcircled{1} \quad Z\text{-Score} = \frac{x_i - \mu}{\sigma}$$

Stats Interview Question



How many standard deviation

4.25 fall from the mean??

$$Z\text{-Score} = \frac{x_i - \mu}{\sigma} = \frac{4.25 - 4}{1} = 0.25$$

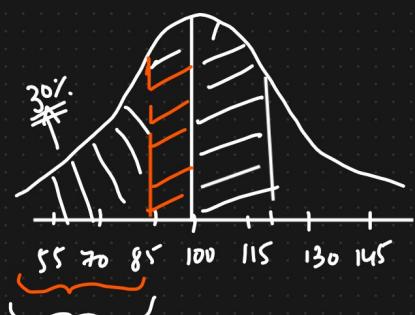
Question : What percentage of scores fall above 4.25?

$$1 - 0.59871 = 0.4013 \Rightarrow 40.13\%$$

2 In India the average IQ is 100, with a standard deviation of 15.

What is the percentage of the population would you expect to have an IQ lower than 85?

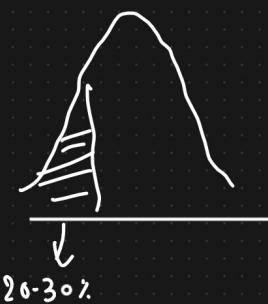
Ans)



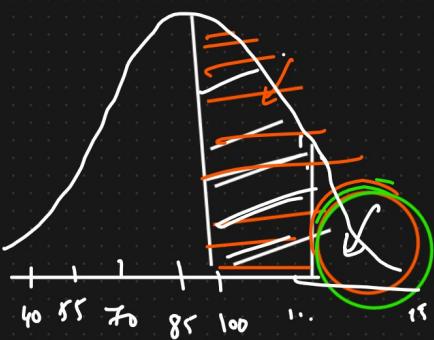
$$Z\text{-Score} = \frac{85 - 100}{15} = \frac{-15}{15} = \boxed{-1}$$

① Area under this curve

$$0.5 - 0.15866 = 0.34143 \Rightarrow \boxed{34.14\%}$$



$$\{ \text{Growth} = 100 \text{ less than } 125 \}$$



$$Z\text{score} = \frac{125 - 100}{15} = \frac{25}{15} = 1.667$$

$$\text{Ans} = 0.4515 \Rightarrow 45.15\%$$

1.667

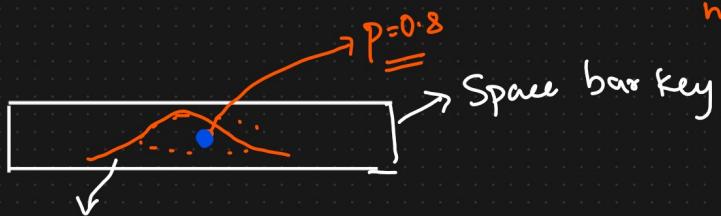
$$\underline{0.5 - 0.4515 = 0.0485} \Rightarrow 4.8\%$$



Left Z-table

P value, Hypothesis Testing, Confidence Interval

Out of all 100 touches, the no. of touches is 80



$$P=0.4$$

Out of all 100 touches, the no. of times 40 times.

Hypothesis Testing, C.I., Significance value Together Fair Coin

Coin  $\rightarrow$  Test whether the coin is a fair coin or not by performing 100 tosses

$$\begin{array}{c} P(H) = 0.5 \\ = \\ P(T) = 0.5 \end{array}$$

## Hypothesis Testing

Criminal is  $\rightarrow$  Court

SMOLAY

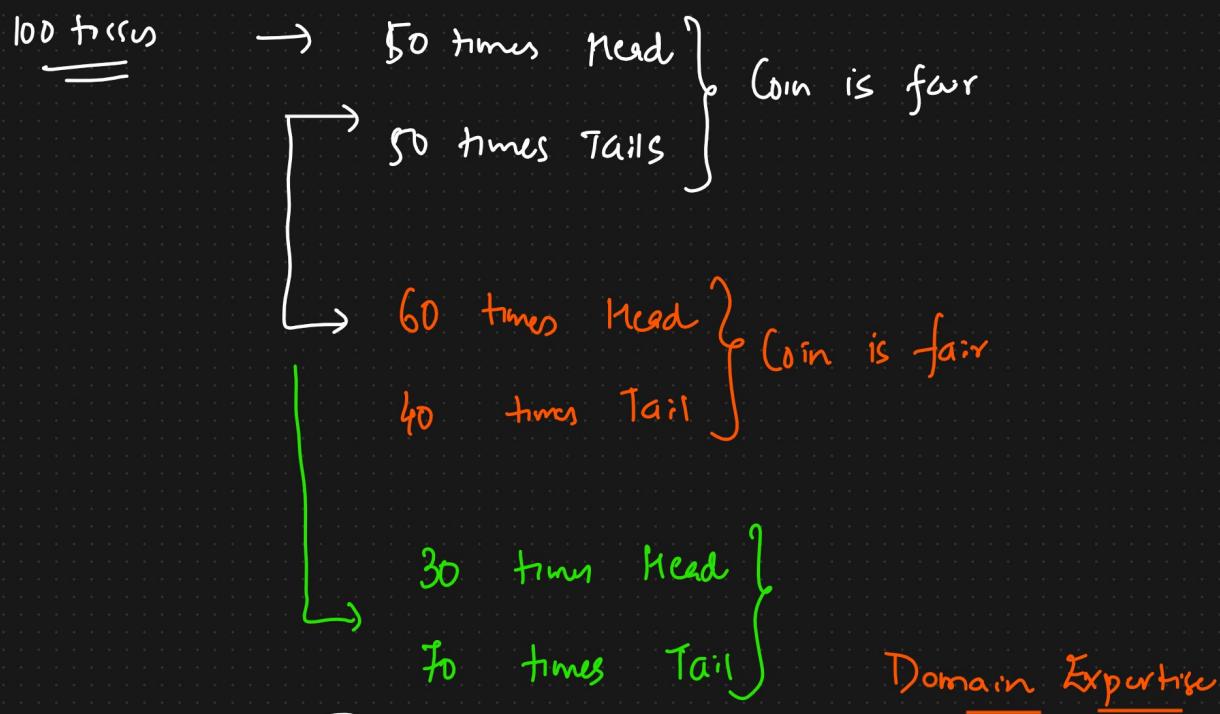
$$P(H) = 100\% \quad P(T) = 0\%$$

① Null Hypothesis — Coin is fair  $\rightarrow (H_0)$

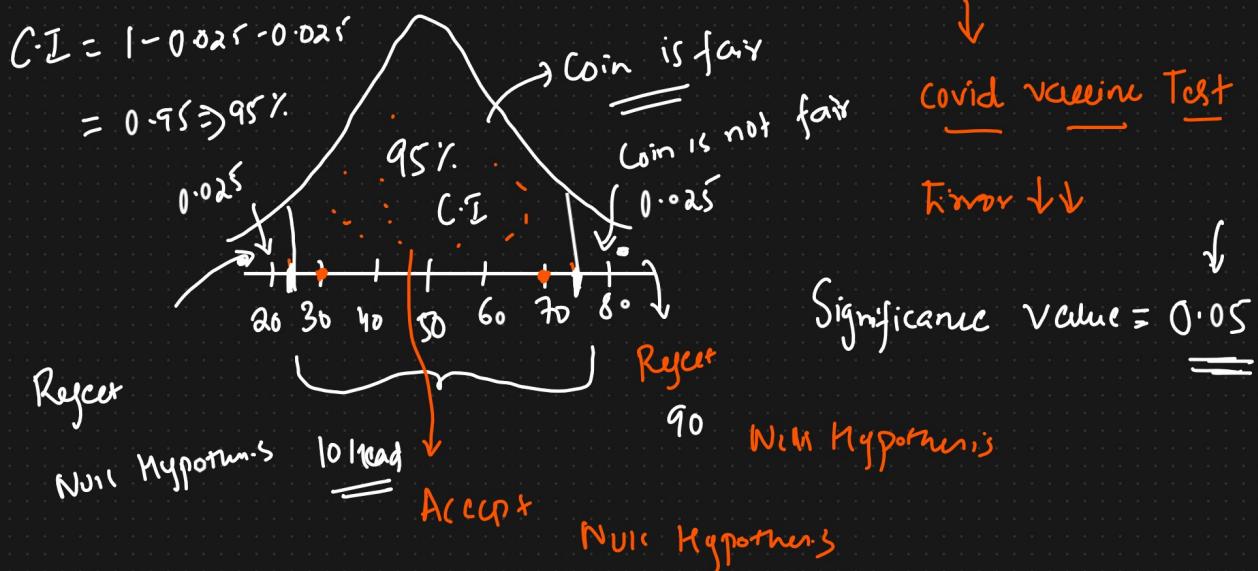
② Alternative Hypothesis — Coin is not fair  $\rightarrow (H_1)$

③ Experiments

④ Reject or Accept the Null Hypothesis



Confidence Interval, Significance Values

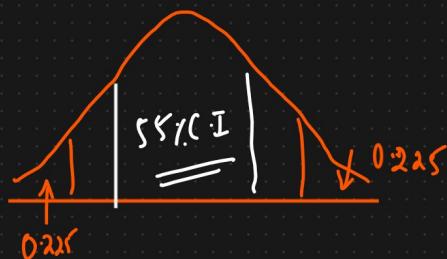


$$\lambda = 0.45$$

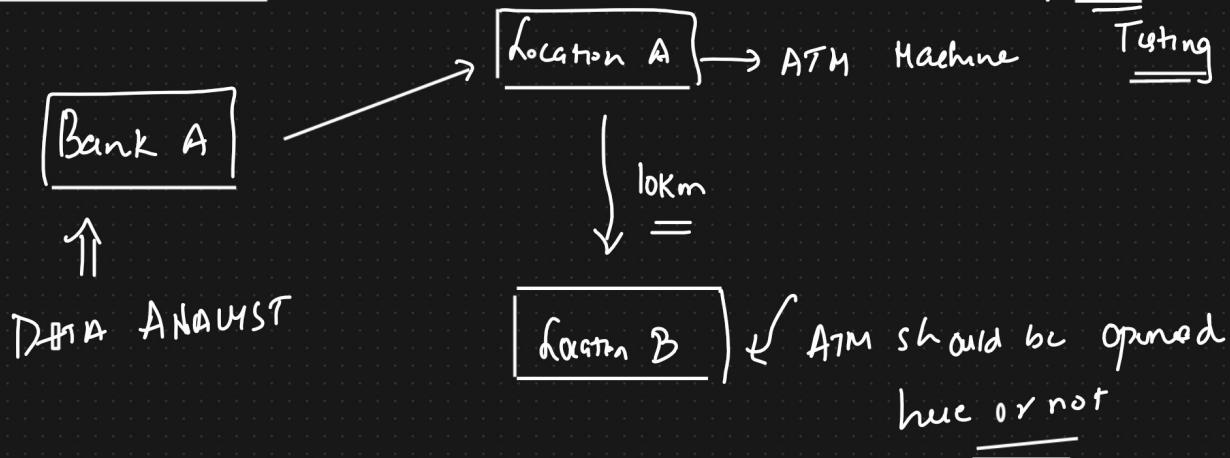
Medical

$f \uparrow \uparrow$

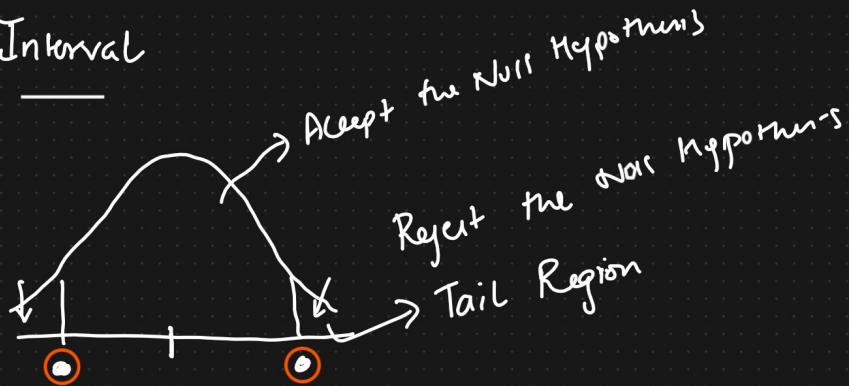
$$\frac{0.45}{2} = 0.225$$



## Real World Project

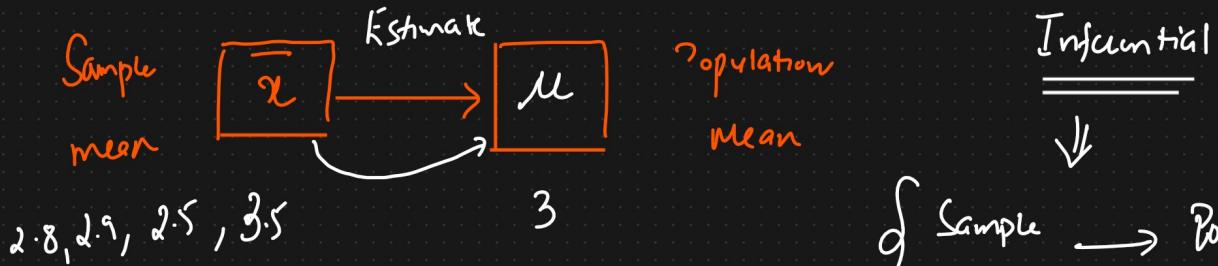


## ① Confidence Interval



## Point Estimate

{ The value of any statistic that estimates the value of a parameter is called Point Estimate.

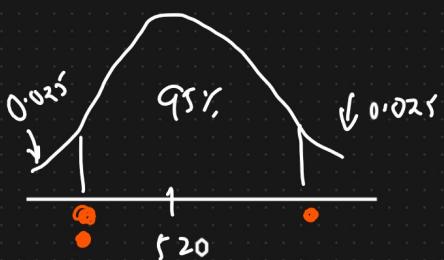


## Confidence Interval

t test Point Estimate  $\pm$  Margin of Error  $\Rightarrow$  Population.

- Q) On the quant test of CAT Exam, the population standard deviation is known to be 100. A sample of 25 test takers has a mean of 520. Construct 95% CI about the mean?

$$\text{Ans) } \sigma = 100 \quad n = 25 \quad \bar{x} = 520 \quad (\cdot I = 95\%) \quad \alpha = 0.05$$



① Population std is given {Z score}  $\rightarrow$  Z-table

Point Estimate  $\pm$  Margin of Error  $\Rightarrow$  C.I. =

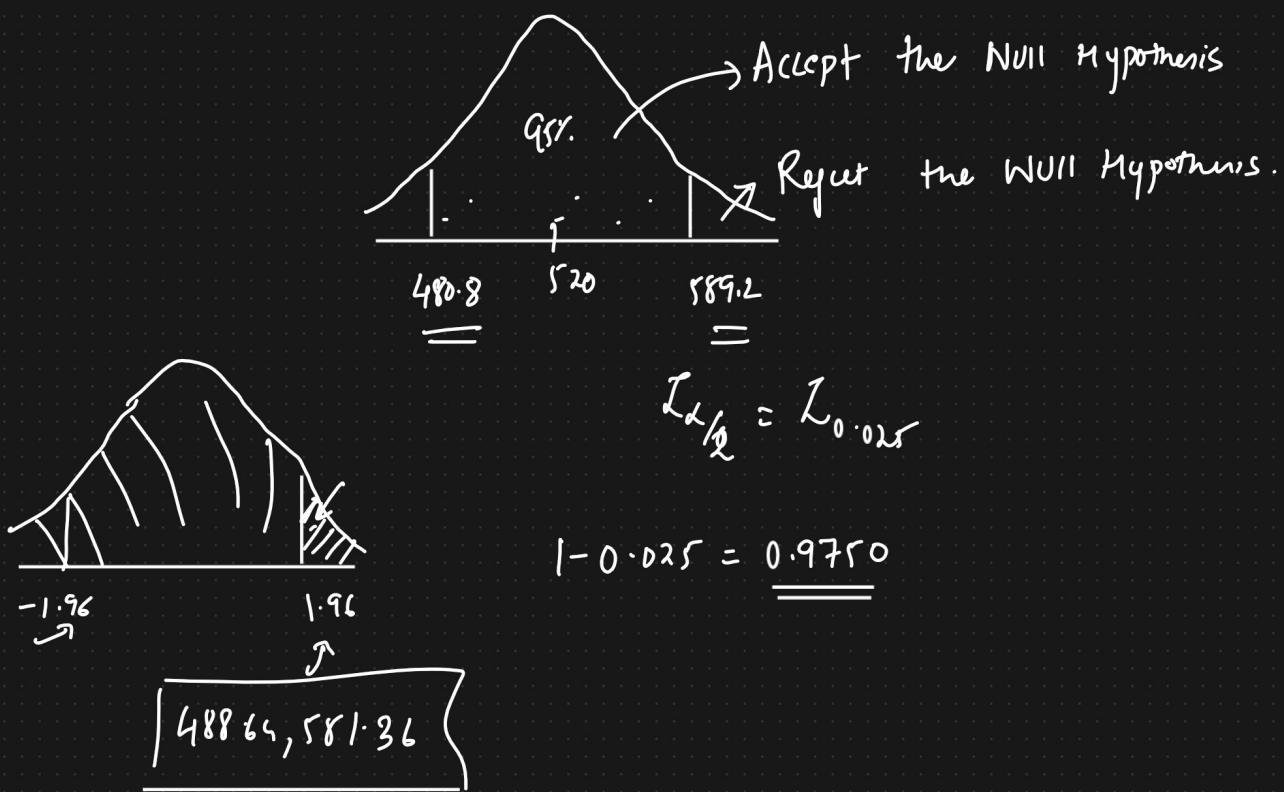
$$\bar{x} \pm Z_{\alpha/2} \left[ \frac{\sigma}{\sqrt{n}} \right] \rightarrow \text{Standard Error}$$

$$\text{Lower fence C.I.} = \bar{x} - Z_{\alpha/2} \left[ \frac{\sigma}{\sqrt{n}} \right] \Rightarrow Z_{0.05} = 1.96$$

$$\text{Higher fence C.I.} = \bar{x} + Z_{\alpha/2} \left[ \frac{\sigma}{\sqrt{n}} \right]$$

$$\text{Lower fence} = 520 - (1.96) \times \frac{100}{\sqrt{25}} = 520 - (1.96) \times 20 = 480.8$$

$$\text{Higher fence} = 520 + (1.96) \times 20 = 559.2$$



- ④ On the quant test of CAT exam, a sample of 25 test-takers has a mean of 520 with a sample standard deviation of 80. Construct 95% C.I about the mean? 2

$$\text{Ans) } \bar{x} = 520 \quad S = 80 \quad f = 0.05 \quad n = 25$$

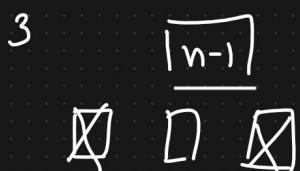
$t$ -test  $\Rightarrow t$  - table { Because population  $Sd$  is not given }

$$\bar{x} \pm t_{\alpha/2} \left( \frac{S}{\sqrt{n}} \right) \rightarrow \text{Standard Error}$$

$$t_{0.025}$$

$t$ -test

$$\textcircled{1} \text{ Degree of freedom} = n-1 = 25-1 = 24$$



3 people

$$\bar{x} \pm 2.064 \left( \frac{80}{5} \right) \Rightarrow 486.976 \leftrightarrow 553.024$$

- (f) Type 1 and Type 2 Error.
- (g) One Tailed vs 2 Tailed Test

## Type 1 and Type 2 Error

### Reality Check

$H_0 \Rightarrow$  Coin is Fair

① Null Hypothesis is True or Null

$H_1 \Rightarrow$  Coin is not Fair

Hypothesis is False

#### Outcome 1:

#### Decision of Experiments?

We reject the Null ✓ Null Hypothesis is True or False.

in reality if it is false → Yes ✓

Null Hypothesis ✓

$H_0 \rightarrow$  The Criminal is not guilty

$H_1 \rightarrow$  " " is guilty

#### Outcome 2:

We reject the Null Hypothesis

when in reality it is true  $\Rightarrow$  No  $\Rightarrow$  Type 1 Error X

#### Outcome 3

We accept the Null Hypothesis,  $\Rightarrow$  Type 2 Error X

When in reality it is false

#### Confusion Matrix

Outcome 4: We accept the Null Hypothesis

when in reality it is True ✓

$\begin{bmatrix} \downarrow \\ \text{Cancer} \\ \text{True} \end{bmatrix} \rightarrow \underline{\text{Not Cancer}}$

{  $\rightarrow$  Stock market is going to crash }

## ② 1 Tail and 2 Tail Test

Eg: College is Karnataka has an 85% placement rate. A new college was recently opened and it was found that a sample of 150 students had a placement rate of 88%. With a standard deviation of 4%. Does this college has a different placement rate?

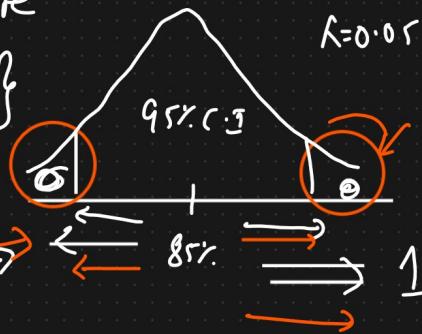
$$\alpha = 0.05 \Rightarrow 95\% \text{ C.I} \rightarrow 85\%$$

of placement rate

less than 85% }



1 tail



{ Placement rate greater than 85% }

2 tail Test

1 tail Test

Saturday

10 min probability

Sunday

① Z test Hypothesis Testing

EDA  $\rightarrow$  3-4 projects

② J Test Hypothesis Testing

FE  $\rightarrow$  \_\_\_\_\_

③ Significance value of P value.

Machine Learning

④ ANOVA TEST

⑤ CHI SQUARE TEST

⑥ Practical