Session-44

- In population
  - mean, variance, std etc } → parameter

- In Sample
  - mean, var, std etc } → statistics

- Inferential statistics -
  - draw conclusion about a population based on data collected from sample

- Point Estimate -
  - a single numerical value used to estimate an unknown population parameter based on sample data
  - Sample mean, variance etc.
  - point estimate is not reliable that's why we use range

- Confidence interval (CI)
  - Range of value which is likely to contain the true population parameter

- confidence level
  - tells how confidence we are that the interval contains true value

CI = Point estimate ± margin of Error

Page No.	
Date	

- Calculating CI –
  1) Z procedure (popln std is known)
  2) t Procedure (" " not ")

- CI is created for Parameter not for Statistics

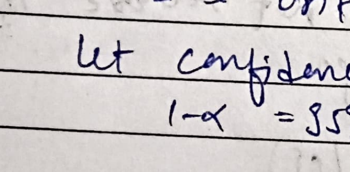
- Z procedure (σ known) :
  Assumptn –
  - Sample should be random
  - σ should be known
  - distributn of populatn is Normal
    - if not normal then apply CLT

$$CI = \bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$\bar{X}$ → Point estimate (By CLT)
σ → Std. deviation
n → Sample size
$(1-\alpha)$ → confidence level
Z → critical value

Let confidence level we want is 95%
$1-\alpha = 95\%$ & $\frac{\alpha}{2} = \frac{0.05}{2} = 0.025\%$

So $CI = \bar{X} + Z_{0.025} \frac{\sigma}{\sqrt{n}}$



Z
0.95
0.025
0.975

using z table we get 1.96

Page No.
Date

So $CI = \bar{X} \pm 1.96 \frac{\sigma}{\sqrt{n}}$

- Similarly if we want 75% accuracy we can use Z table

· confidence level -
suppose 75K pop.
95% - [18, 42]
random sample (50) & 100 times

Means if we take sample of 50 students and calculate mean. And do this 100 times then there is 95% chance that mean will lie b/w the interval.

· If we increase confidence level then the width (range) also increases.

· Margin of Error $= Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$
↳ max. expected diff. b/w point estimate true population value

· Factors affecting it -
i) Sample size → large ↑ then small MOE
ii) confidence level → large ↑ then large MOE
iii) σ → large ↑ more MOE

- t- procedure (σ not known)

Assumption –

- Sampling must be random
- Approx Normal dist. (can apply CLT)
  - ↳ If sample size is small & it is normal then we can also apply t-procedure
- Observation should be indep.

$$CI = \bar{x} + Z_{\alpha/2} \frac{S}{\sqrt{n}}$$ $S \rightarrow$ sample std. deviation

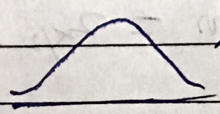
but if we do this then in standardization

$$z' = \frac{\bar{x}' - \mu}{S/\sqrt{n}}$$

variable

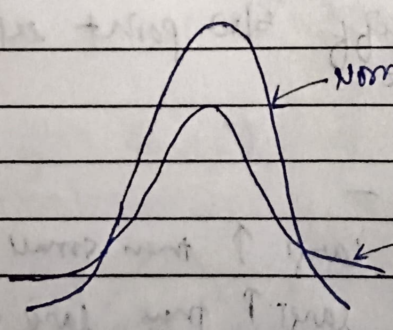
So we will not get normal dist

we will get Student's T- distribution



- It's theoritical
- parameter is degree of freedom
- df = n−1

n = Sample size



$\therefore$ df ↑ then: T-dist goes to Normal

So formula → $CI = \bar{x} + t_{\alpha/2} \frac{S}{\sqrt{n}}$ 't' from t-table