Central Limit Theorem

Normal Distribution

Popularian data → ND / ≠ ND

50 sample size

S1 → $x_1, x_2, x_3 \ldots x_{50}$ → $\bar{X}_1$

S2 → $x_1, x_2, x_3 \ldots x_{50}$ → $\bar{X}_2$

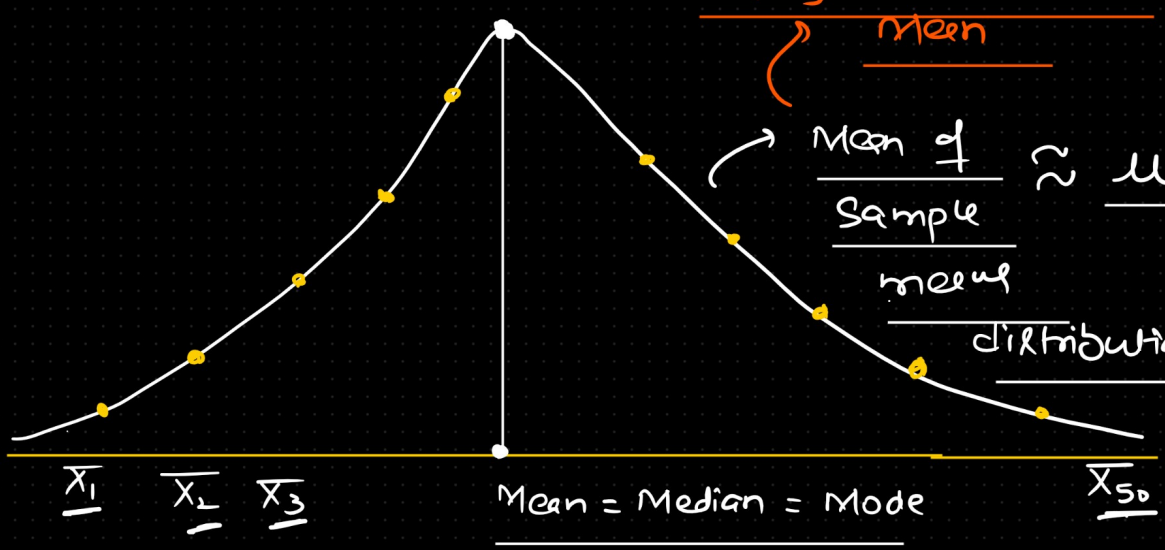
S3 → $x_1, x_2, x_3 \ldots x_{50}$ → $\bar{X}_3$

S_{50} → $x_1, x_2, x_3 \ldots x_{50}$ → $\bar{X}_{50}$

Sampling distribution

Mean

Mean of sample means distribution $\approx \mu$



✓ Standard Error ← Standard Deviation of Sample means distribution = $\frac{\sigma}{\sqrt{n}}$

n = Sample size

✓ Standard Deviation → Measure of the spread of the individual data points in the population

Vs

✓ Standard Error → Measur the spread of Sample means around the population mean

Real Use case ⟶ Zomato or Swiggy

↓

Delivery times of individual orders ⟶ Not a Normal Distribution

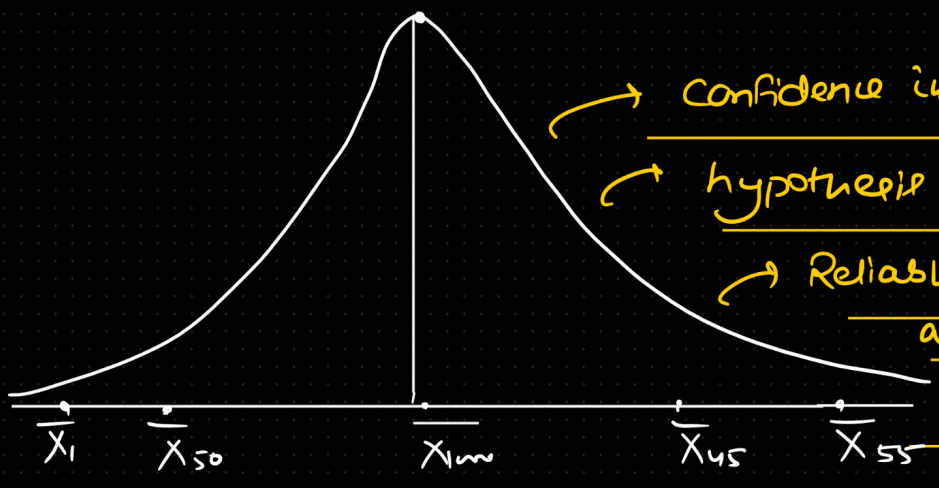
Random Samples of delivery times

S1 : 50 random deliveries ⟶ average time ($\bar{x}_1$)

S2 : 50 random deliveries ⟶ average time ($\bar{x}_2$)

|

S100 : 50 random deliveries ⟶ average time ($\bar{x}_{100}$)



→ Confidence interval

→ hypothesis testing

→ Reliable Predictions about average delivery time

$\mu_{\bar{x}} = \mu$