Central limit Theorem:→ If $\{X_i\}$ $(i=1,2,\ldots n)$ be Independently distributed random variables such that

$$E(X_i) = \mu_i \text{ and } var(X_i) = \sigma_i^2$$

then as $n \to \infty$, the distⁿ of the sum of these random variables, namely

$$S_n = X_1 + X_2 + \ldots + X_n$$

tends to a normal distⁿ with mean μ and variance σ^2, where

$$\mu = \sum_{i=1}^{n} \mu_i \text{ and } \sigma^2 = \sum_{i=1}^{n} \sigma_i^2.$$

Q A Coin is tossed 200 times. Find the appr. probability that the no. of heads obtained is b/w 80 and 120.

Solⁿ→ Let X: no. of heads

$n = 200$ is very large

∴ We will apply Central limit Theorem.

From binomial distⁿ,

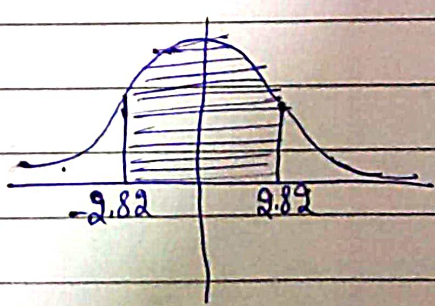
$$mean = np \;;\quad Var = npq$$

$$p = \frac{1}{2},\ q = \frac{1}{2}$$

$$\Rightarrow mean = 100 \;;\quad var = 50$$

$$\text{Req. Prob} = P(80 < X < 120)$$

$$= P\left(\frac{80-100}{\sqrt{50}} < Z < \frac{120-100}{\sqrt{50}}\right)$$



$$= P(-2.82 < Z < 2.82)$$

$$= 0.4976 + 0.4976$$

$$= 0.9952 \quad \text{(from normal Tables)}$$

Ex:- If $X_1, \ldots X_n$ are poisson variables with parameter 2, use the C.L.T. to Estimate $P(120 \le S_n \le 160)$ where

$S_n = X_1 + X_2 + \ldots + X_n$; $n = 75$.

Sol^n $X_1 \ldots X_n \sim P(\lambda)$ where $\lambda = 2$

$\Rightarrow$ Mean $= E(x_i) = 2$ and $V(x_i) = 2$

Req. Prob is $P(120 < S_n \le 160)$

$$E(S_n) = \sum_{i=1}^{75} E(x_i) = 75 \times 2 = 150$$

$$V(S_n) = \sum_{i=1}^{75} V(x_i) = 75 \times 2 = 150$$

$$\therefore P(120 \le S_n \le 160) = P\left(\frac{-30}{\sqrt{150}} < Z \le \frac{10}{\sqrt{150}}\right)$$

$$= P(-2.4495 \le Z \le 0.8165)$$

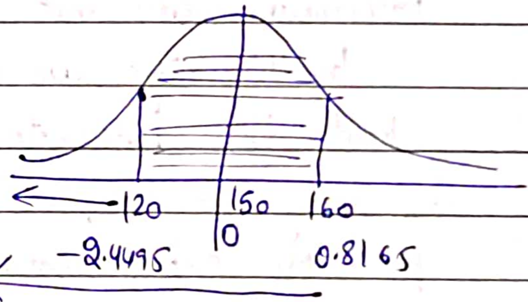
from normal tables

$P(Z < 0.8165) = 0.7939$

and $P(Z < -2.45) = 0.0071$

$\Rightarrow P(-2.45 \le Z \le 0.8165)$

$= 0.7868$.



Ex:- Let X_i's be Independent and Identically distributed r.v's with mean 3 and var $\frac{1}{2}$. Use C.L.T. to Estimate $P(340 \le S_n \le 370)$, where $S_n = X_1 + \ldots + X_n$ and $n = 120$.

Sol^n $E(x_i) = 3$; $V(x_i) = \frac{1}{2}$

$$E(S_n) = \sum_{i=1}^{120} (3) = 120 \times 3 = \cancel{225}\ 360$$

$$V(S_n) = \sum_{i=1}^{120} \left(\frac{1}{2}\right) = 120 \times \frac{1}{2} = 60$$

Thus $Z = \dfrac{S_n - 360}{\sqrt{60}}$

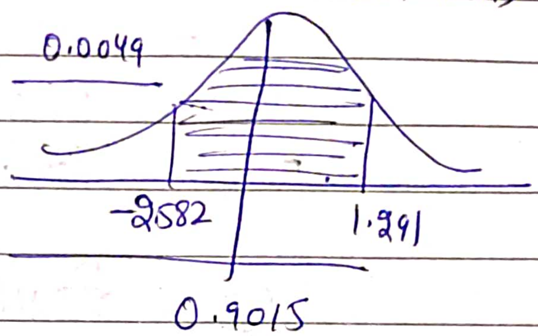
$\therefore$ Req Prob. is

$$P(340 \le S_n \le 370) = P\left(\frac{-20}{\sqrt{60}} \le Z \le \frac{10}{\sqrt{60}}\right)$$

$$= P(-2.582 \le Z \le 1.291)$$

$= 0.9015 - 0.0049$

$= 0.8966$



<u>Sampling Distribution</u>:→ In real life problems, we may be able to Identify which type of probability distribution to be used as a model, but the values of the Parameters are not known.

In this Case, we must rely on the sample to learn about these parameters.

- The way a sample is selected is called the sampling plan or experimental design.
- Simple random sampling is a commonly used sampling plan in which every sample of size n has the same chance of being selected.

Ex. Let we want to select a sample of size $n=2$ from a popⁿ containing $N=4$ Objects x_1, x_2, x_3, x_4.

- So total possible samples of size $2 = {}^4C_2 = 6$.

Sample	Observations in Sample	Prob of Selection
1	x_1, x_2	$1/6$
2	x_1, x_3	$1/6$
3	x_1, x_4	$1/6$

4	x_2, x_3	$1/6$
5	x_2, x_4	$1/6$
6	x_3, x_4	$1/6$

Each of the 6 Samples have same chance of being selected i.e. $1/6$, then the resulting sample is called a Simple random Sample.

Parameter – Numerical descriptive measures of the pop^n.

Statistics – Numerical descriptive measures of the sample

– The statistics are different for each random sample.

Sampling distribution:→ The prob. distribution for statistics.

Pop^n mean – μ
pop^n variance – σ^2
Sample mean – $\bar{x}$
Sample variance – s^2

– Std. deviation of the Sampling dist^n of a statistic is called Standard Error (S.E)

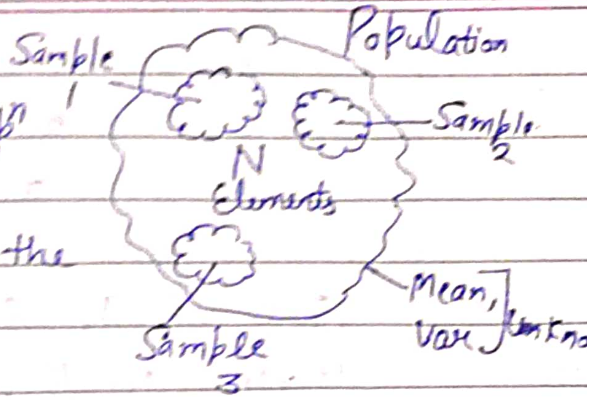
Sampling dist^n of mean:→ If random samples of n observations are drawn from a non-normal pop^n with finite mean μ and std. dev. σ, then when n is large, the sampling dist^n of sample mean $\bar{x}$ is appr. distributed with mean μ and std. deviation $\sigma/\sqrt{n}$. (using Central limit Thm)

i.e. $\bar{x} \sim N\left(\underset{\text{mean}}{\mu}, \underset{\text{var.}}{\frac{\sigma^2}{n}}\right)$ when $n \to \infty$ or n is large.

Types of Statistical Inference

(A) **Estimators:→**

To Estimate the value of the pop^n parameters, we can use the Information from the sample in the form of an estimator.



Estimators are used in two different ways:-

(1) **Point Estimation:→** Based on the sample data, a single number is calculated to estimate the pop^n parameter. The rule or formula that describes this calculation is called the point estimator and the resulting no. is called point estimate.

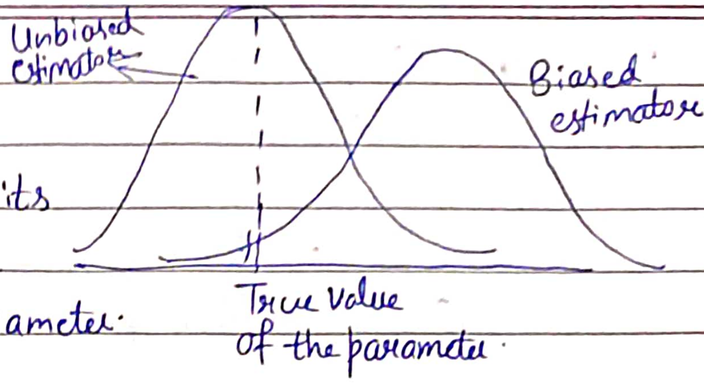
(2) **Interval Estimation:→** Based on the sample data, two nos are calculated to form an Interval within which the pop^n parameter is expected to lie. The rule or formula that describes this calculation is called interval estimator and the resulting nos are called an Interval estimate or Confidence Interval.

Point Estimator:→ There may be several statistics that could be used as point estimators for a pop^n parameter. To decide, which of the several choices is the best estimator.

Sampling dist^n helps to decide the best estimator.

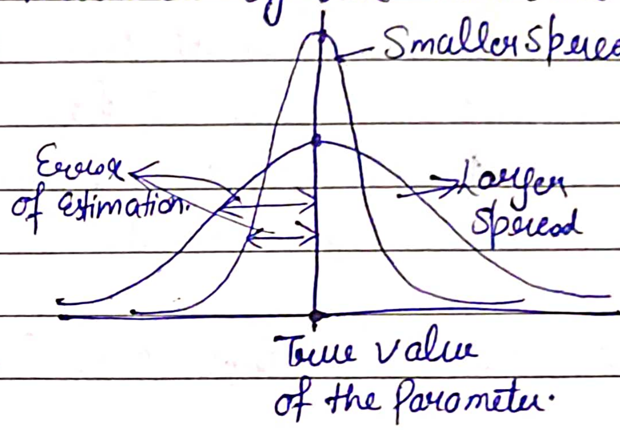
(1) Sampling dist^n of the point estimator should be centered over the true value of the parameter to be estimated.

i.e. the estimator must be unbiased, which means mean of its $dist^n$ is equal to the true value of the parameter.



(2) The spread of the sampling $dist^n$ should be as small as possible. Spread & is measured by the variance.

- The distance b/w the estimate and true value of the parameter is called error of estimation.



For Instance when we find estimate for pop^n mean (μ), we use sample mean ($\bar{x}$)

$$\bar{x} = \frac{\Sigma x}{n} \longrightarrow \text{Estimator}$$

Suppose $\bar{x} = 10$ $\longrightarrow$ our answer is in Single numeric form, it is Called point Estimator.

The whole process is Called Estimation.

Que A random Sample of $n=6$ has the Elements 6, 10, 13, 14, 18 and 20. Compute a Point Estimate of
(1) Pop^n mean
(2) Pop^n Std. dev.

Sol^n Sample mean $\bar{x} = \frac{\Sigma x}{n} = 13.5$

$\bar{x} = 13.5$ is point Estimate of $\boxed{\mu}$.

Date ____
Page ____

(2) The Sample Std. dev. is

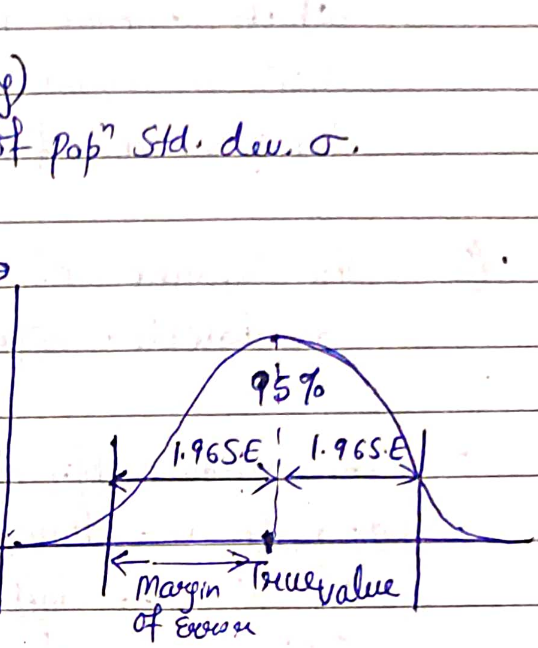
$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$\Rightarrow S = 4.68$ (after Solving)

$S = 4.68$ is the Point Estimate of popⁿ Std. dev. σ.

Sampling distⁿ of unbiased estimator :→

If the numeric value of Point Estimator lies with the Confidence Interval Parameter $\pm 1.96\,SE$, then it is a good estimator.



~~DEF The prob. that a Confidence Interval~~

Point Estimation of Popⁿ mean

- To estimate the Popⁿ mean μ, the point Estimator $\bar{x}$ is an unbiased estimator with Std. error $S.E = \frac{S}{\sqrt{n}}$.

The 95% margin of error is estimated as when $(n \geq 30)$ ↓ n is large.

$$\pm 1.96(SE) = \pm 1.96\left(\frac{S}{\sqrt{n}}\right)$$

$\because$ If $x \sim N(\mu, \sigma^2)$ then $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

where $S.E. = \frac{\sigma}{\sqrt{n}}$

If Std dev of popⁿ is Known then $S.E. = \frac{\sigma}{\sqrt{n}}$

If σ is Unknown then $S.E. = \frac{S}{\sqrt{n}}$ where S is sample Std. dev.

Date ______

Page ______