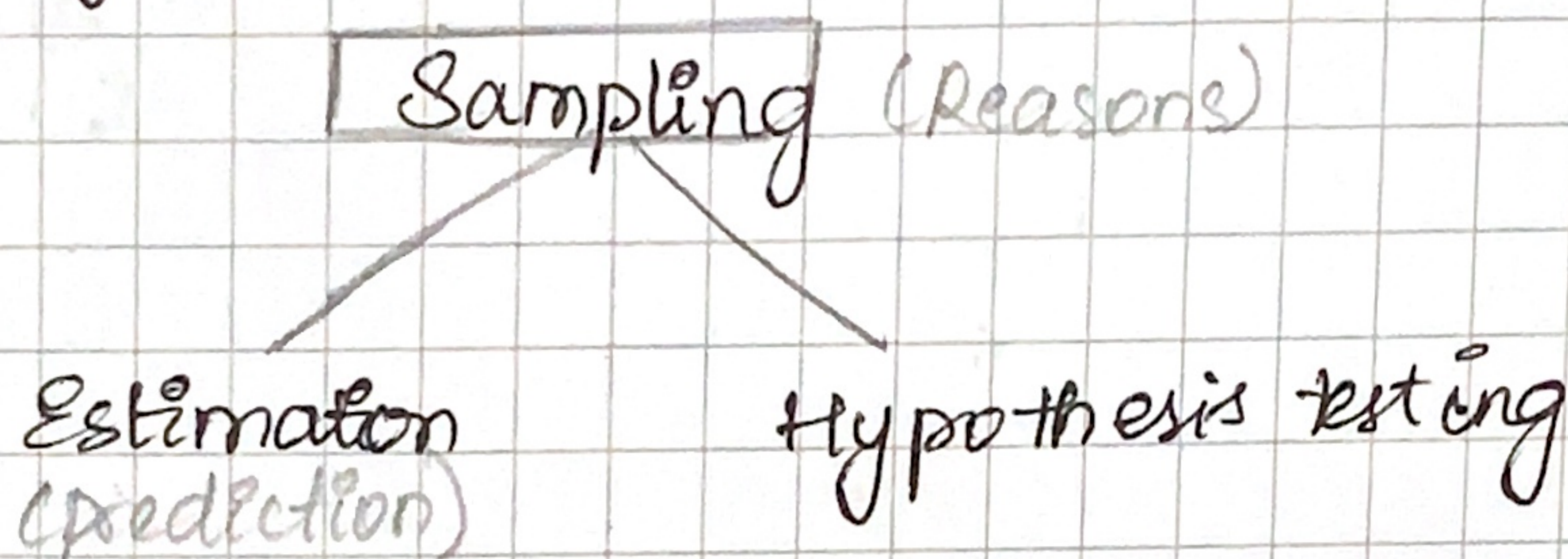


SAMPLING THEORY (1-29ns)

- What is Sampling?
- Purpose of Sampling
- Methods of Sampling
- Laws of Sampling.



A method where we draw the sample, study the population's sample and draw conclusions.

Population is divided into 3 categories

- Finite or Infinite
on counting, exhausted - FINITE
cannot be counted, or does not get exhausted - INFINITE
- Real or Imaginary
physically present - REAL
not physically existent - IMAGINARY/HYPOTHESIAL
- Homogeneous or Heterogeneous
 - population is same - HOMOGENEOUS
 - population is different & mixed - HETEROGENOUS

Calculation for population - PARAMETER μ or NP
Calculation for sample - STATISTIC \bar{x} or np

- Estimation of statistic helps to estimate the parameter
Ex: Sample Mean gives an idea about population mean

THEORY (1-2 qns)

asons)

thesis testing

le, study conclusions.

ries

sted - INFINITE

PROTHESIAL

ROGENOUS

ER - μ NP

- \bar{x} SNP

parameter

tion mean

Hypothesis - statement, idea, belief about the population which is statistically verified using sampling

Eg: A tubelight company claims that avg life of tubelight is 5000 hours which is statistically verified using sampling [being statement is right or wrong]

Hypothesis

Null

Alternate

[Sampling helps in auditing]

Sampling is done —

- with replacement
- without replacement

with replacement — chances of repetition (of element)
without replacement — no chances of repetition (as elements are removed out of popln)

∴ ∴ ∴

No of samples drawn with replacement = N^m

where N = population size
 m = sample size

No of samples drawn without replacement = $N C n$

Eg: $N=5$ 1, 2, 3, 7, 8
 $n=2$

No of samples drawn WR = $N^n = 5^2 = 25$

(1,1)	(2,1)	(3,1)	(7,1)	(8,1)
(1,2)	(2,2)	(3,2)	(7,2)	(8,2)
(1,3)	(2,3)	(3,3)	(7,3)	(8,3)
(1,7)	(2,7)	(3,7)	(7,7)	(8,7)
(1,8)	(2,8)	(3,8)	(7,8)	(8,8)

repetition is seen

combination

$$\text{No of samples drawn w/o R} = {}^N C_n = {}^5 C_2 = \frac{5 \times 4}{2 \times 1} = \underline{\underline{10}}$$

(1, 2) (2, 3) (3, 7) (7, 8)

(1, 3) (2, 7) (3, 8)

(1, 7) (2, 8)

(1, 8)

No repetition

*

- No of samples with replacement is more/greater than no of samples without replacement

$$N^n > {}^N C_n$$

In real time, only one sample is studied (not all samples)
- RANDOM SAMPLING

$N = 3$ a, b, c

$n = 2$

(a, a) (b, a) (c, a)

(a, b) (b, b) (c, b)

(a, c) (b, c) (c, c)

$$\text{NO of samples WR} = N^n = 3^2 = \underline{\underline{9}}$$

$$\text{WOR} = {}^N C_n = {}^3 C_2 = \frac{3 \times 2}{2 \times 1} = 3$$

(a, b) (a, c) (b, c)

There are two types of estimation

- Point estimation
- Interval estimation

Point Estimation: Parameter = Statistic
exactly the same. Eg: $\mu = \bar{X}$

- cannot calculate the error (not accurate)
* diff in actual & estimated value.
error due to variation

Interval Estimation: Parameter = Statistic \pm Marginal error

less error = Accurate info

% Accuracy is measured in terms of error

Point Estimate - 1 value of estimate

Interval Estimate - 2 values of estimate

(upper limit and lower limit)

LAWS OF SAMPLING [*]

- Law of Statistical regularity: sample size should be large on an average will give clear information about the population (affects the accuracy)

Sampling fraction: $\frac{n}{N} > 0.05$ (5% of pop'n)

- Law of Inertia of large number: if sample size is large, error in calculation is less.

* Sample size is inversely proportional to error.

$$SE_x = \frac{\sigma}{\sqrt{n}}$$

Standard error of mean

- Principle of Optimisation: when sample size is large, the time and cost of survey will increase. so select adequate sample size.

* Sample size is directly proportional to time & cost
Minimum cost = Maximum & clear information.

Sample size $n = \left(\frac{\sigma Z}{E} \right)^2$

E = permissible error
(diff b/w actual & estimated value)

Z = Marginal error.

σ - Standard deviation

- Principle of validity: Valid/Better results if we do random sampling/probability sampling [less time, cost, max info]

Every element selected in sampling gets an equal chance - PROBABILITY SAMPLING.

METHODS OF SAMPLING [•%]

- Random Sampling (Probability Sampling) VALID
- Non-Random Sampling (Non-probability Sampling)
- Mixed Sampling.

PROBABILITY

Random Sampling

- Lottery method
- Stratified Sampling
- Multistage sampling
- Systematic sampling
- Random number table method.

NON PROBABILITY

Non-Random Sampling

- Quota Sampling
- Convenience Sampling
- Judgemental Sampling
- Purposive Sampling

• Lottery method - equal chance - make charts and select random (when there is large sample size, numbering is difficult, bias might come out), therefore use random number table method

• Stratified sampling - Divide the area in small groups called strata which are homogeneous to do survey

• Multistage Sampling - Stage wise population selection to get overall view of population

• Systematic sampling - Numbering the population and decide sample size, get the interval and do the sampling. (popn/sample)
 * Periodicity is observed

• Sur
 • Non
 and

results if
sampling

gets an
LING

VALID
ity sampling

ILITY
Sampling
Sampling
Sampling
Sampling

and
age, number
use random

small groups

ation

on
and
able)

• Quota Sampling - not all data gets equal chance due to predetermined quota. ↳ to represent

• Convenience Sampling - Survey is done in a convenient location where respondents are available. ↳ where we get the information required.

• Judgemental Sampling - Selecting ^{sample from} population as per a person's understanding / judgement

• Purposive Sampling - Selecting sample from population according to the purpose.

Mixed Sampling - mixture of Random and Non-Random Sampling

• Cluster sampling : area which has heterogeneous population, divided into clusters. → used to select the sample.

ERRORS IN SAMPLING / STATISTICS

- Sampling Error
- Non sampling error

Error due to sampling process - SAMPLING ERROR

Errors caused due to other - NON SAMPLING ERROR
reason (not sampling) (also called bias)

* due to human error, calculation error, faulty use of technique/instrument. - NON SAMPLING ERROR.

- Sampling errors happens in sample survey
- Non sampling errors happens in both census survey and sample survey.