SNS COLLEGE OF TECHNOLOGY

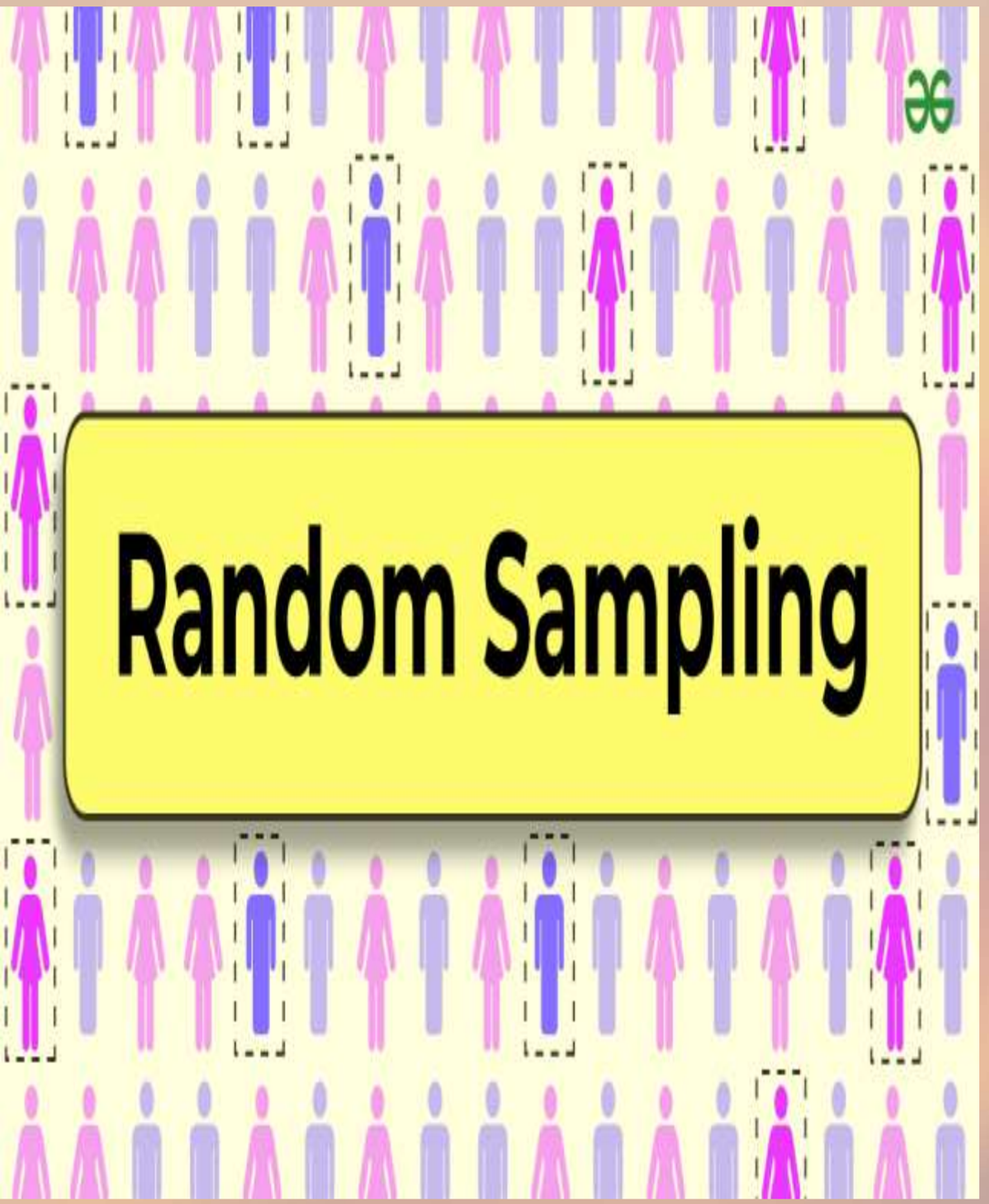
An Autonomous Institution
Coimbatore-35

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT III – RANDOM SAMPLING AND SAMPLING DISTRIBUTION



EMPATHY:

What is Random Sampling?

- It is a method used in statistics to select a subset from a larger population in such a way that each member of the population has an equal and independent chance of being included in the sample.

- It is a fundamental technique for conducting surveys and experiments.

DEFINE:

Random Sampling Definition

- Random sampling is a method employed for selecting observations from a population, facilitating generalization about the entire population.

- Random Sampling is sometimes referred to as probability sampling, distinguishing it from non-probability sampling.

- This method encompasses various techniques, including simple random sampling, stratified sampling, cluster sampling, and multistage sampling.

Types of Random Sampling

Simple Sampling

•Involves randomly selecting items without any specific pattern or criteria.

Systematic Sampling

•Selects individuals at regular intervals, offering an organized yet random way to choose a sample

Stratified Sampling

• Divides the population into distinct strata or subgroups and then randomly samples from each stratum, enhancing representation.

Cluster Sampling

•Organizes the population into clusters, randomly selects some of these clusters, and samples all individuals within the chosen clusters.

IDEATE:

How to Perform Simple Random Sampling?

Define the Population: Clearly identify the entire group from which you want to draw a sample.

Assign Numbers: Assign a unique number to each member of the population.

Use a Random Method: Select members using a random method, such as:

Random Number Generator: Use software or a calculator to generate random numbers corresponding to the assigned numbers.

Lottery Method: Write numbers on slips of paper, mix them, and draw the required number of slips.

Select Sample Size: Determine the sample size you need and select that number of individuals based on the random method used.

Collect Data: Gather data from the selected individuals to conduct your study.

PROTOTYPING:

Random Sampling Formula

Formula of random sampling is mentioned as below:

$$P = 1 - [(N-1)/N] \times [(N-2)/(N-1)] \times \ldots \times [(N-n)/(N-(n-1))]$$

Where,
P represents probability,
n represents sample size, and
N represents population.
In above formula cancelling 1-(N-n/n), it will yield a value of P = n/N.

So, sample getting selected for a chance of more than once

$$P = 1 - (1 - (1/N))^{n}$$

Where,
P represents probability,
n represents sample size, and
N represents population.

TESTING:

Random Sampling Examples

Example 1: A company has 500 products, and they want to randomly select 20 of them for quality testing. What is the probability of any single product getting selected?

Solution:

The chance of one-time selection is:

$P = n/N$

$\Rightarrow P = 20/500$

$\Rightarrow P = 4\%$

Example 2: In a conference with 200 attendees, 50 will be randomly chosen for a survey. What is the probability that one attendee gets selected more than once?

Solution:

The probability of getting selected more than once is:

$$P = 1 - (1 - (1/N))^n$$

$$\Rightarrow P = 1 - (1 - (1/200))^{50}$$

$$\Rightarrow P \approx 9.56\%$$