



Centurion
UNIVERSITY
Shaping Lives...
Empowering Communities...

SESSION-3

SAMPLES AND POPULATION

- **Population**-The population consists of the set of all measurements in which the investigator is interested. It is also called as **universe**.
- **Sample**-A sample is a subset of measurements selected from the population for analysis.
- **Statistic**- A statistic is a summary measure computed from sample data.
- **Parameter**- It is a summary measure that describes a characteristic of an entire population.
- **Statistical Inference**- A conclusion drawn about a population based on the information in a sample from the population is called a statistical inference.

Inferential methods use sample data to calculate summary measures called **statistics** that decision makers can use to estimate the **parameters** characteristics of the entire population.



Centurion
UNIVERSITY
*Shaping Lives...
Empowering Communities...*

Example: Students of CUTM practicing Yoga.

- Population
- Sample
- Sample & Population relative (If compared to all Universities of Odisha)



Centurion
UNIVERSITY
*Shaping Lives...
Empowering Communities...*

SAMPLING VS. CENSUS

- A sample is less time-consuming than a census
- A sample is less costly to administer than a census
- A sample is less cumbersome and more practical to administer than a census.



TYPES OF SAMPLING

Non-probability samples -Here the items or individuals are chosen without regard to their probability of selection.

- Judgment sampling
- Convenience sample

Advantage- convenience, speed, lower cost.

Disadvantage-lack of accuracy due to selection bias, lack of generalizability of the results.

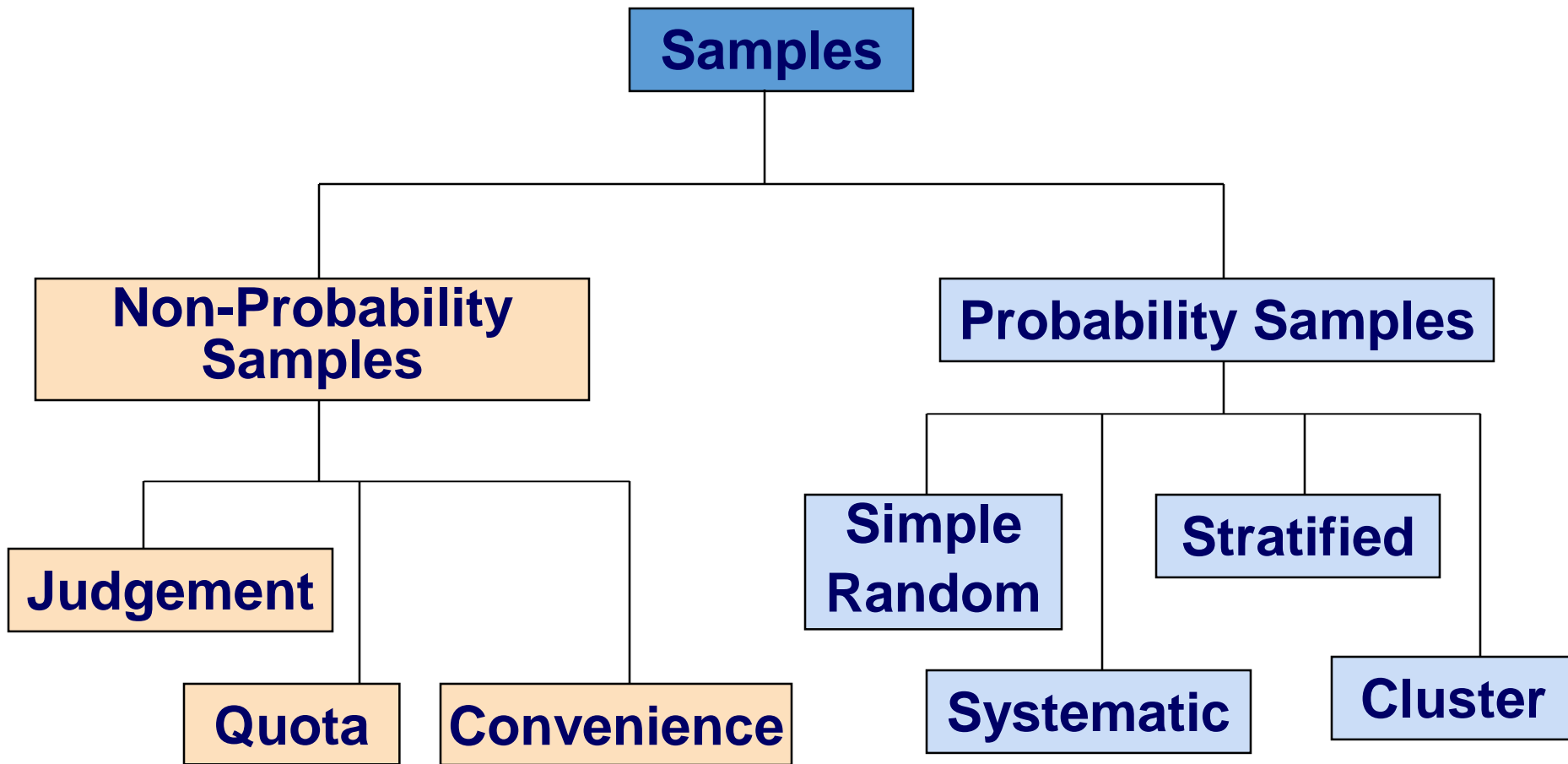
Probability samples – A probability sample is one in which the subjects of the sample are chosen on the basis of their probability.

- Simple Random sample
- Systematic sample
- Stratified sample
- Cluster sample



TYPES OF SAMPLES USED

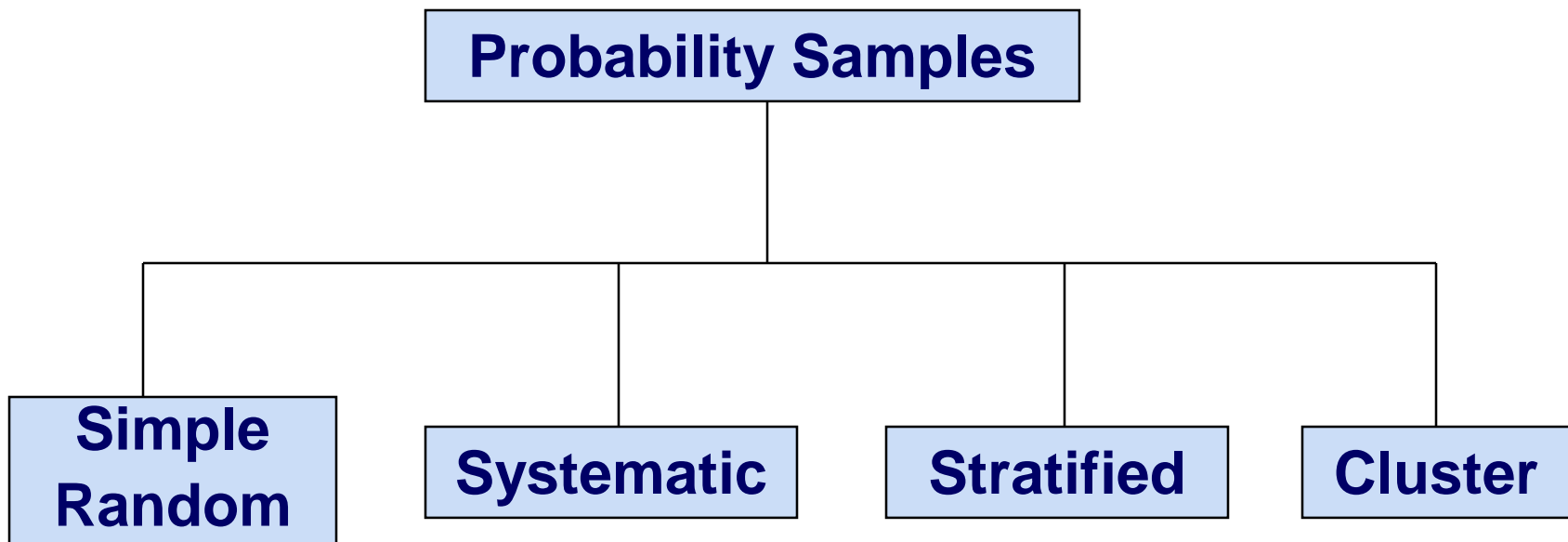
(continued)





PROBABILITY SAMPLING

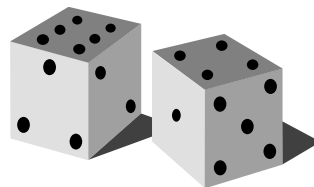
- Items in the sample are chosen based on known probabilities





SIMPLE RANDOM SAMPLES

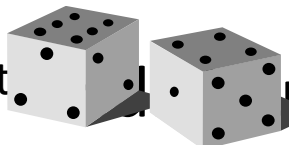
- Every individual or item from the frame has an equal chance of being selected
- Selection may be **with replacement** or **without replacement**
- Samples obtained from table of random numbers or computer random number generators





SYSTEMATIC SAMPLES

- Decide on sample size: n
- Divide frame of N individuals into groups of k individuals: $k=N/n$
- Randomly select one individual from the 1st group
- Select every k^{th} individual thereafter



$N = 64$
 $n = 8$
 $k = 8$

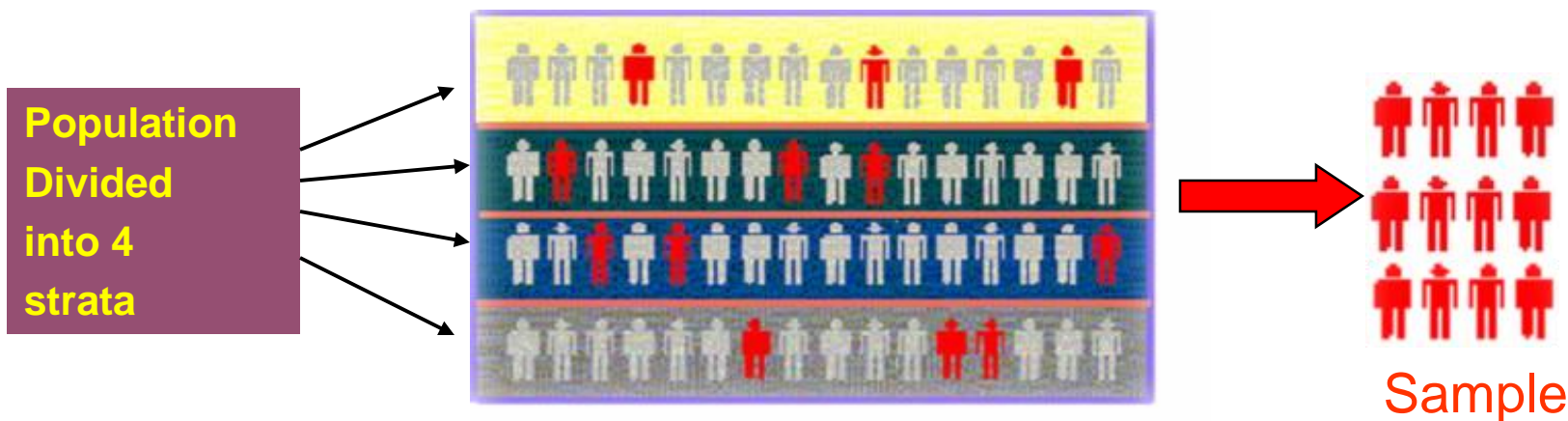
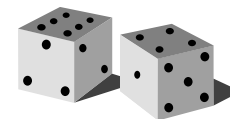
First Group





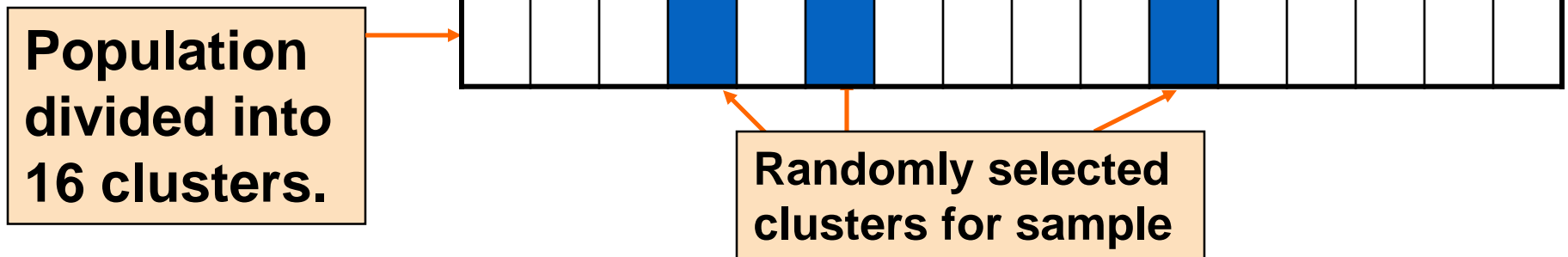
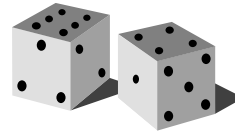
STRATIFIED SAMPLES

- Divide population into two or more subgroups (called *strata*) according to some common characteristic
- A simple random sample is selected from each subgroup, with sample sizes proportional to strata sizes
- Samples from subgroups are combined into one



CLUSTER SAMPLES

- Population is divided into several “clusters,” each representative of the population
- A simple random sample of clusters is selected
 - All items in the selected clusters can be used, or items can be chosen from a cluster using another probability sampling technique





ADVANTAGES AND DISADVANTAGES

- Simple random sample and systematic sample
 - Simple to use
 - May not be a good representation of the population's underlying characteristics
- Stratified sample
 - Ensures representation of individuals across the entire population
- Cluster sample
 - More cost effective
 - Less efficient (need larger sample to acquire the same level of precision)

EVALUATING SURVEY WORTHINESS

- What is the purpose of the survey?
- Is the survey based on a probability sample?
- Coverage error – appropriate frame?
- Nonresponse error – follow up
- Measurement error – good questions elicit good responses
- Sampling error – always exists



TYPES OF SURVEY ERRORS

- Coverage error or selection bias
 - Exists if some groups are excluded from the frame and have no chance of being selected
- Nonresponse error or bias
 - People who do not respond may be different from those who do respond
- Sampling error
 - Variation from sample to sample will always exist
- Measurement error
 - Due to weaknesses in question design, respondent error, and interviewer's effects on the respondent



TYPES OF SURVEY ERRORS

(continued)

- Coverage error



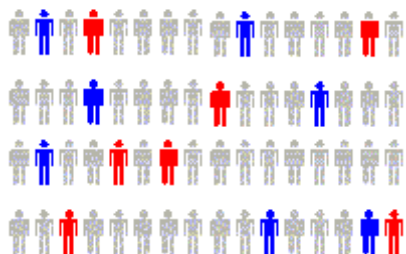
**Excluded from
frame**

- Non response error



**Follow up on
nonresponses**

- Sampling error



**Random
differences from
sample to sample**

- Measurement error



**Bad or leading
question**