

Chapter-5. SAMPLE DESIGN

What is sample design???

- **Sample design:** is a definite plan for obtaining a sample from a given population.
- It refers to the technique or the procedure the researcher would adopt in selecting items for the sample.
- Sample design is determined before data are collected.
- There are many sample designs from which a researcher can choose. Some designs are relatively more precise and easier to apply than others.
- Researcher must select/prepare a sample design which should be **reliable** and **appropriate** for his/her research study.

Characteristics of a good sample design

From what has been stated above, we can list down the characteristics of a good sample design are:

- It must result in a truly **representative sample**.
- It must be such which results in a **small sampling error**.
- It must be **feasible** in the context of **funds** available for the research study.
- Sample should be such that the results of the sample study **can be applied**, in general, for the universe with a **reasonable level of confidence**.

STEPS IN SAMPLE DESIGN

1. *Type of universe:* clearly define the set of objects, technically called the Universe, to be studied.

2. *Sampling unit:* A decision has to be taken concerning a sampling unit before selecting sample.

3. *Source list:* It is also known as “sampling frame” from which sample is to be drawn.

4. *Size of sample:* This refers to the number of items to be selected from the universe to constitute a sample

5. *Parameters of interest:* In determining the sample design, one must consider the question of the specific population parameters which are of interest.

6. *Budgetary constraint:* Cost considerations

7. *Sampling procedure:* Finally, the researcher must decide the type of sample he/she will use.

TYPES OF SAMPLING

- Non-probability samples
- Probability samples

QUESTIONS???

I. We want to estimate the total income of adults living in a given street. We visit each household in that street, identify all adults living there, and randomly select one adult from each household.

II. We visit every household in a given street, and interview the first person to answer the door.

TYPES OF SAMPLE DESIGNS

1. Non-probability samples

Is any sampling method where some elements of the population have *no* chance of selection (these are sometimes referred to as '**out of coverage**'/'**under covered**'), or where the probability of selection can't be **accurately** determined.

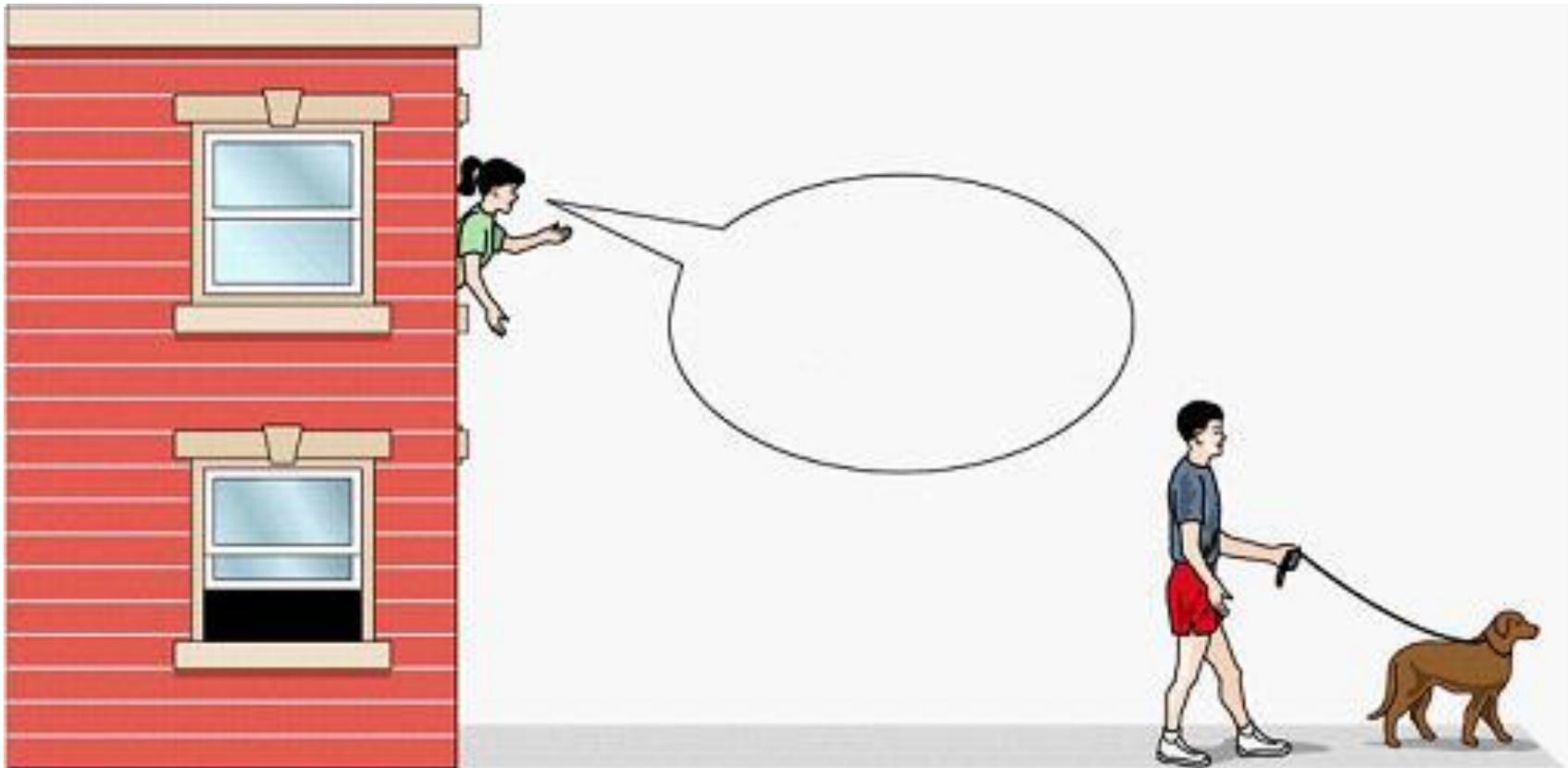
2. Probability samples

A probability sampling scheme is one in which every unit in the population has a **chance of being selected** in the sample, and this probability can be accurately determined.

TYPE OF NON-PROBABILITY SAMPLING

Convenience

Use who is available Selection of whichever individuals are easiest to reach



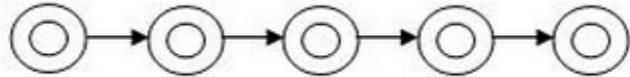
Purposive

Is a sampling technique in which researcher relies on his or her own judgment when choosing members of population to participate in the study.



SNOWBALL

Finding a small group of initial respondents and using them to recruit more respondents.



QUOTA

Keep going until the sampled are reach

	Chocolate Buyers	Respondent quota (sample size = 200)
Men	40%	80
Women	60%	120

QUESTION???

I. An interviewer may be told to sample 200 females and 300 males between the age of 45 and 60.

II. If the interviewer were to conduct such a survey at a shopping center early in the morning on a given day, the people that he/she could interview would be limited to those given there at that given time,

III. You may be conducting a study on why high school students choose community college over university. You might canvas high school students and your first question would be —Are you planning to attend college? People who answer —No, would be excluded from the study.

III. Meeting a homeless person, interviewing that person, and then asking him/her to introduce you to other homeless people you might interview.

TYPES OF PROBABILITY SAMPLING

- Simple Random Sample
- Stratified Random Sample
- Cluster sampling
- Systematic



I. SIMPLE RANDOM SAMPLING

- Every subset of a specified size “ n ” from the population has an equal chance of being selected



A subset of the population.

- The procedure for selecting a Simple Random Sample is as follows

HOW TO SELECT THE SAMPLE?

- List all the units in the population (construct a sampling frame if one does not exist already), say from 1,...,N
- In other words, give each element a unique Identification (ID) starting from 1 to the number of elements in the population N
- Using random numbers or any other random mechanism (eg Lottery or goldfish bowl), select the sample of n units from the list of N units.

- Sampling with replacement (WR) - units can be selected more than once.
- Sampling without replacement (WOR) - units cannot be selected more than once
- It can be done either using the lottery method or table of random numbers
- A table of random numbers consists of digits from 0 to 9, which are equally represented with no pattern or order, produced by a computer random number generator.

HOW TO USE A RANDOM NUMBER TABLE?

This process of selecting a large sample using random number tables is tedious

Example

8442 5653 8775 1891 7666 6483 9711

6941 8092 3875 4200 6543 9063 1003

8754 2564 8890 4195 8888 6490 3476

Suppose we want to select 5 elements from a population of 8000.

Start anywhere in the table and going in any direction choose a number/(s)

Suppose for convenience we start at the top left hand corner and read across

We need to use four digits at a time as there is a minimum of four digits in our sample ID's.

Let us start. The first set of 4 digits is 8442

8442 not in our population-ignore

5653-use for sample -1

8775-ignore

1891-use for sample -2

7666- use for sample -3

6483- use for sample -4

9711-ignore

6941-use for sample- 5

EXERCISE

Exercise: Consider a population with size $N = 4000$. Suppose it is desired to take a sample of 25 items out of 4000 without replacement, if we start from row one, column six and read down columns. (from random table)

We Obtain

1501, 3997, 0690, 1434, 3685, 1276, 2138, 3236, 2700,
3306, 2059, 0421, 2641, 0471, 1858, 0752, 2827, 2951,
0035, 1049, 1015, 3255, 0225, 3062, 2636 by ignoring all numbers greater than 4000.

II. STRATIFIED SAMPLING

- It is done when the population is known to have heterogeneity with regard to some factors and those factors are used for stratification
- Using *stratified sampling*, the population is divided into homogeneous, mutually exclusive groups called strata.

Example: region, rural/urban, zone age, sex, province of residence, income, etc.).

Women



Men



CONT.

Why do we need to create strata?

- To reduce sampling error
- When there is separate estimates are required at the stratum level.
- Stratified sampling is administratively convenient.
- Sometimes, different parts of the population may call for different sampling procedures.

➤ **III. SYSTEMATIC SAMPLING**

- Means the selection of units at fixed interval from a list, starting from a randomly determined point.
- Every k^{th} member (for example: every 3rd person) is selected from a list of all population members.
- Population size N , desired sample size n , sampling interval $k=N/n$.
- Randomly select a number j between 1 and k , sample element j and then every k^{th} element thereafter, $j+k$, $j+2k$, etc.

CONT.

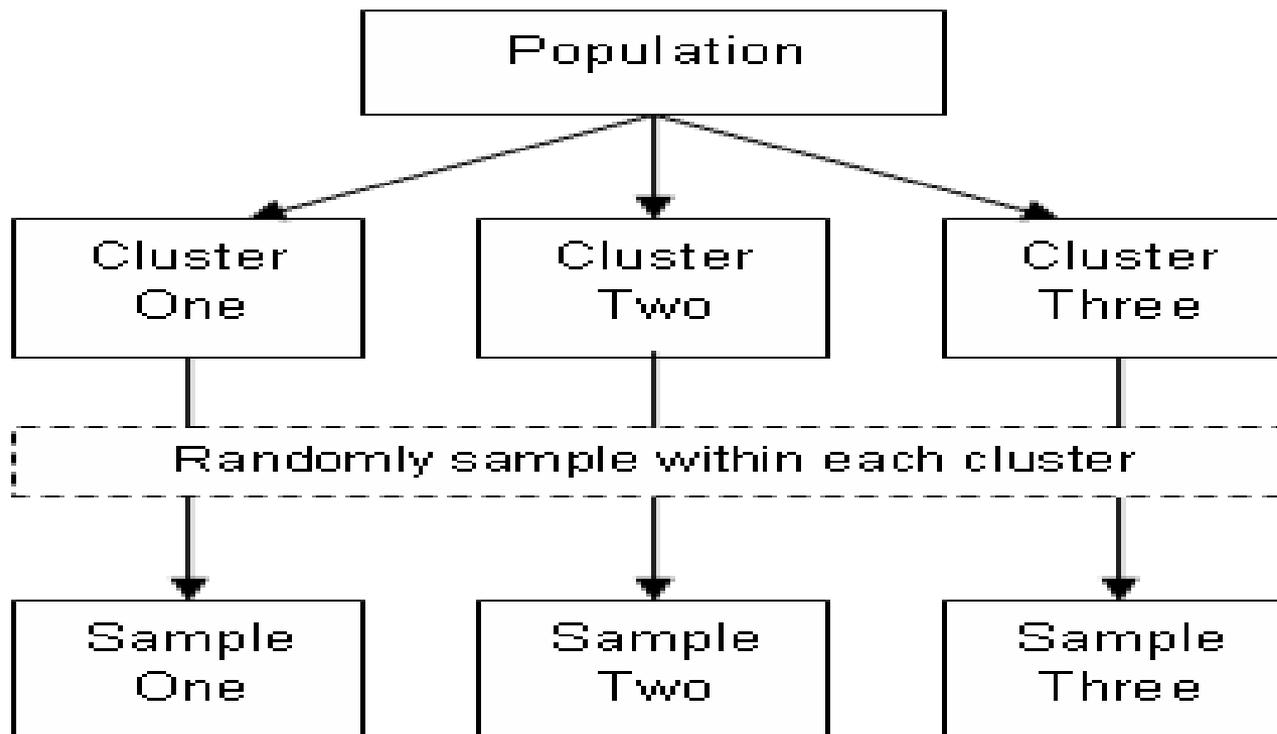
Example

Let $N=150$ and $n=10$ then $k=N/n=150/10=15$, then select one of the first 10 elements at random. If 2 is the selected number, since *2 is between 1 and 15* then the list of all the selected elements are:

2,17,32,47,62,77,92,107,122,137

IV. CLUSTER SAMPLING:

- The population is divided into subgroups (clusters) like families. A simple random sample is taken of the subgroups and then all members of the cluster selected are surveyed.



DIFFERENCE BETWEEN STRATA AND CLUSTERS

- Although strata and clusters are both non-overlapping subsets of the population, they differ in several ways.
- All strata are represented in the sample; but only a subset of clusters are in the sample.
- With stratified sampling, the best survey results occur when elements within strata are internally homogeneous.
- However, with cluster sampling, the best results occur when elements within clusters are internally heterogeneous

TYPES OF CLUSTER SAMPLING?

Single-stage cluster sampling

Is a sampling plan in which cluster are chosen by SRS in only one step and every unit with in the selected clusters are include in z sample.

Two-stage cluster sampling

obtained by first selecting a sample of clusters, and then selecting a sample of elements from each sampled cluster.

multistage cluster sampling.

Is a process of sampling by which several stages of sampling are often involved

CONT.

- In *single-stage* cluster sampling, you divide the entire sample frame into clusters, usually based on some naturally occurring geographic grouping (e.g. city, town village, hospital).
- Then you **sample** these clusters and measure every element within the **selected** clusters.
- In *two-stage* cluster sampling, the sample of elements is obtained as a result of two stages of sampling.
- The population elements are first grouped into disjoint subpopulations, called primary sampling units (**PSU**). Then, in a first-stage sampling, a sample of PSU is drawn.

CONT.

- In the second-stage sampling units (**SSU**) may be clusters of elements, for each PSU in the first-stage sample.

For example,

- In single-stage sampling, you might take a SRS of cities. Within each city, you would measure characteristics of all schools.
- In a two-stage sampling plan, you would take a SRS of cities, and then within each city, you would list out every schools. Then you would take a SRS of schools.