1. Introduction

Welcome to the series of E-learning modules on Population. In this module we are going to cover the basic concept and definitions of population, various types of population and its characteristics, population in real life and statistical population, role of population Sample theory, target population and sample population.

By the end of this session, you will be able to:

- Explain Statistical Population
- Discuss various types of population
- Describe the nature of these populations
- Explain the role of population in Sample theory
- Describe the target population and sample population
- Explain frame and units

In any statistical investigation the interest usually lies in studying the various characteristics relating to items or individuals belonging to a particular group. This group of individuals under study is known as the population.

For example: If an enquiry is intended to determine the average per capita income of the people in a particular city, the population will comprise all the earning people in the city.

On the other hand if we want to study the expenditure habit of the families in that city, then the population will consist of all the house-hold in that city.

Further, if we want to study the quality of the manufactured product in an industrial concern during the day then the population will consist of the day's total production.

In Statistics and in general life, population has different meanings.

For example:
In daily life, population means a number of people living in a particular geographical area whereas Statistics population consists of all the units currently under study or all the number of observations connected with the study.

Hence, from a statistical point of view the population refers to the total of items or units in any field of enquiry about which information is desired is also known as a Universe.

Population may be viewed as the collection of living as well as non-living things. Hence, a population may be that of people, plants, books, farms, soldiers etc.

The attributes that are the object of study are referred to as characteristics and the units possessing them are called elementary units. The aggregate of which is generally described as population. Thus, all the elementary units (on the basis of one characteristic or more) constitute population.

For example: If we want to know the average marks in Statistics of five hundred students then the population in this study will consists of five hundred students. Or if we want to know the total population of India then all the people residing in India will constitute the Universe or Population.

2. Various Definition ,Types of Population (Part 1)

Let us discuss various definitions of population:
In the words of M. Hamberg “The Universe or population consists of the total collection of items and elements that fall within the scope of a statistical investigation”.
In the words of Simpson and Kafka “A Universe or a population may be defined as an aggregate of items possessing a common trait or traits”.

As per Kalton “In statistical usage, the term population does not necessarily refer to people but is a technical term used to describe the complete group of persons or objects for which the results are to apply”.

Thus, “In Statistics, population is the aggregate of objects, animate or inanimate, under study in any statistical investigation.” In sampling theory the population means a large group from which the samples are drawn. On the basis of sample study we can predict and generalize the behavior of the mass of phenomena. This is possible because there is no statistical population whose elements would vary from each other without limit.

For example: Wheat varies to a limited extent in colour, protein content, length, weight etc. It can always be identified as wheat. Similarly the apples of the same tree may vary in size, colour, shape, taste, weight etc. but they can always be identified as apples.
Thus, we find that although diversity is a universal quality of mass of data every population has characteristic properties with limited variation.

Let us discuss the types of population
Population is of the following four types:

1. Finite Population
2. Infinite Population
3. Existing population
4. Experimental population or Hypothetical population

1. Finite population: Finite population is one in which the number of units of information is definite and limited.
For example: The college is a finite population and in this case the number of units of information that is students is definite and limited to five hundred.

Similarly the population of patients in a hospital under treatment is a finite population because there is definite and finite number of patients in a hospital.
Hence, the students in a college, the day’s production in an industrial concern, the scooters in Agra, the population of a city or a town, the customers in a shopping mall, etc. are the example for finite population.

2. Infinite Population: An infinite population refers to a population in which the number of units comprised therein cannot be definitely ascertained. An infinite population is theoretically

impossible to observe all the elements. Thus, in an infinite population the number of items is infinite. For a practical consideration we then use the term infinite population for a population that cannot be enumerated in a reasonable period of time. This way we use the theoretical concept of infinite population as an approximation of a very large finite population. For example: The population of fish in an ocean, the population of temperature at various points of the atmosphere, the production of wheat in India, the population of heights and ages of the people in a country , the population of weights of people in a city, are the examples of infinite population.
Even though it may be possible to measure the heights of all the persons of India or the production of wheat in this country or the weights of people in a city, the actual values would always vary within certain limits.

The series that we shall get in such cases would be continuous as exact measurements are not possible. Further if a population is very large, it is also regarded as an infinite population such as number of leaves of a tree. In fact Infinite population are better for Sampling studies and also the probabilities of various events can be better estimated if the population is infinite. In case of studies relating to biological sciences we often come across with infinite population.

For example: Studies related to the population dynamics such as growth rate in terms of length, weight, etc. of a particular species of any group of insects, animals , birds, fish, etc. we have to deal with an infinite population wherein the total number of these categories are not definite.
In case of agricultural studies such as study related to the diseased plants in a region due to the attack of wilt, pests etc. the population of plants can not be considered as finite.

3. Existing population:
An existing population is the one which already exists with all its units in the form of concrete objects. The investigator has nothing to do for its creation except its discovery and location.

For instance: a college, an University, a library, Indian air force, a country and a state consisting of its concrete objects like students, books, airplanes, individuals respectively are examples of existing universe.

In such population the values of probabilities p and q remains constant in various trials and this is a very important property of existing populations. The probability of a dice falling with number six upwards will always remain one by six in all possible throws and this property enables us to fit a particular curve to such data with a high degree of accuracy. Such a population also may come either under a finite or under an infinite population.

3. Types of Population (Part 2)

4. Experimental (Hypothetical) Population:
An Experimental Population is one which is constituted through experiments being conducted by an investigator and is not found already in existence. This population consists of imaginary objects.

For example: A record made of the number of heads and tails got upward by tossing a coin for a number of times say, hundred, two hundred, five hundred or one thousand times is a case of experimental population. We can construct a population by throwing the dice large number of times and recording its results.

This population consists of an infinite number of units because we can go on throwing the dice any number of times we like, unless of course, it wears out.

Here, the universe is not already in existence but is created by the investigator himself through his experiments for a number of times. An Experimental Population is generally of infinite character as there is no limit to the number of times an experiment may be conducted for recording the happenings of a particular event.

In sampling, our main objective is to study the characteristic of the population by collecting information about the population. A population represents a Universal set and we can define any number of subsets from the universal set. In the process of studying the characteristics of the population if we enumerate each and every unit of the population and then draw conclusions about the population, such a technique of drawing information is known as census survey or complete enumeration survey.

For example: Population census of India which takes place once in ten years. In this census each and every individual of our country provide required information to the enumerators.
But in case of very large population or an infinite population it is impossible for anybody to investigate each and every unit of the population. In such cases we study only a representative portion of the population. Then generalise the results that we get for the part of the population to the entire mass of data. In which case sampling techniques plays a major role in studying the characteristics of the population.

In most of the research work and surveys, the data are generalized or inferences are drawn based on a cross section of the population.
The researcher selects only a few items from the population for his study purposes. All this is done on the assumption that the portion of the data will enable him to estimate the population constants.

But the population units that we select should truly represent the population characteristics.
Then without any bias our study may result in valid and reliable conclusions.
Consider a situation where the Government is debating some drug control laws among the college students. We are asked to conduct an opinion survey. Here one has to first decide about the population.

Whom we have to interview?
Whether all the college students of our country? Or the students of a particular region? We can go for either census survey or survey of only a portion of the college students. But whatever the type of technique we use to infer about the population under consideration, one has to define the population of study in clear words. That is in terms of geographical region, age group, sex etc.

Hence, among the principal steps in developing any sampling design, first and the most important stage is the definition of the population.
One has to clearly define the set of objects about which study has to be made. For example: The definition of the population may present no problem when we have to study the average life of bulbs in a manufactured lot.

In the study of agricultural farms, clear rules must be framed regarding the size, shape etc. of the farm giving importance for the border line cases. So that an investigator will be able to decide in the field without any hesitation whether to include or not to include a given farm in the population.

Our main objective must be to extract maximum information about the characteristics of the population with the available sources at our disposal in terms of time, manpower, and money and to obtain the best possible estimates of the population constants with limited resources.

Hence, before studying the population one should have an idea about the type of the population and some knowledge about characteristics of the units of the population.
The symbol N is generally used to indicate how many elements or items are there in case of a finite population known as population size. If a population consists of N units then they may be represented as
Y one, Y two, Y three, etc, Y N. Capital Letters are used to denote the population units.

Statistical constants parameters explain the characteristic of a population under study. These constants are in the form of a function of population observations such as a population mean (mue), the variance (sigma square), the skewness (beta one), kurtosis (beta two), correlation coefficient (rho) etc.

4. Target Population and Sample Population (Part 1)

Let us now discuss what is Target population and Sample population
The population to be sampled is known as a sample population and the population about which information is wanted is known as the target population. The sample population should coincide with the target population. Sometimes for reasons of practicability or convenience the sampled population is more restricted than the target population.
Due to certain practical limitations and problems in dealing with certain units of the population, they are usually eliminated from the scope of the survey.

It should be remembered that the conclusions drawn from the part of the population is known as sample applied to the sample population. Judgement about extent to which these conclusions will also apply to the target population must depend on other sources of information. Any supplementary information that can be gathered about the nature of the differences between sampled and target population like errors in estimation, precision in estimation etc. may be helpful to generalize sample population results to the target population.

Units and a Frame
A population must be capable of division into what are called sampling unit or units. These units must cover the entire population and must not overlap in the sense that every unit in the population should belong to one and only one unit. These are the units on which the observations are made. Sometimes the appropriate unit is obvious as in the population of electric bulbs in which the unit is the single bulb.

Sometime there is a choice of unit. In a study of people in a city the unit might be an individual person, the members of a family, or all persons living in the same street, block in a locality, livestock, etc. In an agricultural experiment the unit might be a field, a farm or an area of land whose and dimensions are at our disposal. A sampling unit should be unambiguous, specific, stable and appropriate to the enquiry.

The construction of this list of sampling units is called a frame. Frame can be defined as some list or map or other acceptable material which covers the population decided upon and which serves as a guide for the population to be covered. This contains detailed information of the units of the population.

Since, frame is the one which determines the structure of the survey, it has become one of the major practical problems. A frame which has been already prepared for some other purpose has to be scrutinized and should be checked to see that it is free from all sort of defect like unknown amount of duplication and should be brought up-to-date before using them. A good experience helps in constructing a good frame but a good frame may be hard to come.

5. Target Population and Sample Population (Part 2)

For example:
1) In order to make a study on the villages, we must have a map of districts and villages of the state
2) For selecting households we must have a list of blocks in the locality
3) For selecting a group of students in a college the list of students enrolled in a college is needed.

Populations that can be sampled from a good organizational list include elementary school, High school and University students and faculty, church members, factory workers, fraternity or sorority members, members of social, service or political clubs and members of professional associations, Government registers on births and deaths.

Telephone directories are frequently used for “quick” public opinion polls. Undeniably they are easy and inexpensive to use – no doubt the reason for their popularity.

But telephone directory is not a good frame to be referred to cover the complete population for all types of studies because directories will not include new subscribers or those who have requested unlisted numbers. The class bias inherent in telephone directory. Hence, one should be very much careful while selecting a list for the population to be covered.

In practical situation we come across with various types of populations. But sampling techniques are the powerful tools available in Statistics to deal with any type of populations and arrive at valid conclusions about the mass of data.

Here’s a summary of our learning in this session:

- Statistical definition of population
- Importance of population
- Four types of population
- Difference between them with examples
- Target population and sample population
- Units and a Frame
- Importance of population in a sample theory