Course Title: Statistics for Economics
Course Code:ECON4008
Topic: Sampling Theory
M.A. Economics (2nd Semester)

Dr. Kailash Chandra Pradhan

Mahatma Gandhi Central University,
Department of Economics

Population

- Population is an aggregate of objects, animate or inanimate, under study in any statistical investigation. In sampling theory, the population means the larger group from which the samples are drawn.
- The population may be finite or infinite.
- A population containing a finite number of objects or items is known as finite population. For example: the students in a college, the day's production in an industrial concern, a population of a city or a town, etc.
- A population having an infinite number of objects is termed as an infinite population. For example: the population of temperatures at various point of thermosphere, the population of the heights, weights, of the people of the country (the values may be taken from particular intervals), the population of stars in the sky etc.

Sampling

- A finite subset of statistical individuals in a population is called a sample and the number of individuals in a sample is called the sample size.
- For the purpose of determining population characteristics, instead of gathering information of the entire population, the sample only can be observed. Then the sample characteristics are utilised to approximately determine or estimate the population.
- For example, on examining. the sample of a particular stuff we arrive at a decision of accepting or rejecting that stuff. The error involved in such approximation is known as sampling error and which is inherent and unavoidable in any and every sampling scheme.

Objectives of the Sampling theory

1. To obtain the optimum results, i.e. the maximum information about the characteristics of the population with the available sources at disposal in terms of time, money and manpower by studying the sample values only.
2. To obtain the best possible estimates of the population parameter.

Sampling

Characteristics of a Good Sample Design

- Sample design must result in a truly representative sample.
- Sample design must be such which results in a small sampling error.
- Sample design must be viable in the context of funds available for the research study.
- Sample design must be such so that systematic bias can be controlled in a better way.
- 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.

Types of Sampling

Some of the commonly known and frequently used types of sampling are:

1. Purposive sampling
2. Random sampling
3. Stratified sampling
4. Systematic sampling

Purposive/Subjective/Judgement Sampling:

- Purposive sampling is one in which the sample units are selected with definite purpose in view.
- For example, if we want to give the picture that the standard of living has increased in the city of New Delhi, we may take individuals in the sample from rich and posh localities like Defence Colony, South Exttension, Golf Links, Jor Bagh, Chanakyapuri, Greater Kailash etc. and ignore the localities where low income group and the middle class families live.
- This sampling suffers from the drawback of favouritism and nepotism and does not give a representative sample of the population.

Types of Sampling

Random Sampling:

- In this case the sample units are selected at random and the drawback of purposive sampling, viz. favouritism or subjective element, is completely overcome.
- A random sample is one in which each unit of population has an equal chance of being included in it.
- Suppose we take a sample of size n from a finite population of size N. Then there are ${}^{N}C_{n}$ possible samples.
- A sampling technique in which each of the ${}^{N}C_{n}$ samples has an equal chance of being selected is known as random sampling and the sample obtained by this technique is termed as a random sample.
- Proper care has to be taken to ensure that the selected sample is random.
- Human bias, which varies from individual to individual, is inherent in any sampling scheme administered by human beings.
- The simplest method, which is normally used, is the lottery system
- Suppose we want to select 'r' candidates out of n. We assign the numbers one to n, one number to each candidate and write these numbers (1 to n) on n slips which are made as homogeneous as possible in shape, size, etc. These slips are then put in a bag and thoroughly shuffled and then ‘r’ slips are drawn one by one. The ‘r’ candidates corresponding to the numbers on the slips drawn, will constitute the random sample.

Types of Sampling

Simple Random Sampling:

- Simple random sampling is random sampling in which each unit of the population has an equal chance, say p, of being included in the sample and that this probability is independent of the previous drawings.
- Thus a simple random sample of size n from a population may be identified with a series of n independent trials with constant probability 'p' of success for each trial.
- If the unit selected in any draw is not replaced in the population before making the next draw, then it is known as *simple random sampling without replacement (srswor)* and if it is replaced back before making the next draw, then the sampling plan is called *simple random sampling with replacement (srswr*).
- To ensure that sampling is simple random, it must be done with replacement, if population is finite. However, in case of infinite population no replacement is necessary.

Types of Sampling

Simple Random Sampling cntd…

- If a sample of size n is drawn without replacement from a population of size N then there are ${}^{N}C_{n}$ possible samples. Simple random sampling is the technique of selecting the sample so that each of these ${}^{N}C_{n}$ samples has an equal chance or probability

$$\frac{1}{{}^{N}C_{n}}$$

  of being selected in the sample.

- If sample is done with replacement, then there are N^n possible samples of size n. in this case, simple random sampling (srswr) gives equal chance

$$P = \frac{1}{N^n}$$

  of each of the N^n samples to be selected.

Types of Sampling

Stratified Random Sampling:

- Here the entire heterogeneous population is divided into a number of homogeneous groups usually termed as strata, which differ from one another but each of these groups is homogenous within itself.
- Then units are sampled at random from each of these stratum, the sample size in each stratum varies according to the relative importance of the stratum in the population.
- The sample, which is the aggregate of the sampled units of each of the stratum, is termed as stratified sample and the technique of drawing this sample is known as stratified sampling.

Allocation of Sample size in Stratified Sampling:

(i) Proportional Allocation: The method of proportional allocation under which the sizes of the samples from the different strata are kept proportional to the sizes of the strata. That is, if P_i represents the proportion of population included in stratum i, and n represents the total sample size, the number of elements selected from stratum i is n. Let us suppose that we want a sample of size n = 30 to be drawn from a population of size N = 8000 which is divided into three strata of size N_1 = 4000, N_2 = 2400 and N_3 = 1600. Adopting proportional allocation, we shall get the sample sizes as under for the different strata:

For strata with N_1 = 4000, we have P_1 = 4000/8000 and hence $n_1 = n \cdot P_1 = 30\ (4000/8000) = 15$

Similarly, for strata with N_2 = 2400, we have $n_2 = n \cdot P_2 = 30\ (2400/8000) = 9$, and for strata with N_3 = 1600, we have $n_3 = n \cdot P_3 = 30\ (1600/8000) = 6$

Types of Sampling

Proportional Allocation cntd….

- The sample sizes for different strata are 15, 9 and 6 respectively which is in proportion to the sizes of the strata viz., 4000 : 2400 : 1600.
- Proportional allocation is considered most efficient and an optimal design when the cost of selecting an item is equal for each stratum, there is no difference in within-stratum variances, and the purpose of sampling happens to be to estimate the population value of some characteristic.

(ii) Optimum Allocation

- In cases where strata differ not only in size but also in variability and it is considered reasonable to take larger samples from the more variable strata and smaller samples from the less variable strata,
- we can then account for both (differences in stratum size and differences in stratum variability) by using disproportionate sampling design by requiring:

$$n_1 / N_1\sigma_1 = n_2 / N_2\sigma_2 = \ldots\ldots\ldots = n_k / N_k\sigma_k$$

Types of Sampling

Optimum Allocation cntd….

- where s_1, s_2 , … and s_k denote the standard deviations of the *k* strata, N_1, N_2,…, N_k denote the sizes of the *k* strata and n_1, n_2,…, n_k denote the sample sizes of *k* strata. This is called *optimum allocation* in the context of disproportionate sampling.
- The allocation in such a situation results in the following formula for determining the sample sizes different strata:

$$n_i = \frac{n.N_i\sigma_i}{N_1\sigma_1 + N_2\sigma_2 + ... + N_k\sigma_k}$$

- In addition to differences in stratum size and differences in stratum variability, we may have differences in stratum sampling cost, then we can have cost optimal disproportionate sampling design by requiring

$$\frac{n_1}{N_1\sigma_1\sqrt{C_1}} = \frac{n_2}{N_2\sigma_2\sqrt{C_2}} = ... = \frac{n_k}{N_k\sigma_k\sqrt{C_k}}$$

Types of Sampling

- where

  C_1 = *Cost of sampling in stratum 1*

  C_2 = *Cost of sampling in stratum 2*

  C_k = *Cost of sampling in stratum k*

  and all other terms remain the same as explained earlier. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n.N_i\sigma_i / \sqrt{C_i}}{N_1\sigma_1 / \sqrt{C_1} + N_2\sigma_2 / \sqrt{C_2} + ... + N_k\sigma_k / \sqrt{C_k}}$$

(iii) Disproportionate Allocation:

- In this case an equal number of items is taken from every stratum regardless of how the stratum is represented in the population .
- Sometimes, the proportion may vary from stratum to stratum also.

Types of Sampling

- where

  C_1 = *Cost of sampling in stratum 1*

  C_2 = *Cost of sampling in stratum 2*

  C_k = *Cost of sampling in stratum k*

  and all other terms remain the same as explained earlier. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n.N_i\sigma_i / \sqrt{C_i}}{N_1\sigma_1 / \sqrt{C_1} + N_2\sigma_2 / \sqrt{C_2} + ... + N_k\sigma_k / \sqrt{C_k}}$$

(iii) Disproportionate Allocation:

- In this case an equal number of items is taken from every stratum regardless of how the stratum is represented in the population .
- Sometimes, the proportion may vary from stratum to stratum also.

Types of Sampling

Systematic Sampling:

- Systematic sampling is slight variation of the simple random sampling. In systematic sampling, only the first unit is selected randomly and the remaining units of the sample are selected at fixed intervals.
- Let us suppose that N sampling units in the population are arranged in some systematic order and serially numbered from 1 to N and we want to draw a sample of size n fro it such as

  N=nk or k=N/n

  where k is usually called sample interval
- Systematic sampling consists in selecting any unit at random from the first k units numbered from 1 to k and then selecting every kth unit in succession subsequently.

Cluster sampling:

- If the total area of interest happens to be a big one, a convenient way in which a sample can be taken is to divide the area into a number of smaller non-overlapping areas and then to randomly select a number of these smaller areas (usually called clusters).
- In the cluster sampling, the total population is divided into a number of relatively small subdivisions which are themselves clusters of still smaller units and then some of these clusters are randomly selected for inclusion in the overall sample.

Area sampling:

- If clusters happen to be some geographic subdivisions, in that case cluster sampling is better known as area sampling.

Theory of Sampling

Reference:

- Gupta, S. C. (2015), *Fundamentals of Statistics*, Himalaya Publishing House.
- Kothari, C.R.,1985, Research Methodology-Methods and Techniques, New Delhi, Wiley Eastern Limited