## UNIT-II

| Unit-II | Sampling Methods |
| :---: | :--- |
| 2.1 | Sampling Methods |
| 2.2 | Advantages and Limitations |
| 2.3 | Sampling and Non-Sampling Errors |
| 2.4 | Random sampling methods |
| 2.5 | Simple Random Sampling |
| 2.6 | Systematic Sampling |
| 2.7 | Stratified Sampling |
| 2.8 | Non-Radom sampling methods |
|  | - (No Derivations, Only Concepts). |

### 2.1 Sampling Methods

Sampling is very often used in our daily life. For example, while purchasing food grains from a shop we usually examine a handful from the bag to assess the quality of the commodity. A doctor examines a few drops of blood as sample and draws conclusion about the blood constitution of the whole body. Thus, most of our investigations are based on samples. Let us see the importance of sampling and the various methods of sample selections from
 the population.

## Population:

In a statistical study, the set of all items, which are under the study or under consideration, is known as Population or Universe. In other words, the population is a complete set of all possible observations of the type which is to be investigated.

Examples of population: -
> Set of all students studying in a school or college,
> Set of all books in a library,
> All houses in a village or town of population.

## Complete Enumeration, or Census Study:

Sometimes it is possible and practical to examine every person or item in the population we wish to describe. We call this a Complete enumeration, or Census Study. We use sampling when it is not possible to measure every item in the population. Statisticians use the word
population to refer not only to people. Finally, the population is set of all items/units that have been chosen for the study.

## Finite population and infinite population:

A population is said to be finite if it consists of finite number of units. Eg., number of workers in a factory, production of articles in a particular day for a company are examples of finite population. The total number of units in a population is called population size.

A population is said to be infinite if it has infinite number of units. For example, the number of stars in the sky, the number of people seeing the Television programmes etc.,

## Census Method:

Information on population can be collected in two ways - census method and sample method. In census method, every element of the population is included in the investigation. For example, if we study the average annual income of the families of a particular village or area, and if there are 1000 families in that area, we must study the income of all 1000 families. In this method no family is left out, as each family is a unit.

Population census of India:
The population census of our country is taken at 10 yearly intervals. The latest census was taken in 2011. The first census was taken in 1871-72.

POPULATION OF INDIA 2001

| India/State/ <br> Union <br> territories* | POPULATION OF INDIA 2001 |  | Population | Sex ratio <br> (females per <br> (hariationsand <br> males) |  |
| :--- | ---: | ---: | ---: | :---: | :---: |
|  | PERSONS | MALES | FEMALES |  |  |
| INDIA 1,2 | $1,027,015,247$ | $531,277,078$ | $495,738,169$ | 21.34 | 933 |
|  <br> Nicobar Is.* | 356,265 | 192,985 | 163,280 | 26.94 | 846 |
| Andhra Pradesh | $75,727,541$ | $38,286,811$ | $37,440,730$ | 13.86 | 978 |
| Arunachal <br> Pradesh | $1,091,117$ | 573,951 | 517,166 | 26.21 | 901 |
| Assam | $26,638,407$ | $13,787,799$ | $12,850,608$ | 18.85 | 932 |
| Bihar | $82,878,796$ | $43,153,964$ | $39,724,832$ | 28.43 | 921 |
| Chandigarh* | 900,914 | 508,224 | 392,690 | 40.33 | 773 |
| Chhatisgarh | $20,795,956$ | $10,452,426$ | $10,343,530$ | 18.06 | 990 |
| Dadra \& Nagar <br> Haveli* | 220,451 | 121,731 | 98,720 | 59.20 | 811 |
| Daman \& Diu* | 158,059 | 92,478 | 65,581 | 55.59 | 709 |
| Delhi* | $13,782,976$ | $7,570,890$ | $6,212,086$ | 46.31 | 821 |


| Goa | $1,343,998$ | 685,617 | 658,381 | 14.89 | 960 |
| :--- | ---: | ---: | ---: | :---: | :---: |
| Gujarat 5 | $50,596,992$ | $26,344,053$ | $24,252,939$ | 22.48 | 921 |
| Haryana | $21,082,989$ | $11,327,658$ | $9,755,331$ | 28.06 | 861 |
| Himachal <br> Pradesh 4 | $6,077,248$ | $3,085,256$ | $2,991,992$ | 17.53 | 970 |
|  <br> Kashmir 2, 3 | $10,069,917$ | $5,300,574$ | $4,769,343$ | 29.04 | 900 |
| Jharkhand | $26,909,428$ | $13,861,277$ | $13,048,151$ | 23.19 | 941 |
| Karnataka | $52,733,958$ | $26,856,343$ | $25,877,615$ | 17.25 | 964 |
| Kerala | $31,838,619$ | $15,468,664$ | $16,369,955$ | 9.42 | 1,058 |
| Lakshadweep* | 60,595 | 31,118 | 29,477 | 17.19 | 947 |
| Madhya Pradesh | $60,385,118$ | $31,456,873$ | $28,928,245$ | 24.34 | 920 |
| Maharashtra | $96,752,247$ | $50,334,270$ | $46,417,977$ | 22.57 | 922 |
| Manipur | $2,388,634$ | $1,207,338$ | $1,181,296$ | 30.02 | 978 |
| Meghalaya | $2,306,069$ | $1,167,840$ | $1,138,229$ | 29.94 | 975 |
| Mizoram | 891,058 | 459,783 | 431,275 | 29.18 | 938 |
| Nagaland | $1,988,636$ | $1,041,686$ | 946,950 | 64.41 | 909 |
| Orissa | $36,706,920$ | $18,612,340$ | $18,094,580$ | 15.94 | 972 |
| Pondicherry* | 973,829 | 486,705 | 487,124 | 20.56 | 1,001 |
| Punjab | $24,289,296$ | $12,963,362$ | $11,325,934$ | 19.76 | 874 |
| Rajasthan | $56,473,122$ | $29,381,657$ | $27,091,465$ | 28.33 | 922 |
| Sikkim | 540,493 | 288,217 | 252,276 | 32.98 | 875 |
| Tamil Nadu | $62,110,839$ | $31,268,654$ | $30,842,185$ | 11.19 | 986 |
| Tripura | $3,191,168$ | $1,636,138$ | $1,555,030$ | 15.74 | 950 |
| Uttar Pradesh | $166,052,859$ | $87,466,301$ | $78,586,558$ | 25.80 | 898 |
| Uttaranchal | $8,479,562$ | $4,316,401$ | $4,163,161$ | 19.20 | 964 |
| West Bengal | $80,221,171$ | $41,487,694$ | $38,733,477$ | 17.84 | 934 |

## Notes:

1. The population of India includes the estimated population of entire Kachchh district, Morvi, Maliya-Miyana and Wankaner talukas of Rajkot district, Jodiya taluka of Jamanagar district of Gujarat State and entire Kinnaur district of Himachal Pradesh where population enumeration of Census of India 2001 could not be conducted due to natural calamity.
2. For working out density of India, the entire area and population of those portions of Jammu and Kashmir which are under illegal occupation of Pakistan and China have not been taken into account.
3. Figures shown against Population in the age-group 0-6 and Literates do not include the figures of entire Kachchh district, Morvi, Maliya-Miyana and Wankaner talukas of

Rajkot district, Jodiya taluka of Jamanagar district and entire Kinnaur district of Himachal Pradesh where population enumeration of Census of India 2001 could not be conducted due to natural calamity.
4. Figures shown against Himachal Pradesh have been arrived at after including the estimated figures of entire Kinnaur district of Himachal Pradesh where the population enumeration of Census of India 2001 could not be conducted due to natural calamity.
5. Figures shown against Gujarat have been arrived at after including the estimated figures of entire Kachchh district, Morvi, Maliya-Miyana and Wankaner talukas of Rajkot district, Jodiya taluka of Jamnagar district of Gujarat State where the population enumeration of Census of India 2001 could not be conducted due to natural calamity.

## Merits and limitations of Census method:

## Merits:

- The data are collected from each and every item/unit of the population.
- The results are more accurate and reliable, because every item of the universe is required.
- Intensive study is possible.
- The data collected may be used for various surveys, analyses etc.


## Limitations:

- It requires a large number of enumerators and it is a costly method
- It requires more money, labour, time, energy etc.
- It is not possible in some circumstances where the universe is infinite.


## Sampling:

The theory of sampling has been developed recently but this is not new. In our everyday life, we have been using sampling theory as we have discussed in introduction. In all those cases, we believe that the samples give a correct idea about the population. Most of our decisions are based on the examination of a few items only, ie., sample studies.

Sample:
A sample is a portion chosen from the population. A finite subset of statistical individuals defined in a population is called a sample. The number of units in a sample is called the sample size.

Sampling units:
The constituents of a population which are individuals to be sampled from the population and cannot be further subdivided for the purpose of the sampling at a time are called sampling units. For example, to know the average income per family, the head of the family is a sampling unit. To know the average yield of rice, each farm owner's yield of rice is a sampling unit.

Sampling frame:
For adopting any sampling procedure, it is essential to have a list identifying each
sampling unit by a number. Such a list or map is called sampling frame. A list of voters, a list of house holders, a list of villages in a district, a list of farmers etc. are a few examples of sampling frame.
Reasons for selecting a sample:
Sampling is inevitable in the following situations:

- Complete enumerations are practically impossible when the population is infinite.
- When the results are required in a short time.
- When the area of survey is wide.
- When resources for survey are limited particularly in respect of money and trained persons.
- When the item or unit is destroyed under investigation.


## Parameters and Statistics:

We can describe samples and populations by using measures such as the mean, median, mode and standard deviation. When these terms describe the characteristics of a population, they are called Parameters. When they describe the characteristics of a sample, they are called Statistics. A parameter is a characteristic of a population and a statistic is a characteristic of a sample. Since samples are subsets of population statistics provide estimates of the parameters. That is, when the parameters are unknown, they are estimated from the values of the statistics.

In general, we use Greek or capital letters for population parameters and lower case letters to denote sample statistics.

- $\mathrm{N}, \mu, \sigma$ are the standard symbols for the size, mean, S.D, of population.
- $n, x$ bar, $s$ are the standard symbol for the size, mean, S.D of sample respectively.


## Principles of Sampling:

Samples have to provide good estimates for the Population Parameters. The following principles tell us that the sample methods provide such good estimates.
i. Principle of statistical regularity: A moderately large number of units chosen at random from a large group are almost sure on the average to possess the characteristics of the large group.
ii. Principle of Inertia of large numbers: Other things being equal, as the sample size increases, the results tend to be more accurate and reliable.
iii. Principle of Validity: This states that the sampling methods provide valid estimates about the population parameters.
iv. Principle of Optimisation: This principle takes into account the desirability of obtaining a sampling design which gives optimum results. This minimizes the risk or loss of the sampling design. The foremost purpose of sampling is to gather maximum information about the population under consideration at minimum cost, time and human power. This is bestachieved when the sample contains all the properties of the population.

### 2.2 Advantages and Limitations of Sampling:

There are many advantages of sampling methods over census method.

1. Sampling saves time and labour.
2. It results in reduction of cost in terms of money and manhour.
3. Sampling ends up with greater accuracy of results.
4. It has greater scope.
5. It has greater adaptability.
6. If the population is too large, or hypothetical or destroyable, the sampling is the only method to be used.

The limitations of sampling are given below:

1. Sampling is to be done by qualified and experienced persons. Otherwise, the information will be unbelievable.
2. Sampling method may give the extreme values sometimes instead of the mixed values.
3. There is the possibility of sampling errors. Census survey is free from sampling error.

### 2.4 Sampling errors and non-sampling errors:

There are two types of errors in a sample survey.
They are sampling errors and non - sampling errors.
i. Sampling errors: Although a sample is a part of population, it cannot be expected generally to supply full information about population. So, there may be in most cases difference between Statistics and Parameters. The discrepancy between a parameter and its estimate due to sampling process is known as sampling error.
ii. Non-sampling errors: In all surveys, some errors may occur during collection of actual information. These errors, other than sampling errors, are called Non-sampling errors.

## Types of Sampling:

The technique of selecting a sample is of fundamental importance in sampling theory and it depends upon the nature of investigation. The sampling methods (procedures) which are commonly used may be classified as

- Probability sampling.
- Non-probability sampling.
- Mixed sampling.

Probability sampling (Random sampling):
A probability sample is one where the selection of units from the population is made according to known probabilities. (eg.) Simple random sample, probability proportional to sample size etc.

## Non-Probability Sampling:

It is the one where discretion is used to select 'representative' units from the population (or) to infer that a sample is 'representative' of the population. This method is called judgement or purposive sampling. This method is mainly used for opinion surveys; A common type of judgement sample used in surveys is quota sample. This method is not used in general because of prejudice and bias of the enumerator. However, if the enumerator is experienced and expert, this method may yield valuable results. For example, in a market research survey of the performance of the cars, suppose the sample has all new car purchasers.

## Mixed Sampling:

Here samples are selected partly according to some probability and partly according to a fixed sampling rule; they are termed as mixed samples and the technique of selecting such samples is known as mixed sampling.

### 2.4 Probability (Random) Sampling Methods of selection of samples:

There are several methods of random sampling, of which let us study one three methods:
a) Simple random sampling.
b) Stratified random sampling.
c) Systematic random sampling.

a) Simple Random Sampling:

Simple random sampling is the sampling method used to select a random sample from a finite population under study such that each possible sample unit has an equal probability of being chosen. It is also called unrestricted random sampling.

There are two approaches of Simple Random Sampling (SRS) and are

- Without Replacement (SRS WOR) and
- With Replacement (SRS WR).
i. Simple random sampling without replacement:

In this method, the population elements can enter the sample only once (ie.,) the unit once selected is not returned to the population before the next draw.
ii. Simple random sampling with replacement:

In this method, the population units may enter the sample more than once. (ie.,) the units once selected is returned back to the population before the next draw.

### 2.5 Simple Random Sampling

Methods of selection of a simple random sampling:
The following are some methods of selection of a simple random sampling.
a) Lottery Method:

This is the most popular and simplest method. In this method all the items of the population are numbered on separate slips of paper of same size, shape and colour. They are folded and mixed up in a container. The required numbers of slips are selected at random for the desire sample size. For example, if we want to select 5 students, out of 50 students, then we must write their names or their roll numbers of all the 50 students on slips and mix them. Then we make a random selection of 5 students. This method is mostly used in lottery draws. If the universe is infinite this method is inapplicable.
b) Table of Random numbers:

As the lottery method cannot be used, when the population is infinite, the alternative method is that of using the table of random numbers. There are several standard tables of random numbers of which commonly used three tables are:
i. Tippett' s table
ii. Fisher and Yates' table
iii. Kendall and Smith' s table

A random number table is so constructed that all digits 0 to 9 appear independent of each other with equal frequency. If we have to select a sample from population of size $\mathrm{N}=100$, then the numbers can be combined three by three to give the numbers from 001 to 100 .

Random Number Table

|  |  |  | RAN <br> DOM <br> NUM <br> BERS |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 4652 | 3819 | 8431 | 2150 | 2352 | 2472 | 0043 | 3488 |
| 9031 | 7617 | 1220 | 4129 | 7148 | 1943 | 4890 | 1749 |

18BGE24A: Allied: Statistics-II UNIT-II
I BSc Geography (English Medium)

| 2030 | 2327 | 7353 | 6007 | 9410 | 9179 | 2722 | 8445 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0641 | 1489 | 0828 | 0385 | 8488 | 0422 | 7209 | 4950 |
| 8479 | 6062 | 5593 | 6322 | 9439 | 4996 | 1322 | 4918 |
| 9917 | 3490 | 5533 | 2577 | 4348 | 0971 | 2580 | 1943 |
| 6376 | 9899 | 9259 | 5117 | 1336 | 0146 | 0680 | 4052 |
| 7287 | 0983 | 3236 | 3252 | 0277 | 8001 | 6058 | 4501 |
| 0592 | 4912 | 3457 | 8773 | 5146 | 2519 | 3931 | 6794 |
| 6499 | 9118 | 3711 | 8838 | 0691 | 1425 | 7768 | 9544 |
| 0769 | 1109 | 7909 | 4528 | 8772 | 1876 | 2113 | 4781 |
| 8678 | 4873 | 2061 | 1835 | 0954 | 5026 | 2967 | 6560 |
| 0178 | 7794 | 6488 | 7364 | 4094 | 1649 | 2284 | 7753 |
| 3392 | 0963 | 6364 | 5762 | 0322 | 2592 | 3452 | 9002 |
| 0264 | 6009 | 1311 | 5873 | 5926 | 8597 | 9051 | 8995 |
| 4089 | 7732 | 8163 | 2798 | 1984 | 1292 | 0041 | 2500 |
| 9376 | 7365 | 7987 | 1937 | 2251 | 3411 | 6737 | 0367 |
| 3039 | 3780 | 2137 | 7641 | 4030 | 1604 | 2517 | 9211 |
| 8971 | 8653 | 1855 | 5285 | 5631 | 2649 | 6696 | 5475 |
| 0375 | 4153 | 5199 | 5765 | 2067 | 6627 | 3100 | 5716 |
| 9092 | 4773 | 0002 | 7000 | 7800 | 2292 | 2933 | 6125 |
| 2464 | 1038 | 3163 | 3569 | 7155 | 2029 | 2538 | 7080 |
| 3027 | 6215 | 3125 | 5856 | 9543 | 3660 | 0255 | 5544 |
| 5754 | 9247 | 1164 | 3283 | 1865 | 5274 | 5471 | 1346 |
| 4358 | 3716 | 6949 | 8502 | 1573 | 5763 | 5046 | 7135 |
| 7178 | 8324 | 8379 | 7365 | 4577 | 4864 | 0629 | 5100 |
| 5035 | 5939 | 3665 | 2160 | 6700 | 7249 | 1738 | 2721 |
| 3318 | 0220 | 3611 | 9887 | 4608 | 8664 | 2185 | 7290 |
| 9058 | 1735 | 7435 | 6822 | 6622 | 8286 | 8901 | 5534 |
| 7886 | 5182 | 7595 | 0305 | 4903 | 3306 | 8088 | 3899 |
| 3354 | 8454 | 7386 | 1333 | 5345 | 6565 | 3159 | 3991 |
| 3415 | 7671 | 0846 | 7100 | 1790 | 9449 | 6285 | 2525 |
| 3918 | 5872 | 7898 | 6125 | 2268 | 1898 | 0755 | 6034 |
| 6138 | 9045 | 6950 | 8843 | 6533 | 0917 | 6673 | 5721 |
| 3828 | 1704 | 2835 | 4677 | 4637 | 7329 | 3156 | 3291 |
| 1349 | 0417 | 9311 | 9787 | 1284 | 0769 | 8422 | 1077 |
| 4234 | 0248 | 7760 | 6504 | 2754 | 4044 | 0842 | 9080 |
| 6880 | 3201 | 7044 | 3657 | 5263 | 0374 | 7563 | 6599 |
| 0714 | 5008 | 5076 | 1134 | 5342 | 1608 | 5179 | 0967 |
| 3448 | 6421 | 3304 | 0583 | 1260 | 0662 | 7257 | 0766 |
| 5711 | 7373 | 7539 | 3684 | 9397 | 5335 | 4031 | 1486 |
| 2588 | 3301 | 0553 | 2427 | 3598 | 2580 | 7017 | 9176 |
| 8581 | 4253 | 7404 | 5264 | 5411 | 3431 | 3092 | 8573 |
| 8475 | 6322 | 3949 | 9675 | 6533 | 1133 | 8776 | 2216 |
| 0272 | 5624 | 8549 | 5552 | 7469 | 2799 | 2882 | 9620 |
| 7383 | 7795 | 7939 | 2652 | 4456 | 6993 | 2950 | 8573 |

Procedure to select a sample using random number table:
Units of the population from which a sample is required are assigned with equal number of digits. When the size of the population is less than thousand, three digit number $000,001,002$,
..... 999 are assigned. We may start at any place and may go on in any direction such as column wise or row- wise in a random number table. But consecutive numbers are to be used.

On the basis of the size of the population and the random number table available with us, we proceed according to our convenience. If any random number is greater than the population size N , then N can be subtracted from the random number drawn. This can be repeatedly until the number is less than N or equal to N .

Example 1:
In an area there are 500 families. Using the following extract from a table of random numbers select a sample of 15 families to find out the standard of living of those families in that area.

| 4652 | 3819 | 8431 | 2150 | 2352 | 2472 | 0043 | 3488 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 9031 | 7617 | 1220 | 4129 | 7148 | 1943 | 4890 | 1749 |
| 2030 | 2327 | 7353 | 6007 | 9410 | 9179 | 2722 | 8445 |
| 0641 | 1489 | 0828 | 0385 | 8488 | 0422 | 7209 | 4950 |

## Solution:

In the above random number table, we can start from any row or column and read three digit numbers continuously row-wise or column wise.

Now we start from the third row, the numbers are:

| 203 | 023 | 277 | 353 | 600 | 794 | 109 | 179 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 272 | 284 | 450 | 641 | 148 | 908 | 280 |  |

Since some numbers are greater than 500 , we subtract 500 from those numbers and we rewrite the selected numbers as follows:

| 203 | 023 | 277 | 353 | 100 | 294 | 109 | 179 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 272 | 284 | 450 | 141 | 148 | 408 | 280 |  |

c) Random number selections using calculators or computers:

Random number can be generated through scientific calculator or computers. For each press of the key get a new random number. The way of selection of sample is similar to that of using random number table.

## Merits of using random numbers:

1. Personal bias is eliminated as a selection depends solely on chance.
2. A random sample is in general a representative sample for a homogenous population.
3. There is no need for the thorough knowledge of the units of the population.
4. The accuracy of a sample can be tested by examining another sample from the same universe when the universe is unknown.
5. This method is also used in other methods of sampling.

## Limitations:

1. Preparing lots or using random number tables is tedious when the population is large.
2. When there is large difference between the units of population, the simple random sampling may not be a representative sample.
3. The size of the sample required under this method is more than that required by stratified random sampling.
4. It is generally seen that the units of a simple random sample lie apart geographically. The cost and time of collection of data are more.

### 2.6 Stratified Random Sampling

In Stratified random sampling, the heterogeneous population of size $N$ units is subdivided into $L$ homogeneous non overlapping sub populations called Strata, the $i^{\text {th }}$ stratum having $N_{\mathrm{i}}$ units ( $i=1,2,3, \ldots, \mathrm{~L}$ ) such that $N_{1}+N_{2}+\ldots+N_{L}=N$.


A sample size being $n_{i}$ from $i^{\text {th }}$ stratum $(i=1,2, \ldots, L)$ is independently taken by Simple Random Sampling in such way that $n_{1}+n_{2}+\ldots+n_{L}=n$. A sample obtained using this procedure is called a Stratified Random Sample.

To determine the sample size for each stratum, there are two methods:
i)Proportionate Allocation: In proportionate allocation, sample size is determined as proportionate to stratum size. If the stratum size is large, that stratumivill get more representation in the sample. If the stratum size is small, that stratum will get less representation in the sample.

The sample size for the $\mathrm{i}^{\text {th }}$ stratum can be determined using the formula $n_{i}=(n / N) * N_{i}$. If we put $\mathrm{i}=1,2,3, \ldots$ L , we get $n_{l}=(n / N) * N_{l}$,

$$
n_{2}=(n / N) * N_{2}, \ldots \ldots \ldots n_{L}=(n / N) * N_{L} .
$$

ii) Optimum Allocation: The optimum allocation method uses variation in the stratum and cost to determine the stratum sample size $n_{i}$.

Example 1: To study about the introduction of NEET exam, the opinions are collected from 3 schools. The strength of the schools are 2000, 2500 and 4000. It is fixed that the sample size is 170 . Calculate the sample size for each school?

## Solution

Here $N=2000+2500+4000=8500$ and $n=170$ then $n_{1}=n_{2}=n_{13}=$ ?
$N_{1}=2000, \quad N_{2}=2500, \quad N_{3}=4000$
School 1: $n_{1}=(n / N) \times N_{1}=(170 / 8500) \times 2000=40$
School 2: $n_{2}=(n / N) \times N_{2}=(170 / 8500) \times 2500=50$
School 3: $n_{3}=(n / N) \times N_{3}=(170 / 8500) \times 4000=80$
Therefore, 40 students from school 1 ;
50 students from school 2 and
80 students from school 3
are to be selected using Simple Random Sampling to obtain the required stratified random sample.

The main objective of stratification is to give a better cross-section of the population for a higher degree of relative precision. The criteria used for stratification are States, age and sex, academic ability, marital status etc,. In many practical situations, when it is difficult to stratify with respect to the characteristic under study, administrative convenience may be considered as the basis for stratification.

## Merits

- It provides a chance to study of all the sub-populations separately.
- An optimum size of the sample can be determined with a given cost, precision and reliability.
- It is a more precise sample.
- Representation of sub groups in the population.
- Biases reduced and greater precise.


## Limitations

- There is a possibility of faulty stratification and hence the accuracy may be lost.
- Proportionate stratification requires accurate information on the proportion of population in each stratum.


### 2.7 Systematic Sampling

In systematic sampling, the population units are numbered from 1 to $N$ in ascending order. A sampling interval, denoted by $\boldsymbol{k}$, is determined as $k=\frac{N}{n}$, 四here $n$ denotes the required sample size. Then $n-1$ such sampling intervals each consisting of $k$ units will be formed. A number is selected at random from the first sampling interval. Let it be number ' i ' where $i \leq k$. This number is the random starting point for the whole selection of the sample. The unit corresponding to $i$ is the first unit in the sample. The subsequent sampling units are the units in the following positions:

$$
i, k+i, 2 k+i, 3 k+i, \ldots,(n-1) k+i .
$$

| The layout for systematic sampling |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | ---: | ---: | ---: | :---: | :---: | :---: |
| Sampling interval 1 | 1 | 2 | 3 | 4 | $i$ - Random <br> start |  | k |  |
| Sampling interval 2 | $k+1$ | $k+2$ | $k+3$ | $k+4$ | .. | $k+i$ | $\ldots$ | 2 k |
| Sampling interval 3 | $2 k+1$ | $2 k+2$ | $2 k+3$ | $2 k+4$ | .. | $2 k+i$ | $\ldots$ | 3 k |
| Sampling interval 4 | $3 k+1$ | $3 k+2$ | $3 k+3$ | $3 k+4$ | .. | $3 k+i$ | $\ldots$ | $4 k$ |
|  | $\ldots$ | $\ldots$ | $\ldots$ | $\ldots$ | . | $\ldots$ | $\ldots$ | $\ldots$ |
| Sampling interval $n$ | $(n-1) k+1$ | $(n-1) k+2$ | $(n-1) k+3$ | $(n-1) k+4$ | . | $(n-1) k+i$ | $\cdots$ | $n k$ |

Thus, with selection (random) of the first unit, the whole sample is selected automatically. As the first unit could have been any of the $k$ units, the technique will generate ' $k$ ' systematic sample units with equal probability. If $N$ is not an integral multiple of $n$, then sizes of a few possible systematic samples may vary by one unit.

## Note:

This sample is also called quasi random sample, since, the first unit only is selected at random and all the subsequent units are not selected randomly.

## Merits:

- This method is simple and convenient.
- Less time consuming.
- It can be used in infinite population.


## Limitation:

- Since, it is a quasi random sampling, the sample may not be a true representative sample of the population under study.


## Example 2

Suppose a systematic random sample of size $n=10$ is needed from a population of size $N=200$, the sampling interval $k=\frac{N}{n}=\frac{200}{10}=20$.

The first sampling interval consists of numbers 1 to 20 . If the randomly selected number (random starter) is 7 , the systematic sample will consist units corresponding to positions (serial numbers) : 7, 27, 47, 67, 87, 107, 127, 147, 167, 187.

## Applications:

- Systematic sampling is preferably used when the information is to be collected from trees in a highway, houses in blocks, etc,.
- This method is often used in industry, where an item is selected for testing from a production line (say, every fifteenth item in the order of production) to ensure that equipments are working satisfactorily.
- This technique could also be used in a sample survey for interviewing people. A market researcher might select every 10th person who enters a particular store, after selecting a person at random as a random start.


### 2.8 Non-probability (Random) Sampling:

Non-probability sampling is the sampling procedure in which samples are selected based on the subjective judgment of the researcher, rather than random selection. This method isused when the representativeness of the population is not the prime issue. Convenience or judgments of the investigators play an important role in selecting the samples. In general, there are four types of non probability sampling called convenience sampling, judgment sampling, quota sampling and snowball sampling.


### 2.8.1 Convenience Sampling:

The samples are drawn at the convenience of the investigator. The investigator pickup cases which are easily available units, not keeping the objectives in mind for the study.

## Merits:

- Useful for pilot study.
- Use the results that are easily available.
- Processes of picking people in the most convenient and faster way to immediately get their reactions to a certain hot and controversial topic.
- Minimum time needed and minimum cost incurs.


## Limitations:

- High risk of selection bias.
- May provide misleading information.
- Not representative sample. Errors occur in the form of the members of the population who are infrequent or non-users of that location and who are not.

Note: The use of convenience sampling technique is discouraged by many researchers dueto inability to generalize research findings, the relevance of bias and high sampling error. Nevertheless, convenience sampling may be the only option available in certain situations. For example, "it may be that a researcher intends to study the 'customer satisfaction of Jet Airways' he uses the convenient sampling because he has been able to negotiate access through available contacts".

### 2.8.2 Snowball Sampling:

In this type, initial group of respondents are selected. Those respondents are requested to provide the names of additional respondents who belong to the target population of interest. It is a sampling method that involves the assistance of study subjects to identify other potential subjects in studies where subjects are hard to locate such as sex workers, drug abusers, etc. This type of sampling technique works like a chain referral. Therefore, it is also called chain referral sampling.

## Merits:

- Appropriate for small specialized population.
- Useful in studies involving respondents rare to find.


## Limitations:

- It takes more time
- Most likely not representative
- Members of the population, who are little known, disliked or whose opinions conflict with the respondents, have low probability of being included.


### 2.8.3 Judgement Sampling.

The investigator believes that in his opinion, some objects are the best representative of the population than others. It involves "hand picking" of sampling units. That is the interviewer uses his judgment in the selection of the sample that who should represent the population to serve the investigator's purpose. It is usually used when a limited number of individuals possess the trait of interest. This type of sampling is also known as purposive sampling. This is useful when selecting specific people, specific events, specific prices of data, etc. For example, selecting members for a competition like quiz, oratorical contest to represent a school.

## Merits:

- Low expense.
- Minimum time needed.
- Easy


## Limitations:

- Highly subjective.
- Generalization is not appropriate.
- Certain members of the population will have a smaller chance or no chance of selection compared to others.
- This method does not give representative part of the population, since favoritism is involved.


### 2.8.4 Quota Sampling

This is another non-probability sampling method. In this method, the population is divided into different groups and the interviewer assign quotas to each group. The selection of individuals from each group is based on the judgment of the interviewer. This type of sampling is called quota sampling. Specified sizes of number of certain types of people are included in the sample.

## Merits:

- The selection of the sample in this method is quick, easy and cheaper.
- May control sample characteristics.
- More chance of representative.


## Limitations:

- Selection bias.
- The sample is not a true representative and statistical properties cannot beapplied.


## Example

A selection committee wants to compose a cricket team (11 players) for a test match.

| Groups | Pace bowlers | Spinners | All-rounders | Batsmen | Wicket keepers |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | Players in this category | Players in this category | Players in this category | Players in this category | Players in this category |
| Quota | 2 | 3 | 3 | 2 | 1 |
| Select the players using judgment sampling to fulfill the requirement of the respective quota. |  |  |  |  |  |

In the composition of a cricket team, the selection committee forms groups compartmentalize as pace bowlers, spinners, all-rounders, batsmen and wicket keepers.The committee fixed quota for each group based on the pitch and the opponent teams' strength. Then, from each group they select the required number of players using judgement.

