# UNIT 12 SIMPLE RANDOM SAMPLING AND SYSTEMATIC SAMPLING

**Structure** **Page No.**



## 12.1 INTRODUCTION

In our day to day routine we often have to make certain judgements about a large bulk or population after studying a small portion, or a sample of it. For example, a house wife tastes a spoonful of soup to see whether a little more salt is required before it is served to guests; a quality control inspector inspects a sample from a tin of oil before passing the whole tin as acceptable quality; or a doctor takes a few drops of blood from a patient to decide if the patient has malarial infection. These are typical examples where it is not practical to examine the entire lot or population and decision making is done on the basis of sample information. Essentially this is sampling. It saves time and if you have based your judgement on judicious observations, it serves your purpose. Think of situations, when decisions are to be taken at macro levels, say, national level. Sampling has become an effective tool for generating information for policy formulation at different administrative levels. The above examples have the special phenomenon that a spoon of soup, tinful of oil and the blood in the patient are known to be perfectly **homogeneous** material so that every part of the material represents the material exactly. Often, however, we are not in this simple situation. For example, suppose we are interested in knowing the average weight of an adult Indian male. Obviously, it will not be satisfactory to measure just a few adult males, as not all adult males are of the same weight. They show considerable **hetereogeneity**. So how do we take a part of a large heterogeneous mass to draw valid conclusions from it? However, careful considerations are needed for selection of samples and making valid inferences from these samples. In the subject of 'sampling', these considerations and criteria are developed on a scientific basis.

In this unit, we shall start by introducing the basic concepts of sampling. We shall then discuss the simplest procedure for sample collection i.e., simple random sampling (SRS). In the process you will get introduced to the concept of random numbers, simple random sampling as a method of selection, estimation as well as considerations for determination of sample size.

Further, the concept of systematic sampling, which is also a method of selection based on random sampling procedure, will be introduced alongwith the estimation approach.

**Objectives**

After reading this unit, you should be able to

- select samples using simple random sampling and systematic sampling procedures;
- estimate population mean, population variance, proportion, alongwith their efficiency in SRS;
- select samples using systematic sampling by linear and circular systematic procedures.

## 12.2 SAMPLING – WHAT AND WHY?

You have already learnt in Block 1, the concept of a statistical population. Often we are interested in studying a specified characteristics of the individuals in a finite population. For example, we may be interested in studying the annual income and size of the households in Delhi for the year 2000-2001. In this case our population is the collection of all households in Delhi during the year 2000-01 and the individuals are the households. It may be enough for our purpose to find out the average size and annual income of the house-holds in Delhi. How do we go about obtaining this information? One way is to visit each household, note down the number of members and find out from the members how much was their total income during the year 2000-01 and calculate a simple average of all the figures obtained. You can very well imagine the difficulties and expenses involved in undertaking this huge task. The task of enumerating all the households in Delhi, visiting them and noting their sizes and annual income will be termed a census of the population under study. As you may all be aware, such a census is held once in ten years and information is obtained on a large number of characteristics including the ones mentioned above. Given this complexity and the huge expenses involved, this method cannot be adopted every time we need information on a population. Here are a few more instances where we may need to collect information on a characteristic from all the individuals of a population:

i) A telephone company may be interested in figuring out the average number of calls and average duration of a call made by the households in a locality or a city
ii) A large fruit store may be interested in the quality of the truck load of peaches packed in crates received at the stores from a farm
iii) A company producing electric bulbs may be interested in the average life of bulbs produced by them during a shift.

E1) In the cases mentioned above identify the population and the individuals.

You may notice that observing or inspecting all the individuals in a population can be very expensive in terms of time and money. In certain instances, like observing the life of bulbs, inspection may be destructive, so that observing all the bulbs of a population for calculating their average life is a meaningless exercise. so an alternative approach is called for. Can you think of an alternative? One alternative is to observe or measure only some individuals of a population, but estimate the average for the whole population based on the few measurements made. The idea is roughly as follows: Suppose you want to find the total number of mangoes on a tree. First, it is easy to count the number of mangoes on a "typical" branch. Multiply the number obtained by the number of branches and you get an estimate. Do you think this method will yield a good enough estimate? The method adopted in this approach is called the **method of sampling**. You may realise that in this methodology your final estimate of the total may depend on the number of fruits on the selected branch. Of course, if all the branches happen to bear the same number of fruits, it does not matter which branch you happen to select. You may therefore want to know if all the branches bear nearly the same number of fruits or, if they differ then by how many they may differ. In order to do this, again using the principle of sampling, you may select another branch to get an idea of the difference. If you select the branches with a priori known probabilities, (for example, by ensuring that any set of two branches have the same probability of being selected in the sample as any other set), you may be able to use the theory of probability and statistical inference you have learnt in the earlier blocks to calculate the likely error in your estimate. The theory and method of sampling deal with the issue of how to select the sample, how to estimate the population total or average and how to estimate the error in observing only a sample instead of the entire population. The error that results from estimation based only on a sample of observations is called **sampling error**. In contrast, there may also be nonsampling errors while observing or measuring an individual. A census approach, while free from the sampling error may suffer from nonsampling errors. In fact, in measuring a large number of individuals, due to `inspection fatigue' the nonsampling error may be large. The theory of sampling may allow one to estimate the sampling error but will not be able to assess the extent of nonsampling errors.

In order to have a better grasp of what we have discussed above, you may try this exercise.

---

E2) List the advantages and disadvantages of using a sampling approach instead of a census approach for studying a characteristic.

---

You may also recall at this stage that in Block-2, you learnt about sampling distributions or the derived distribution of functions of a sample of observations. The issues studied in that block are to be distinguished from the issues we will be considering in this block. In Block-2, you had a conceptual infinite population, such as the drying time required for a formulation of paint, and you had ten realizations of the random variable `drying time' on ten pieces of wood on which this formulation of paint was applied. These ten observations were assumed to be ten independent realizations of the same random variable having a theoretical distribution and you derived the distribution of their average and other functions. The issues to be studied in this block relate to sampling from a finite population and no theoretical distribution of the characteristic is assumed.

Before proceeding further to discuss the methods of simple random sampling, we shall introduce some concepts and definitions, which we shall be using frequently in our discussicn.

## 12.3 PRELIMINARIES

Let us assume that we wish to find out the proportion of votes a particular party A is expected to get in an election in a particular constituency.

An **element** is a unit for which information is sought. In this example the element is a registered voter in the constituency. The study variable will be measured as one if the voter prefers to vote for the party A, otherwise, the measurement will be taken as zero.

The population as you already know, is an aggregate of elements about which the inference is to be made; the collection of all the registered voters of the constituency in this case constitutes the population. A population is **finite** if it consists of a finite number of elements. In this unit we shall consider the case of finite population only.

For studying a population we select some of the elements or a collection of elements (i.e. a sample) of the population on which observations are made. These elements are called the **sampling units**. In our example, if households, which has got a number of elements i.e., individual voters, are to be selected then households are sampling units. Sampling units are non-overlapping collections of elements of the population.

For selection purposes, identity of sampling units is necessary. Usually, a list of sampling units of the population provides such an identity. A complete list of sampling units which represents the population to be covered is called a **sampling frame**. The number of units in the sample is the **sample size**.

In the entire theory of sampling, the approach for selecting a **representative** sample and making a good **estimate** from the sample is addressed. In the example considered here, preference for the party is asked only from the registered voters selected in the sample. This information is then used to determine the proportion of all votes that party A is expected to get in the election. It is therefore necessary to exercise a great deal of care in selecting sampling units for a sample survey. Other examples of results based on sample surveys are quite common in practice. Whenever you read about important figures like production of important crops, or average income of people in rural or urban areas, it must be realised that such figures are invariably based on results of well planned sample surveys.

If the units in the sample are selected using some random mechanism then such a procedure is called **random sampling** or **probability sampling.** In this method samples are selected according to certain laws of probability in which each unit of the population has some definite probability of being selected in the sample. All other sampling procedures, which are not based on random procedures but are based on subjective judgement or convenience of the sampler, are known as **non-random sampling** or **non-probability sampling.** They are also termed as **purposive sampling** or **judgement sampling.** Clearly, inferences drawn on the basis of a purposive sample can often be subjective and biased.

Random sampling is preferred over non-random sampling for a variety of reasons. Besides eliminating the subjectivity in selection, it provides a measure of reliability associated with the estimates developed from the samples Thus, one can make inferences from the sample with a known level of confidence. As stated earlier, in random sampling procedure, every unit in the population is assigned definite probability of selection. The randomness associated with the sampling procedure is the key to make valid inferences from the sample.

Samples are often selected by adopting the procedure of **one after the other draw procedure** or **unit by unit** selection. If the units selected at one draw are replaced in the population before the next draw then the procedure is called **with replacement** (WR) procedure. If the units are not replaced in the population and the selection is made from the remaining units then the selection is called **without replacement** (WOR). If a population consists of N units and a sample of size n is to be selected, number of possible samples for with replacement procedure is $N^n$. In case of without replacement sample, if the order of sample units is ignored then there are $\binom{N}{n}$ possible samples.

You may now try to solve the following exercises to see whether you have grasped the basic concepts of sampling discussed above.

E3) Define population, sampling unit and sampling frame for conducting surveys on each of the following subjects.

a) Measurement of the volume of timber available in a forest.
b) Annual yield of apple fruit in a hilly district.
c) Study of nutrient contents of food consumed by the residents in a city.

E4) Consider a population consisting of 5 villages, the areas (in hectares) of which are given below

| Village | A | B | C | D | E |
|---|---|---|---|---|---|
| Area | 760 | 343 | 657 | 550 | 480 |

Enumerate all possible WOR samples of size 3. Also write the values of the study variable (area) for the sampled units.

List all the WR samples of size 3 along with their area values.

Now, we consider the simplest of the random sampling procedures i.e., simple random sampling.

## 12.4 SIMPLE RANDOM SAMPLING

If each sample among the all possible samples has the same chance of being selected, then the associated method is called **simple random sampling**.

In simple random sampling procedure with unit by unit selection, every unit has got equal chance (probability) of selection at every draw. However, the converse is not true i.e. there are sampling schemes in which every unit gets the same chance of selection but they are not simple random sampling methods e.g. the systematic sampling. You shall learn about such sampling methods later in this unit. We now try to answer the question which may be occurring in your mind.

**How to select a simple random sample (SRS)?**

We consider the selection of a simple random sample through unit by unit selection method. At every draw equal probabilities are to be assigned to the available sampling units of the population. Thus, a pre-requisite for the selection is a random device by which selections are to be made. The most commonly used procedures for selecting a SRS are (1) lottery method, (2) through the use of random number tables. Let us discuss these methods one by one..

**Lottery method**

As the name suggests, units from the given frame (of size N) are selected using any procedure of generating a number randomly through lottery procedures. The simplest method may be writing down N numbers on identical slips of papers and drawing one of the slips after thoroughly mixing the slips. The number on the selected slip indicates the unit selected. For instance, suppose we have a population of 400 individuals and we wish to draw a random sample of 40 individuals. We can number the individuals of the population serially from 1 to 400. We can than take 400 identical slips of paper, write numbers 1 to 400 on them, put them in a box, mix them thoroughly and pick out 40 slips, one by one without looking. This gives us a random sample of 40 individuals. In with replacement procedure, the slip is replaced before the next draw while in case of without replacement it is not replaced. The sampling is continued till desired number of units are selected.

Any other randomization device such as pack of cards or random disc etc, may be used. However, the procedure becomes cumbersome if large number of selections is to be made as numbering of the slips become inconvenient and one has to be careful to see that the slips are thoroughly mixed after each draw. This method is not so common in random selections. We now discuss another method which makes use of random number tables.

**Through random number tables**

Before discussing the method based on the use of random number tables you may like to know what random numbers are? Random numbers are numbers generated by a random procedure involving repeated independent trials. Such numbers are generated with the help of random digits 0 through 9. When we say random digits 0 through 9 it is assumed that a trial of the procedure yields each of the ten digits with probability 0.1. One simple way of generating these random digits is to take ten cards of the same size and write the digits 0 through 9 on the cards so that each card has a different digit. Then take a large hat, say, toss in the cards and mix them well. Now choose a card at random from the hat. Write down on a piece of paper the digit appearing on the card you have chosen. Put the card back into the hat and mix the cards again. Repeat the procedure by choosing a card at random, writing down the digit appearing on the card, replacing, mixing, choosing again, and so on. The string of digits we write down constitutes a string of random digits because it has been produced by a random device supposed to yield each digit with probability 0.1 in independent trials. Random digits can also be produced by using a modified roulette wheel in which the wheel is divided into ten equal parts, each one corresponding to one of the ten digits.

Given random digits, we can get more complicated random numbers. Suppose we have generated the sequence 3217900597 of ten digits. Then each digit is random, and also the two digit numbers 32, 17, 90, 05, 97 obtained by taking the numbers two at a time are random numbers because they have been produced by a random procedure so that each of one hundred two-digit numbers 00 through 99 has the probability 0.01 of appearing, and moreover selection of these two digit numbers are independent. Taking the original ten digit sequence and choosing the two digits at a time going backward to get 79,50,09,71,23 also gives random two-digit numbers. In the same way you might think of some other ways to get two-digit numbers using the generated string, as long as the method does not use the same selection more than once.

---

E5) Give two more sequences of 5 two digit random numbers obtained by using the string 3217900597 of ten digits.

---

In a similar manner, random numbers with three, four or even more digits can be obtained by using the given string of random digits. There are several standard random number tables available which give the arrangement of these numbers in a rectangular manner. Some of these which are commonly used are prepared by Tippett (1927), Fisher and Yates (1938), Kendall and Smith (1939), Rand Corporation (1955), and Rao et al. (1974). One such random number tables is reproduced in Appendix A. You may note that this random number table can be used as single digited numbers or two digited numbers or, three or four digited numbers depending on the size of the population you are sampling from. We shall now illustrate the use of random number tables for selecting samples.

We shall discuss here three commonly used methods of using random number tables for selection of simple random samples.

**Direct Approach.** The first step in the method is to assign serial numbers 1 to N to the N population units. If the population size N is made up of K digits, then consider K digit random numbers, either row wise or column wise, in the random number table. The sample of required size is then selected by drawing, one by one, random numbers from 1 to N, and including the units bearing these serial numbers in the sample.

**Problem 1:** Consider a population of 56 households. Select a simple random sample of 10 households by with replacement as well as without replacement methods.

**Solution:** Here the sampling unit is a household. The first step is to serially arrange the households if they are not already arranged so. Since, the population size is a number consisting of two digits we have to use two digited random number table. Alternatively, in the table of random numbers (Appendix A) the first two digits of any column can be used randomly. While using these tables, it is advisable to take a blind start on the table by placing your finger on the table with closed eyes – let it be column 7, row 6 of the first page of Appendix A. Then the first number is 20 and going down the page subsequent random numbers between 1 to 56 are 12, 03 etc.

By selecting first 10 random numbers from 1 to 56, without discarding repetitions for **with replacement procedure** (WR), we obtain the serial numbers of the households in the sample. These are given below:

20 12 03 16 30 15 24 37 01 15

For **without replacement procedure** (WOR), repetitions have to be avoided and the number 15, when appears again at the tenth draw, should be dropped. The next two digits are then chosen. Thus, a WOR sample will consist of:

20 12 03 16 30 15 24 37 01 07

——— x ———

This procedure may involve number of rejections of random numbers, since zero and all the numbers greater than 56 appearing in the table are not considered for selection. The use of random numbers has, therefore, to be modified. We now discuss two of the commonly used modified procedures,

**Remainder Approach.** In this method if the population size N is a K digit number, then first we have to determine the highest K digit multiple of N. Let it be N'. Then a random number r is selected, such that $1 \leq r \leq N'$. This number r is then divided by N and let the remainder be R. The unit bearing the serial number equal to the remainder R, is then considered as selected. If remainder is zero, the last unit is selected. As an illustration, let N=24. Here N is a two digit number. The highest two digit multiple of 24 is 96. Let us now choose a number between 1 and 96 say, 83. On dividing 83 by 24, we get the remainder as 11. Therefore, the unit bearing serial number 11 is selected in the sample. Then another number between 1 and 96 is selected and the process is repeated till the sample of required size is selected. As before, the repeated selections of population units in the sample are permitted for WR sample, whereas they are rejected and only distinct units selected for a WOR sample. With a little variation in this method we consider another method.

**Quotient Approach.** As before, let N be a K digit number and N' be the highest K digit multiple of N, such that $N' = Nm$ for some integer m. Select a random number r from 0 to N'-1. Then the unit having serial number (Q+1) is included in the sample, where Q is the quotient when r is divided by m. For instance, if N=24 then N'=96 and m=4. Let a random number r=49 be chosen from 0 to 95. Then dividing 49 by 4 one gets the quotient Q=12. The unit bearing serial number (Q+1)=13 is then selected in the sample. The process is repeated by selecting each time a new number r from 0 to 95 till a sample of required size is obtained.

As we have mentioned earlier while using the random number tables, any starting point can be used, and one can move in any predetermined direction along the rows or columns. However, normally as a convention, column-wise selection is followed. If more than one sample is to be selected in any problem, each should have an independent starting point.

Besides the methods discussed above, some more methods for sample selection are available in the literature. However, being operationally inconvenient, these are usually not employed in practice.

In order to get conversant with the methods discussed above for selecting a simple random sample, you may try the following exercise. While doing this exercise you will also get convinced that the number of rejections in quotient and remainder approach are much less as compared to the direct approach.

E6) Select a simple random sample of 10 households from the same population of 56 households by WR and WOR methods, using remainder approach and quotient approach.

The whole purpose of sampling is to collect information about the population from which the sample is drawn. It is used to study the unknown characteristics in the population called **parameters**. However, a sample cannot tell us about the population parameters exactly, it can only estimate parameters of the population. In the next section we shall see how these estimations are done.

## 12.5 ESTIMATION OF POPULATION PARAMETERS

In order to infer about population parameters, we compute various quantities from the sample. These computed quantities from the sample are called **statistics**. In general, we can say that a statistic is any quantity computed from a sample. Values such as mean, variance and standard deviations derived from samples are sample statistics which are then used to estimate population parameters and hence are called **estimators**.

Some of the important population parameters required to be estimated are population mean, variance and proportion. When the population is of size N, comprising of units with variate values $Y_1, Y_2, \ldots, Y_N$, then **population mean** and **variance** are

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i = N^{-1}Y, \text{ where } Y = \text{Total population and } \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \overline{Y}\right)^2 \qquad (1)$$

respectively. The corresponding formulas for **sample mean** and **variance** are given by

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ and } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2 \qquad (2)$$

respectively for a sample of size n with variate values $y_1, y_2, \ldots y_n$. . You may also note here that a **paramter is a fixed unknown quantity**. For example, the average height of the population of adult Indian males at a given time, has a single fixed value.

A **statistic** on the other hand is a **variable quantity**. The value of a statistic (or, an estimator) computed from different samples would differ from sample to sample. We know that the heights of adult Indian males vary. If two samples of the same size are drawn from this population then it may happen that one sample has a few more taller people than the other. Hence, the average height computed from one sample is likely to be different from that computed from the other. Thus, there is a need to study this variability in the statistic if it is to be used as an estimator. In other words, we can say that there is a need to study the distribution known as sampling distribution of an estimator.

The sampling distribution of an estimator helps in defining certain desirable criteria for goodness of an estimator. One of the most important criteria is unbiasedness. The estimator is said to be unbiased for the parameter t, if E(t) = t. where E(.) stands for expectation. This expectation is computed by averaging the value of t over all possible samples. The criterion of unbiasedness ensures that on an average the estimator will take value equal to the unknown population parameter t. We now illustrate the concept of a sampling distribution through an example.

**Problem 2 :** Consider a simple random sample (WOR) of two households from a population of five households having monthly income (in rupees) as follows :

| Household | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Income (rupees) | 1560 | 1490 | 1660 | 1640 | 1550 |

Enumerate all possible samples (WOR) of size 2 and show that the sample mean gives an unbiased estimate of population mean.

**Solution:** Population mean $\bar{Y} = \frac{1560+1490+1660+1640+1550}{5} = \frac{7900}{5} = 1580$

All possible samples and their corresponding sample means in this case are presented in Table-2.

**Table 2 : All samples and their corresponding sample means in SRSWOR**

**(N=5, n=2)**

| Sample No. | Units in sample | Probability | Sample observations | | Sample mean |
|---|---|---|---|---|---|
| | | | $y_1$ | $y_2$ | $\bar{y} = \frac{y_1+y_2}{2}$ |
| 1 | 1,2 | 1/10 | 1560 | 1490 | 1525 |
| 2 | 1,3 | 1/10 | 1560 | 1660 | 1610 |
| 3 | 1,4 | 1/10 | 1560 | 1640 | 1600 |
| 4 | 1,5 | 1/10 | 1560 | 1550 | 1555 |
| 5 | 2,3 | 1/10 | 1490 | 1660 | 1575 |
| 6 | 2,4 | 1/10 | 1490 | 1640 | 1565 |
| 7 | 2,5 | 1/10 | 1490 | 1550 | 1520 |
| 8 | 3,4 | 1/10 | 1660 | 1640 | 1650 |
| 9 | 3,5 | 1/10 | 1660 | 1550 | 1605 |
| 10 | 4,5 | 1/10 | 1640 | 1550 | 1595 |
| Average | | | | | 1580 |

It may be seen that the average of sample means (1580) is equal to the population mean. This shows the unbiased nature of the sample mean as an estimator of population mean.

———— × ————

In the same way unbiasedness can be shown in the case when all possible samples of size 2 are drawn with replacement. We are leaving this for you to do it yourself.

---

E7) In Problem 2 above, enumerate all possible samples (WR) of size 2 and show that the sample mean is an unbiased estimator of the population mean.

---

The **sampling variance** is the variance of the sampling distribution of the estimator. It measures the divergence of the estimator from its expected value. If $\hat{\theta}$ is an estimator of $\theta$ then,

$$V\left(\hat{\theta}\right) = E\left(\hat{\theta} - E\hat{\theta}\right)^2$$

The positive square root of sampling variance is termed **standard error** (SE).

$$SE\left(\hat{\theta}\right) = \sqrt{V\left(\hat{\theta}\right)}$$

Thus, standard error is the standard deviation of the sampling distribution. It measures the precision of the estimator particularly in view of the fluctuations due to specific sampling design.

In case of **simple random sampling with replacement** (SRSWR). **Variance** of $\bar{y}$ can be written as

$$V(\bar{y}) = \frac{\sigma^2}{n} \text{ where, } \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2 \tag{3}$$

and **estimated variance** is

$$v(\bar{y}) = \frac{s^2}{n} \text{ where, } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{4}$$

For simple **random sampling without replacement** (SRSWOR)

variance of $\bar{y}$ is

$$V(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)S^2 \text{ where, } S^2 = \frac{N\sigma^2}{N-1} = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 \tag{5}$$

and estimated variance is

$$v\,(\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)s^2 \text{ where, } s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 \tag{6}$$

To have a better understanding of the above formulas, you may try the following exercises.

E8) Suppose we have a population of 5 students enrolled for statistics course and a counsellor wants to find the average amount of time spent by each student in preparing for classes each week. The amount of time (in hours) each student spends per week is given by 7,3,6,10 and 4. If the counsellor takes a sample of three students WOR, obtain the sampling distribution of the sample mean. Compute the population mean and the mean and standard error of the sampling distribution.

E 9) In the data of Example 2, show that the sample mean square ($s^2$) is an unbiased estimator of the population mean square ($S^2$).

**Estimation Of Proportion**

Sometimes interest lies in estimating the population proportion. Examples, such as, proportion of persons below poverty line or proportion of female members in a particular group or proportion of persons getting degrees through distance education etc. are very common. In all these examples the population is considered as divided in two parts on the basis of an attribute.

For instance, a crop field may be irrigated or not irrigated. If it is irrigated, we say that it possesses the characteristic 'irrigation'. If it is not irrigated, we say that it does not posses the particular characteristic of irrigation. If we are interested in estimating the proportion of irrigated fields, the population of N fields can be defined with variate $y_i$ as

$$y_i = 1, \text{ if the field is irrigated}$$

$$= 0, \text{ otherwise}$$

If the total number of irrigated fields be $N_1$ out of N then

$$\sum_{i=1}^{N} y_i = N_1$$

Thus, $\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} y_i = \frac{N_1}{N} = P$ = proportion of irrigated fields.

Thus, the problem of estimating a population proportion becomes that of estimating a population mean by defining the variate as above. Now if a simple random sample of size n is taken from the population and if $n_1$ units out of n possess that characteristic, then sample proportion is given by

$$p = \frac{n_1}{n}$$

Thus, $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i = \frac{n_1}{n} = p.$

It follows that, in SRSWR as well as SRSWOR, p is an unbiased estimator of P. The variances and estimator of variances in case of WR and WOR procedures are given as follows.

For **SRSWR**

**Variance** of p is

$$V(p) = \frac{PQ}{n}, \quad \text{where } Q = 1 - P, \tag{7}$$

and **estimated variance** is

$$V(p) = \frac{pq}{n-1}, \ q = 1 - p \tag{8}$$

For **SRSWOR**

**Variance** of p is

$$V(p) = \frac{N-n}{N-1}\left(\frac{PQ}{n}\right), \ Q = 1 - P \tag{9}$$

and **estimated variance** is

$$v(p) = \frac{N-n}{N}\left(\frac{pq}{n-1}\right) \ q = 1 - p \tag{10}$$

We now take up an example to illustrate what we have discussed above.

**Problem 3:** Obtain the sampling distribution of the sample proportion and the standard error of the proportion of households having monthly income more than Rs.1550 in the population given in Problem 2 by considering simple random sample (WOR) of two households.

**Solution:** In the population of 5 households the income of household 1,3 and 4 is more than Rs.1550. Population proportion of interest to us is $P = 3/5$.

Let us now work out the sampling distribution of the sample proportion by getting the sample proportion of households with income exceeding Rs.1550 from each of 10 possible samples listed in Table 2. In each sample we score a household as 0 if its income is Rs.1550 or less and as 1 if it exceeds Rs.1550. Then the mean score in each sample shown in Table-3 below gives the sample proportions.

**Table-3: Computation of the sample proportions**

| Sample No. | Score | Mean Score |
|---|---|---|
| 1 | 1,0 | ½ |
| 2 | 1,1 | 1 |
| 3 | 1,1 | 1 |
| 4 | 1,0 | ½ |
| 5 | 0,1 | ½ |
| 6 | 0,1 | ½ |
| 7 | 0,0 | 0 |
| 8 | 1,1 | 1 |
| 9 | 1,0 | ½ |
| 10 | 1,0 | ½ |

Standard error SE(p) of the proportion given by Formula (9) is

$$SE(p) = \sqrt{\frac{N-n}{N-1}\frac{P(1-P)}{n}}$$
$$= \sqrt{\frac{5-2}{5-1}\left(\frac{3}{5}\right)\left(1-\frac{3}{5}\right)\frac{1}{2}}$$
$$= \sqrt{\frac{9}{100}} = \frac{3}{10} = 0.3$$

——— x ———

Why don't you try this exercise now

E10) IT facility committee of IGNOU has a total of eight members whose ages in years are 27,32,33,26,43,52,28 and 25. The committee has a rule which requires a minimum age of 33 for a member to be the chairperson. Assume that a simple random sample of size 4 is selected to provide an estimate of the population proportion eligible to be chairperson. Find the mean and standard deviation of the sampling distribution.

**Sample size determination**

While planning a sample survey, a decision has to be made regarding the sample size in the initial stages. Having decided about the method of selection, one has to determine about the sample size in view of the resources available as well as the desired level of precision of the estimators. Larger sample sizes will involve more cost in data collection as well as data analysis while smaller sample sizes will reduce the precision of estimate. Therefore, a balance has to be struck between cost and precision, while deciding about the sample sizes. We consider a simple example to explain the idea behind principles involved in determination of sample size.

An anthropologist wants to study the inhabitants of some island. He wishes to estimate the percentage of inhabitants belonging to blood group O. A simple random sample is to be selected. How large should the sample be?

Some related questions are pertinent here. How accurately does the anthropologist wish to know the percentage of people with blood group O? Suppose, he answers that he will be contented if the percentage is correct within a tolerance limit d of ± 5%. It is also to be understood that even with this specification of tolerable limit of error it is not possible to ensure that the estimates are obtained in this margin in 100% of cases. A level of confidence has therefore to be attached with the estimates. Let confidence level $1-\alpha$ be 95% associated with the estimates.

Let us assume that p is normally distributed about P. It will then lie in the range $(P \pm 2\sigma_p)$ where $\sigma_p$ is the standard deviation of p, apart from a one in twenty chance (i.e. apart from a probability of $\alpha$).

In case of **SRSWR**, $\sigma_p = \sqrt{\frac{PQ}{n}}$. Hence we can put

$$2\sqrt{PQ/n} = \frac{5}{100} \text{ or } n = \frac{4PQ}{25} \text{ x}100 \text{ x } 100 .$$

At this stage some idea about P is needed in order to determine the sample size n. Fortunately, we do not need very accurate estimate of P for this purpose. In fact, if P lies between 0.3 to 0.6 then, the determined sample size lies between 336 to 400. To be on the safe side, 400 may be taken as the initial estimate of **n**.

**Note** that the maximum value of PQ with $0 \le P \le 1$, $Q = 1-P$ is attained at $P = \cdot 5$ and its value is 0.25

In case of **SRSWOR,** for estimation of P, the formula for sample size is given as

$$n = \frac{t^2PQ/d^2}{1+\frac{1}{N}\left(\frac{t^2PQ}{d^2}-1\right)}$$

where d is the margin of tolerable error and t is the abscissa of the normal curve that cuts of an area of $\alpha$ at the tails, $(1-\alpha)$ being the confidence level of the estimate. If N is large, a first approximation to n is $n_0 = t^2P\,Q/d^2$.

In the example considered above tolerable error is within 5% so d = 0.05. We want this at the level of confidence of 95%, or in other words $\alpha$=0.05. Assume that P = 0.5. From the standard normal distribution, the value of the variate corresponding to the two-sides tail of 5% is 1.96 or approximately 2 and hence

$$t=2 \text{ Thus, } n_0=4 \times \frac{0.5 \times 0.5}{(.05)} = 4 \times 100 = 400.$$

In simple random sampling, units are selected randomly at each draw. Now we shall discuss a sampling technique which has a nice feature of selecting the whole sample with just one random start.

## 12.6 SYSTEMATIC SAMPLING

In simple random sampling units are drawn randomly at every draw. In many situations, it may be desirable to select a sample in a systematic way. For example, if we want to have an even spread in terms of spatial distribution, a systematic selection may ensure that units maintain a uniform distance between selected units. In **systematic sampling,** one unit is selected randomly and subsequent units are selected according to a pre-determined system. Invariably uniform distance is adopted for pre-assigned system. Systematic samples actually provide an improvement over simple random samples as the samples are spread more evenly over the entire populations. We now discuss sample selection procedures for systematic sampling.

### 12.6.1 Linear Systematic Sampling

The most commonly adopted procedure of systematic sampling is **linear systematic sampling.** We shall explain the method through an example.
**Problem 4 :** Consider a population of 12 households from which a sample of 3 households is to be selected.

**Solution:** Let the households be arranged serially from 1 to 12. These households are now rearranged in 3 rows of 4 columns as follows:

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 |
| 9 | 10 | 11 | 12 |

Then, for selecting a systematic sample of size 3, we select a random number r (say) between 1 to 4. Starting with r, every 4th unit is selected. Thus, if r=3, then the units selected are 3, 7(=3+4), and 11(=7+4). Thus, if r is selected the entire column of units consisting of r is selected.

———— × ————

In general, this method is applicable if the population size N is a multiple of the sample size n i.e. N = nk where k is an integer. The random number r is selected between 1 to k. Here, r is called a **random start** and k is called **sampling interval**. The sample then comprises of the units r, r+k, r+2k + ..... + r + (n-1)k. The technique will generate k systematic samples with equal probability. The method is known as linear systematic sampling as N units are assumed to be arranged sequentially on a line.

The method is specially suitable in forestry where for estimating the volume of timber this method is used for selection of area units. Some other applications are in industries where items for sample checks are selected systematically in a production process. The concept of systematic sampling is not only confined to spatial distributions. It can also be done over time. In fact in one of the applications in estimation of fish catch from marine resources, sampling of boats on landing centres is carried out systematically over time. Boats arriving every two hourly on selected landing centres are observed.

In the method described above, if N is not a multiple of n, then it may not be possible to get samples of equal size. For example, if N = 14 and n = 3 then the method described above would lead to following arrangement of units in a n x k table as follows:

| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 |
| 9 | 10 | 11 | 12 |
| 13 | 14 | | |

In this case if randomly selected number r between 1 to 4 is 1 then the sample is 1,5,9,13 while if r=3, then the sample is 3,7,11. Thus, samples are not of the same size. Sample size is either 4 or 3 depending on the value of r. As an improvement to this method we shall discuss circular systematic sampling

### 12.6.2 Circular Systematic Sampling

To overcome the difficult of varying sample size in a situation when $N \neq nk$ the procedure is modified slightly by which a sample of constant size is always obtained. This procedure is known as circular systematic sampling

In this method, the N units may be regarded as arranged round a circle. A random start is taken between 1 to N and thereafter every $k^{th}$ unit, k being an integer nearest to $\frac{N}{n}$, in a circular manner is selected until a sample of n units is chosen. Suppose that a unit with random number i is selected. The sample will then consists of the units corresponding to the serial numbers.

$i + jk$; if $i + jk \leq N$

$i + jk - N$; if $i + jk > N$ for $j = 0,1,\ldots,(n-1)$

This method is applicable, even if $N \neq nk$ To illustrate the method, we consider the following example:

**Problem 5:** Consider a population of 14 households from which a sample of size 5 is to be selected .

**Solution:** Here N=14, n=5, k=3. Consider Fig.1:

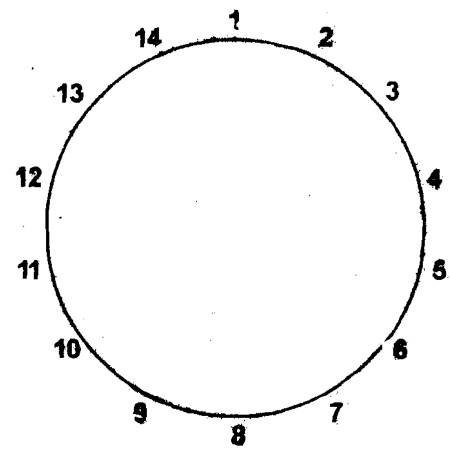


**Fig.1**

Let the random start be 7. Then the selected sample is 7,10,13,2,5. If we start from 9 then the selected sample is 9, 12, 1, 4, 7. Like this we can have 12 more samples as the total number of possible samples in this case is N = 14.

———— × ————

The above method has got the advantage of providing samples of the given size irrespective of the random start. In case of linear systematic sampling the number of possible samples is k while in case of circular systematic sampling it is N. When N is a multiple of n then linear systematic sampling is normally preferred although one could also go for circular systematic sampling. However, when N is not a multiple of n then one should necessarily go for circular systematic sampling.

### 12.6.3 Advantages and Limitations of Systematic Sampling

The systematic sampling has the nice feature of operational convenience because the selection of the first unit determines the whole sample. This operation is easier to understand and can be speedily executed in relation to simple random sampling. Secondly, systematic samples are well spread over the population and there is no risk that any large part of the population will be left unrepresented. For populations with linear trend, systematic sampling is more efficient in comparison to simple random sampling.

Systematic sampling should, however, be cautiously used in case the population exhibits a periodic trend. For periodic populations, the efficiency of the systematic sampling depends upon the value of the sampling interval. If the sampling interval coincides with the period, the sample will contain identical units and consequently the systematic sampling performance becomes very poor. If, however, the sampling interval is an odd multiple of half the period, the systematic sampling becomes most effective.

A serious limitation of this scheme lies in its use with populations having unforeseen periodicity, which may substantially contribute to the bias in the estimate of mean/total. Another serious limitation of the sampling scheme, as mentioned earlier, is that the variance of the estimator cannot be estimated unbiasedly.

You may now try this exercise.

---

E11) A sample of size 4 is to be selected from a population of 11 households. List all the possible sample by (i) linear systematic sampling ii) circular systematic sampling

---

We now end this unit by giving a summary of what we have covered in it.

## 12.7 SUMMARY

In this unit, we have learnt

1) The preliminary concepts and definitions for simple random sampling
2) Method of selecting a simple random sample
3) How to estimate the population mean/total
4) The method of estimating population proportions
5) How to determine sample size in case of simple random sampling
6) The basic concept of systematic sampling
7) How to select a linear systematic sample
8) The method of selecting a circular systematic sample
9) The advantages and limitations of systematic sampling

## 12.8 SOLUTIONS/ANSWERS

E1). i) Collection of all the households in a locality/city is the population and the individuals are the households

ii) Total number of crates loaded in a truck is the population and the individuals are the crates.

iii) All the bulbs produced by the company during a shift is the population and the individuals are the bulbs.

E2) Advantages: less expensive in terms of time, money and energy, less cumbersome, free of nonsampling errors. Also in case of destructive testing, sampling is the only method.

Disadvantages: may suffer from sampling errors. The estimate is only an approximation to the true value.

E3) a) The collection of all trees in the forest is the population, a tree is the individual sampling unit and a list of all trees is the sampling frame.

b) Total number of apple trees in a district is the population, an apple tree is the sampling unit and list of trees is the sampling frame.

c) Collection of all the household in a city consuming the food is the population, an individual household is the sampling unit and a list of all households selected for a sample is the sampling frame.

E4) a) Table 3: All samples and their corresponding sample means in SRSWOR (N=5, n=3) (order of sample units is ignored)

| Sample No. | Units in Sample | Sample areas | | | Sample means $\bar{y} = \frac{y_1 + y_2 + y_3}{3}$ |
|---|---|---|---|---|---|
| | | $y_1$ | $y_2$ | $y_3$ | |
| 1 | A,B,C | 760 | 343 | 657 | 586.67 |
| 2 | A,B,D | 760 | 343 | 550 | 551 |
| 3 | A,B,E | 760 | 343 | 480 | 527.67 |
| 4 | B,C,D | 343 | 657 | 550 | 516.67 |
| 5 | B,C,E | 343 | 657 | 480 | 493.33 |
| 6 | C,D,E | 657 | 550 | 480 | 562.33 |
| 7 | C,D,A | 657 | 550 | 760 | 655.67 |
| 8 | D,E,A | 550 | 480 | 760 | 596.67 |
| 9 | A,C,E | 760 | 657 | 480 | 632.33 |
| 10 | B,D,E | 343 | 550 | 480 | 457.67 |
| | Average | | | | 558 |

b) Similarly make a table for all samples of size 3 with replacement. There will be $5^3$ samples in all.

E5) One sequence could be 21,79,00,59,73. Similarly give another.

E6) **Remainder approach:** For N=56, the highest two digit multiple of N is N itself. Using Appendix A select a two digit random number r, s.t. $1 \le r \le 56$.

r=44 is one possibility. Also $r/N = \frac{44}{56}$ gives quotient as 0 and remainder is 44. Thus select the unit with serial number 44. Likewise you can select other 9 units also. One such simple random sample (WOR) of 10 households selected could be a sample of households with serial numbers

44, 49, 40, 15, 12, 38, 29, 52, 22, 50

and a sample of 10 households (WR) could be

44, 49, 40, 15, 44, 12, 38, 29, 52, 22

**Quotient approach:** In this case m=1. So if a random number r is selected s.t

$1 \le r \le 56$, say r=44, then dividing r by m we get Q = 44 and selected unit is the unit with serial number 44+1 = 45. Likewise you can select a SRS of 10 households (WR) and (WOR) from the population of 56 households.

E7) Average of sample means (1580) is equal to the population mean and thus the sample mean is an unbiased estimator of the population mean. Calculations are shown in Table 4 in the next page.

Table 4: All samples and their corresponding sample means in SRSWR (N=5, n=2)
(order of sample units ignored)

| Sample No. | Units in Sample | Probability | Sample observations | | Sample mean |
|---|---|---|---|---|---|
| | | | $y_1$ | $y_2$ | |
| 1 | 1,2 | 1/10 | 1560 | 1490 | 1525 |
| 2 | 1,3 | 1/10 | 1560 | 1660 | 1610 |
| 3 | 1,4 | 1/10 | 1560 | 1640 | 1600 |
| 4 | 1,5 | 1/10 | 1560 | 1550 | 1555 |
| 5 | 2,3 | 1/10 | 1490 | 1660 | 1575 |
| 6 | 2,4 | 1/10 | 1490 | 1640 | 1565 |
| 7 | 2,5 | 1/10 | 1490 | 1550 | 1520 |
| 8 | 3,4 | 1/10 | 1660 | 1640 | 1650 |
| 9 | 3,5 | 1/10 | 1660 | 1550 | 1605 |
| 10 | 4,5 | 1/10 | 1640 | 1550 | 1595 |
| 11 | 1,1 | 1/10 | 1560 | 1560 | 1560 |
| 12 | 2,2 | 1/10 | 1490 | 1490 | 1490 |
| 13 | 3,3 | 1/10 | 1660 | 1660 | 1660 |
| 14 | 4,4 | 1/10 | 1640 | 1640 | 1640 |
| 15 | 5,5 | 1/10 | 1550 | 1550 | 1550 |
| Average | | | | | 1580 |

**E8)** Distribution of Sample mean in sample size 3

| Sample mean value | No. of Samples giving this means (frequency) | Relative frequency |
|---|---|---|
| 4.33 | 1 | 1/10 |
| 4.67 | 1 | 1/10 |
| 5.33 | 1 | 1/10 |
| 5.67 | 2 | 2/10 |
| 6.33 | 1 | 1/10 |
| 6.67 | 2 | 2/10 |
| 7.0 | 1 | 1/10 |
| 7.67 | 1 | 1/10 |
| Total | 10 | 1 |

$$\text{Population mean } = \mu = \frac{7+3+6+10+4}{5} = \frac{30}{5} = 6.$$

The mean of the above distribution

$$\sigma = \sqrt{\frac{\sum_{i=1}^{5}(y_i-\mu)^2}{N}} = \sqrt{\frac{(7-6)^2+(3-6)^2+(6-6)^2+(10-6)^2+(4-6)^2}{5}}$$

$$= 2.45\ (N=5,\ n=3)$$

$$S.E = \frac{2.45}{\sqrt{3}}\sqrt{\frac{5-3}{5-1}} = 1.00 \text{ (approximately)}$$

$$= \frac{4.33\times1+4.67\times1+5.33\times1+5.67\times2+6.33\times1+6.67\times2+7.0\times1+7.67\times1}{10}$$

$$= \frac{60}{10} = 6$$

$$S.E = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \text{ (Using formula (5))}$$

where $\sigma$ is the population standard deviation

E9) Sample Variance = $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$

$= 2450 + 5000 + 3200 + 50 + 14450 + 11250 + 1800 + 200 + 6050 + 4050$
$= 48500$

average of 10 sample mean squares gives $E(s^2)$.

Here $E(s^2) = 48500/10 = 4850$ (i)

Also, population mean square =

$$S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 = \frac{1}{4}(400+8100+6400+3600+900)$$

$$= \frac{1}{4} \times 19400 = 4850 \qquad \text{ii)}$$

Thus from (i) and (ii)

$$E(s^2) = S^2$$

i.e. sample mean square provides an unbiased estimator of the population mean square.

E10) The mean of the Sampling distribution is the population proportion P,

$$P = \frac{3}{8}$$

S.E for the sampling distribution

$$= \sqrt{\frac{P(1-P)}{n}}\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{\frac{3}{8}(1-3/8)}{4}}\sqrt{\frac{8-4}{8-1}} = \sqrt{\frac{3}{8}\times\frac{5}{8}\times\frac{1}{4}}\sqrt{\frac{4}{7}} = 0.183$$

E11) i) $N=11, n=4, k=\frac{11}{4}=3$ (approx). Arranging the units in 4 rows of 3 columns each (except for the last row) we get table as follows:

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |
| 10 | 11 | |

selecting a number r between 1 and 3, possible samples are 1,4,7,10; 2,5,8,11; 3,6,9 of size 4 or 3.

ii) Here N=11, n=4, k=3. Consider Fig.2

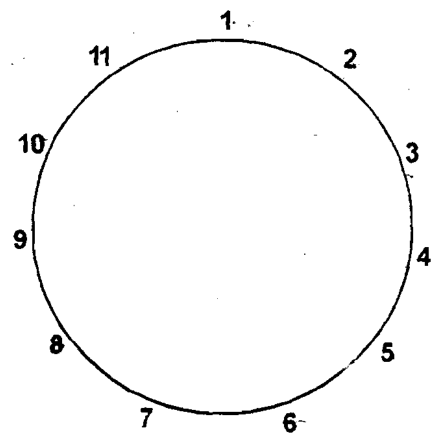


**Fig.2**

Let the random start be 2. Then sample selected is 2,5,7,10. If random start be 5 then sample selected is 5,8,11,3. Likewise you can write the remaining 9 samples as the total number of possible samples is N=11.

# Appendix A : Random Numbers

| Column Numbers | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| 3436 | 6833 | 5809 | 9169 | 5081 | 5655 | 6567 | 8793 | 6830 | 1332 |
| 6133 | 4454 | 2675 | 3558 | 7624 | 5736 | 2184 | 4557 | 0496 | 8547 |
| 9853 | 3890 | 5535 | 3045 | 9830 | 5455 | 8218 | 9090 | 7266 | 4784 |
| 5807 | 5692 | 6971 | 662 | 6751 | 5001 | 5533 | 2386 | 0004 | 2855 |
| 6291 | 0924 | 1298 | 7386 | 5856 | 2167 | 8299 | 9314 | 0333 | 8803 |
| 4725 | 9516 | 8555 | 0379 | 7746 | 9647 | 2010 | 0979 | 7115 | 6653 |
| 7697 | 6486 | 3720 | 6191 | 3552 | 1081 | 6141 | 7613 | 5455 | 3731 |
| 3497 | 2271 | 9641 | 0304 | 4425 | 6776 | 1205 | 2953 | 5669 | 1056 |
| 8940 | 4765 | 1641 | 0606 | 4970 | 7582 | 7991 | 6480 | 2946 | 5190 |
| 1122 | 6364 | 5264 | 1267 | 4027 | 4749 | 0338 | 8406 | 1213 | 5355 |
| 4333 | 0625 | 3947 | 1373 | 6372 | 9036 | 7046 | 4325 | 3491 | 8989 |
| 7685 | 1550 | 0853 | 4276 | 1572 | 9348 | 6893 | 2113 | 8285 | 9195 |
| 0592 | 8341 | 4430 | 0496 | 9613 | 2643 | 6442 | 0870 | 5449 | 8560 |
| 3506 | 0774 | 0447 | 7461 | 4459 | 0866 | 1698 | 0184 | 4975 | 5447 |
| 8368 | 2507 | 3565 | 4243 | 6667 | 8324 | 3063 | 8809 | 4248 | 1190 |
| 2630 | 1112 | 6680 | 4863 | 6813 | 4149 | 8325 | 2271 | 1963 | 9569 |
| 3883 | 3897 | 1848 | 8150 | 8184 | 1133 | 6088 | 3641 | 6785 | 0658 |
| 1123 | 3943 | 5248 | 0635 | 9265 | 4052 | 1509 | 1280 | 0953 | 9107 |
| 1167 | 9827 | 4101 | 4496 | 1254 | 6814 | 2479 | 5924 | 5071 | 1244 |
| 7831 | 0877 | 3806 | 9734 | 3801 | 1651 | 7169 | 3974 | 1725 | 9709 |
| 2487 | 9756 | 9886 | 6776 | 9426 | 0820 | 3741 | 5427 | 5293 | 3223 |
| 1245 | 3875 | 9816 | 8400 | 2938 | 2530 | 0158 | 5267 | 4639 | 5428 |
| 5309 | 4806 | 3176 | 8397 | 5758 | 2503 | 1567 | 5740 | 2577 | 8899 |
| 7109 | 0702 | 4179 | 0438 | 5234 | 9480 | 9777 | 2858 | 4391 | 0979 |
| 8716 | 7177 | 3386 | 7643 | 6555 | 8665 | 0768 | 4409 | 3647 | 9286 |
| 9499 | 5280 | 5150 | 2724 | 6482 | 6362 | 1566 | 2469 | 9704 | 8165 |
| 3125 | 4552 | 6044 | 0222 | 7520 | 1521 | 8205 | 0599 | 5167 | 1654 |
| 3788 | 6257 | 0632 | 0693 | 2263 | 5290 | 0511 | 0229 | 5951 | 6808 |
| 2242 | 2143 | 8724 | 1212 | 9485 | 3985 | 7280 | 0130 | 7791 | 6272 |
| 0900 | 4364 | 6429 | 8573 | 9904 | 2269 | 6405 | 9459 | 3088 | 6903 |
| 7909 | 4528 | 8772 | 1876 | 2113 | 4781 | 8678 | 4873 | 2061 | 1835 |
| 0379 | 2073 | 2680 | 8258 | 6275 | 7149 | 6858 | 4578 | 5932 | 9582 |
| 0780 | 6661 | 0277 | 0998 | 0432 | 8941 | 8946 | 9784 | 6693 | 2491 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 8478 | 8093 | 6990 | 2417 | 0290 | 5771 | 1304 | 3306 | 8825 | 5937 |
| 2519 | 7869 | 9035 | 4282 | 0307 | 7516 | 2340 | 1190 | 8440 | 6551 |
| 2472 | 0823 | 6188 | 3303 | 0490 | 9486 | 2896 | 0821 | 5999 | 3697 |
| 8418 | 5411 | 9245 | 0857 | 3059 | 6689 | 6523 | 8386 | 6674 | 7081 |
| 8293 | 5709 | 4120 | 5530 | 8864 | 0511 | 5593 | 1633 | 4788 | 1001 |
| 9260 | 1416 | 2171 | 0525 | 6016 | 9430 | 2828 | 6877 | 2570 | 4049 |
| 6568 | 1568 | 4160 | 0429 | 3488 | 3741 | 3311 | 3733 | 7882 | 6985 |
| 6694 | 5994 | 7517 | 1339 | 6812 | 4139 | 6938 | 8098 | 6140 | 2013 |
| 2273 | 6882 | 2673 | 6903 | 4044 | 3064 | 6738 | 7554 | 7734 | 7899 |
| 6364 | 5762 | 0322 | 2592 | 3452 | 9002 | 0264 | 6009 | 1311 | 5873 |
| 6696 | 1759 | 0563 | 8104 | 5055 | 4078 | 2516 | 1631 | 5859 | 1331 |
| 3431 | 2522 | 2206 | 3938 | 7860 | 1886 | 1229 | 7734 | 3283 | 8487 |
| 4842 | 3765 | 3484 | 2337 | 0587 | 9885 | 8568 | 3162 | 3028 | 7091 |
| 8295 | 9315 | 5892 | 6981 | 4141 | 1606 | 1411 | 3196 | 9428 | 3300 |
| 4925 | 4677 | 8547 | 5258 | 7274 | 2471 | 4559 | 6581 | 8232 | 7405 |
| 5439 | 0994 | 3794 | 8444 | 1043 | 4629 | 5975 | 3340 | 3793 | 6060 |
| 2031 | 0283 | 3320 | 1595 | 7953 | 2695 | 0399 | 9793 | 6114 | 2091 |

**Source:** Rao, CR, Mitra, S.K., Maitthai, A. and Ramamurthy, K.G. (1974). Formulae and Tables for Statistical Work. Statistical Publishing Society. Indian Statistical Institute Calcutta.