

Block 2

Methods of Data Collection



ignou
THE PEOPLE'S
UNIVERSITY

UNIT 4 SAMPLING*

Structure

- 4.0 Objectives
- 4.1 Introduction
- 4.2 Sampling Process
- 4.3 Types of Sampling
 - 4.3.1 Probability Sampling
 - 4.3.2 Non-Probability Sampling
 - 4.3.3 Mixed Sampling
- 4.4 Selection of a Simple Random Sample
 - 4.4.1 Lottery Method
 - 4.4.2 Random Numbers Table Method
 - 4.4.3 Steps in the Use of RNT
 - 4.4.4 Advantages of SRS
 - 4.4.5 Limitations of SRS
- 4.5 Selection of Systematic Random Sample
 - 4.5.1 Advantages of Systematic Random Sampling
 - 4.5.2 Disadvantages of Systematic Random Sampling
- 4.6 Selection of Stratified Random Sample
 - 4.6.1 Proportional Stratified Sample
 - 4.6.2 Disproportional Stratified Sampling
 - 4.6.3 Advantages of Stratified Sampling
 - 4.6.4 Disadvantages of Stratified Sampling
- 4.7 Selection of a Cluster Sample
 - 4.7.1 Steps in Cluster Sampling
 - 4.7.2 Advantages of Cluster Sampling
 - 4.7.3 Disadvantages of Cluster Sampling
- 4.8 Multistage Sampling
- 4.9 Non-Probability Sampling Procedures
 - 4.9.1 Convenience Sampling
 - 4.9.2 Judgement Sampling
 - 4.9.3 Quta Sampling
 - 4.9.4 Snowball Sampling
- 4.10 Determination of the Sample Size
- 4.11 Let Us Sum Up
- 4.12 Answers to Check Your Progress

* Adapted from Unit 12 of BECC 107: Statistical Methods for Economics; written by Prof. C G Naidu (retd.), School of Vocational Studies, IGNOU

4.0 OBJECTIVES

On completion of this Unit, you should be able to

- explain different methods of drawing a sample;
- describe different types of samples;
- use random number tables to draw a sample; and
- determine the sample size.

4.1 INTRODUCTION

In Unit 2 of this course, you have learned about different types of data, namely, primary data and secondary data. In that Unit, we also discussed the use of different survey techniques like face-to-face interview, telephone survey, postal survey, internet survey, etc. in collecting primary data. In Unit 15, you also have learned the meaning of sampling, advantages of sampling, and sampling error.

In statistics, we often rely on a sample (that is, a small subset of a larger set of data) to draw inferences about the population (that is, the larger set of data). For example, you are interested to know the voting behaviour of Delhi people in the next election. Who will you ask? Naturally, it is not possible for you to ask every single Delhi voter how he or she is likely to vote. Instead, you may query a relatively small number of Delhi voters and draw inferences about entire Delhi from their responses. In this case total voters of Delhi constitute the population and the voters actually queried constitute your sample.

Ideally, the characteristics of a sample should reflect the characteristics of the population from which it is drawn. In such cases, the inferences drawn from a sample are probably applicable to the entire population.

In this unit you will learn how to draw a sample under different population characteristics and how to determine the sample size.

4.2 SAMPLING PROCESS

In conducting a sample survey, the sampling process determines which sampling units will be included in the survey. Sampling makes data collection process more manageable and affordable. It enables the population characteristics to be inferred with minimal errors on the basis of the sample. The sampling process includes defining the target population from which we draw the sample, identification of the sampling frame, selection of the sampling method, and selection of the sampling units.

- 1) **Survey Objectives:** A sample survey begins with the specification of the objectives. We should have a clear and un-ambiguous idea of the objectives of the survey, because all other steps – target population, sampling frame, sampling procedure, etc. – are designed according to survey objectives.

- 2) **Questionnaire Design:** Keeping the objectives of the survey in view we are required to design a questionnaire. We have already learnt the major steps involved in designing of a questionnaire in Unit 1 of this course. In addition to the questionnaire we need to develop training documents for the investigators, particularly when the sample survey is conducted at a larger scale involving a number of investigators.
- 3) **Defining the Target Population:** To draw the samples from a population we must know the target population about which conclusions are to be drawn. The target population is also referred to as the universe. The target population is the group about which we wish to generalize or make inferences from the sample. For example, you want to conduct a sample survey on the family planning methods used by eligible couples in Delhi. Here, all those couples in Delhi in the reproductive age group form the target population.
- 4) **Identifying Sampling Frame:** The sampling frame is a list of cases from the target population. The sampling frame is the actual operational definition of the target population. In our earlier example of eligible couples in Delhi using family planning methods, all the people in the reproductive age group form the sampling frame. Many times we may not be able to list all the cases in the target population for some reason or other. For example, we want to list the people for a survey based on telephone directory. In this case, certainly those people who do not have telephone numbers in the directory will be excluded from the listing. This type of error is called sampling frame error.
- 5) **Selecting Sampling Procedure:** Once the sampling frame is identified, we select appropriate sampling procedure to select the sample for the survey. We will discuss various sampling procedures in detail in the next section of this Unit.
- 6) **Selecting the Sampling Units:** Sampling units are those cases from the sampling frame which are included in the sample by using appropriate sampling procedure. Essentially, a sampling unit is the case on which data is collected. For example, you may decide to take 1000 sampling units from the sampling frame (consisting of all the reproductive age group people in Delhi) for your sample survey.
- 7) **Survey data Processing:** After selection of sampling units the next step is data collection and processing. We need to check the incomplete questionnaires and edit or cross-check the responses wherever there is a doubt. Data entry and tabulation follows.
- 8) **Analysis of Data:** The next step in the sequence is analysis of data. Keeping in view our requirements we analyse the data by using various statistical tools.

- 9) Publication and Dissemination of Results: On the basis of data analysis we prepare technical and research reports. Finally, the socio-economic results of the survey and their implications are discussed in seminars.

4.3 TYPES OF SAMPLING

The method of selecting a sample from a given population is called *sampling*. Basically there are two types of sampling, viz., probability sampling and non-probability sampling. In probability sampling the sampling units are selected according to some chance mechanism or probability of selection. On the other hand, non-probability sampling is based on judgment or discretion of the person making a choice. Thus in non-probability sampling certain units may be selected because of convenience or they serve a purpose or the researcher feels that these units are representative of the population. No random selection on the basis of chance mechanism is involved here.

4.3.1 Probability Sampling

It is also called random sampling. It is a procedure in which every member of the population has a chance or probability of being selected in the sample. It is in this probabilistic sense that the sample is random. The word 'random' does not mean that the sample is obtained in a haphazard manner without following any rule.

Random sampling is based on the well-established principles of probability theory. There are quite a few variants of the random sampling, viz., simple random sampling, systematic random sampling and stratified random sampling. We discuss these types below.

a) Simple Random Sampling

If there is not much variation in the characteristics of the members of a population, we can follow the method of simple random sampling. In this method, we consider the population in its entirety as a homogeneous group and follow the principle of random sampling to choose the members for the sample.

There are two variants of simple random sampling, viz., simple random sampling with replacement (SRSWR) and simple random sampling without replacement (SRSWOR). This difference pertains to the way the sample units are selected. According to the procedure of simple random sampling with replacement (SRSWR), we draw one unit from the population, note down its features and put it back to the whole lot in the sense that the unit again becomes eligible for selection. In this way, the total number of units in the population always remains the same. In other words, the composition of the population remains unchanged, and each member of the population has the *same chance* or probability of being selected in the sample. In fact, if N is the size of the population, this probability is $\frac{1}{N}$. On the other hand, in the case of simple random sampling without replacement, the unit once selected is not returned to the population in the sense

that it becomes ineligible for selection again. Consequently, the composition of the population changes after each draw. Therefore, the probability of any particular unit being selected also changes.

b) Systematic Random Sampling

In this variant of random sampling, only the first unit of the sample is selected at random from the population. The subsequent units are then selected by following some definite rule. For example, suppose we have to choose a sample of agricultural plots. In systematic random sampling, we begin with selecting one plot *at random* and then every j^{th} plot may be selected.

c) Stratified Random Sampling

Stratified random sampling is the appropriate method if the population under consideration consists of heterogeneous units. Here, first we divide the population into certain homogeneous groups or strata from each stratum. Secondly, some units are selected by simple random sampling. Thirdly, after selecting the units from each stratum, they are mixed together to obtain the final sample.

Let us consider an example. Suppose we want to estimate the per capita income of Delhi by a sample survey. It is common knowledge that Delhi is characterised by rich localities, middle class localities and poor localities in terms of the income groups of the people living in these localities. Now, each of these different localities can constitute a stratum from which some people may be selected by adopting simple random sampling procedure.

d) Multi-Stage Random Sampling

Let us consider a situation where we want to obtain information from a sample of households in a large city, say, Delhi. Sometimes, it may not be possible to directly take a sample of households because a list of all the households may not be easily obtained. In such a situation, one may resort to take samples in various stages. Generally, the city is divided into certain geographical areas for administrative purposes. These areas may be termed as city blocks. So in the first stage, some of such blocks may be selected by random sampling. In the next stage, from each of the selected blocks in the first stage, some households may be selected again by the principle of random sampling. In this way, ultimately a sample of households from a large city may be obtained. The above-mentioned example is the case of a two-stage random sampling. However, if the nature of the inquiry so demands, the method of sampling can be extended to more than two stages.

4.3.2 Non-Probability Sampling

We have considered the method of random sampling and some of its variants above. It should be clear that the basic objective of the principle of random

sampling is to eliminate or at least minimise the effect of the subjective bias of the investigator in the selection of the population sample. But for certain purposes, there is a need for using discretion. For example, suppose a teacher has to choose 4 participants from a class of 30 students in a debate competition. Here, the teacher may select the top 4 debaters on the basis of her own conscious judgment about the top debaters in the class. This is an example of purposive sampling. In this method, the purpose of the sample guides the choice of certain members or units of the population.

4.3.3 Mixed Sampling

In mixed sampling, we have some features of both non-probability sampling and random sampling. Suppose an institute has to send 5 students for managerial training in a company during the summer vacation. Initially, it may shortlist about 20 students who are considered to be suitable for the training by applying its own discretion. Then from these 20 students, 5 students may finally be selected by random sampling.

We will discuss the process of drawing various types of samples later in this Unit.

4.4 SELECTION OF A SIMPLE RANDOM SAMPLE

Simple random sampling (SRS) is the basic sampling procedure where we select a sample from the population. In this procedure each unit is chosen entirely through chance mechanism and each unit of the population has an equal chance of being included in the sample. Every possible sample of a given size has the same chance of selection. That is, each unit of the population is equally likely to be chosen at any stage in the sampling process and selection of one unit should in no way influence the selection of another unit in the population.

Simple random sampling should be used with a homogeneous population. That is, all the units in a population should possess the same attributes that we are interested in measuring. The characteristics of homogeneity may include age, sex, income, social status, geographical region, etc.

There are two most commonly used methods to extract a simple random sample. The first is lottery method and the second is a random numbers selection method. Irrespective of which method we decide to use, every element in the sampling frame should be assigned an identifying number.

4.4.1 Lottery Method

The simplest method of selecting a random sample is drawing lottery. In this method, the unit identification numbers are placed in a container and mixed together. Finally, someone draws out numbers at random from the container until the desired sample size is attained. Suppose, we want to select n sample units out of N population units. We assign the numbers 1 to N ; one number to each of the population unit; and write these numbers on N slips. The slips are made as

homogeneous as possible in shape, size, colour, etc. These slips are then put in a container and thoroughly shuffled. Finally, n slips are drawn one by one. The n units corresponding to numbers on slips drawn, will constitute a random sample.

Example 4.1

Let us assume that you are doing some research with a bank branch that wishes to assess customers views on the quality of service in the bank branch. You are asked to select 100 customers as the sample using simple random sampling procedure with lottery method. First, you have to get the sampling frame organized. For this, you will go through the bank records to identify the account holders. You then assign the serial numbers to all the account holders. Suppose there are 1000 account holders and you want to draw $100/1000=10\%$ sample. You could print the serial numbers, tear them into separate strips, put the strips in a container, mix them up real good, close your eyes and pull out the first 100 strips.

The disadvantage of the lottery method is that it would be tedious and the quality of the sample would depend on how thoroughly you mixed them up and how randomly you picked them. Also, as the population size increases, it becomes more and more difficult to draw samples using lottery method.

4.4.2 Random Numbers Table Method

The random numbers are a collection of digits generated through a probability mechanism. The random numbers have the following properties:

The probability that each digit (0,1,2,3,4,5,6,7,8 or 9) will appear at any place is the same, that is $1/10$.

The occurrence of any two digits in any two places is independent of each other.

In this method each unit in the population is assigned a unique number in a sequence. To draw the sample we use a table of random numbers. You can find the random number table (RNT), among other places, in Fisher and Yates (1963): *Statistical Tables for Biological, Agricultural and Medical Research*. An example of a random numbers table is shown in Table 4.1.

Table 4.1: Random Numbers Table

39634 62349 74088 65564 16379 19713 39153 69459 17986 24537
14595 35050 40469 27478 44526 67331 93365 54526 22356 93208
30734 71571 83722 79712 25775 65178 07763 82928 31131 30196
64628 89126 91254 24090 25752 03091 39411 73146 06089 15630

**Methods of Data
Collection**

42831 95113 43511 42082 15140 34733 68076 18292 69486 80468
80583 70361 41047 26792 78466 03395 17635 09697 82447 31405
00209 90404 99457 72570 42194 49043 24330 14939 09865 45906
05409 20830 01911 60767 55248 79253 12317 84120 77772 50103
95836 22530 91785 80210 34361 52228 33869 94332 83868 61672
65358 70469 87149 89509 72176 18103 55169 79954 72002 20582
72249 04037 36192 40221 14918 53437 60571 40995 55006 10694
41692 40581 93050 48734 34652 41577 04631 49184 39295 81776
61885 50796 96822 82002 07973 52925 75467 86013 98072 91942
48917 48129 48624 48248 91465 54898 61220 18721 67387 66575
88378 84299 12193 03785 49314 39761 99132 28775 45276 91816
77800 25734 09801 92087 02955 12872 89848 48579 06028 13827
24028 03405 01178 06316 81916 40170 53665 87202 88638 47121
86558 84750 43994 01760 96205 27937 45416 71964 52261 30781
78545 49201 05329 14182 10971 90472 44682 39304 19819 55799
14969 64623 82780 35686 30941 14622 04126 25498 95452 63937
58697 31973 06303 94202 62287 56164 79157 98375 24558 99241
38449 46438 91579 01907 72146 05764 22400 94490 49833 09258
62134 87244 73348 80114 78490 64735 31010 66975 28652 36166
72749 13347 65030 26128 49067 27904 49953 74674 94617 13317
81638 36566 42709 33717 59943 12027 46547 61303 46699 76243
46574 79670 10342 89543 75030 23428 29541 32501 89422 87474
11873 57196 32209 67663 07990 12288 59245 83638 23642 61715
13862 72778 09949 23096 01791 19472 14634 31690 36602 62943
08312 27886 82321 28666 72998 22514 51054 22940 31842 54245
11071 44430 94664 91294 35163 05494 32882 23904 41340 61185
82509 11842 86963 50307 07510 32545 90717 46856 86079 13769

07426 67341 80314 58910 93948 85738 69444 09370 58194 28207
57696 25592 91221 95386 15857 84645 89659 80535 93233 82798
08074 89810 48521 90740 02687 83117 74920 25954 99629 78978
20128 53721 01518 40699 20849 04710 38989 91322 56057 58573
00190 27157 83208 79446 92987 61357 38752 55424 94518 45205
23798 55425 32454 34611 39605 39981 74691 40836 30812 38563
85306 57995 68222 39055 43890 36956 84861 63624 04961 55439
99719 36036 74274 53901 34643 06157 89500 57514 93977 42403
95970 81452 48873 00784 58347 40269 11880 43395 28249 38743
56651 91460 92462 98566 72062 18556 55052 47614 80044 60015
71499 80220 35750 67337 47556 55272 55249 79100 34014 17037
66660 78443 47545 70736 65419 77489 70831 73237 14970 23129
35483 84563 79956 88618 54619 24853 59783 47537 88822 47227
09262 25041 57862 19203 86103 02800 23198 70639 43757 52064

Source: Adapted from the Table of Random Numbers given at
<http://www.mrs.umn.edu/~sungurea/introstat/public/instruction/ranbox/randomnumbersII.html>

The above random numbers table contains 450 (5 digit) random numbers.

4.4.3 Steps in the Use of RNT

We need to follow the following steps while selecting a SRS by using RNT.

1. Determine the population size (N).
2. Determine the sample size (n).
3. List all the units of the population. Assign the numbers in a serial order. Suppose there are 100 units in the population, assign the serial numbers from 00 to 99.
4. Determine the starting point of selecting the sample by picking up a page from the random number tables and dropping your finger on a number in the page blindly.
 5. Choose the direction in which you want to read the numbers (say from left to right or right to left or top to bottom or bottom to top).

6. Suppose you are looking for two digit numbers (00 to 99) you may not get these numbers by direct reading from the tables since they are 5 digit numbers (see Table 4.1). You can either look at the last two digits or first two digits of the numbers. For example, if the 5 digit number you have chosen is 54245 (that is, the number in the 29th row and 10th column of the random number table given at Table 4.1). Then, the two digit number will be 45 if you chose the last two digits of the number.
7. Look only at the numbers assigned to each population unit. If the number represents one of the unit of the population it becomes part of the sample. Suppose you want to select 10 sample units, the other numbers you will be choosing are 71(11071), 30(44430), 64(94664), 94(91294), 63(35163), 82(32883), 04(23904), 40(41340), 85(61185). Observe that you have omitted 94(05494) since you have already chosen this number.
8. Once a number is chosen, do not select it again.
9. If you reach the end point of the table before obtaining the required sample, pick another starting point in the random number table and select the remaining units for the sample.

Example 4.2

Suppose you have to select 100 account holders as a sample out of total 1000 account holders in the population using random numbers table.

628	126	254	090	752	091	411	146	089	630
831	113	511	082	140	733	076	292	486	468
583	361	047	792	466	395	635	697	447	405
209	404	457	570	194	043	330	939	865	906
409	830	911	767	248	253	317	120	772	103
836	530	785	210	228	869	332	868	672	358
469	149	509	176	169	954	002	582	249	037
192	221	918	437	571	995	006	694	692	581
050	734	652	577	631	184	295	776	885	796
822	973	822	467	013	072	942	917	129	624

Here, first you assign each account holder a number from 000 to 999. To draw a sample of 100 account holders, you need to find 100 three digit numbers in the range 000 to 999. Pick up any row or column in the random numbers table given in Table 4.1. Suppose you have selected the fourth row and first column as starting point to draw the sample, the first digit number is 628(64628) if you chose last 3 digits as the number for your purpose. Here, you read the last 3 digits of the number. If the number is within the range (000 to 999) include the number in the sample. Otherwise skip the number and read the next number in some

identified direction. If a number is already selected omit it. In this example since you have started with fourth row and first column and moving from left to right direction the following 100 numbers are selected for the sample.

If the number of units in the population is very large, neither of the above two methods is feasible. These days by using a computer we can select a random sample in a much easier way. There are many computer program which can generate a series of random numbers if we have the units of the population listed in a computer.

We will explain one way of selecting a sample using computer generated random numbers. In our example, let us assume we can copy and paste the list of account holders into a column in an EXCEL spreadsheet. Then, in the column right to it we paste the function =RAND() which is EXCEL's way of putting a random number between 0 and 1 in the cells. Then, all we have to do is take the first 100 names in the sorted list. The entire process takes a minute if we are familiar with using EXCEL program in computer.

4.4.4 Advantages of SRS

- 1) In simple random sample we assure population units to be homogeneous and thus do not require additional information on the characteristics of the population.
- 2) Using simple random sampling we can select an unbiased sample. This is because, in SRS the chance of including each unit of population in the sample is equal. The bias due to human preferences is completely eliminated.
- 3) Through estimation of sampling error we can assess the accuracy of the results.
- 4) If the population size is not too large, simple random sampling is a simple and easily implementable sampling procedure.

4.4.5 Limitations of SRS

- 1) The greatest limitation of simple random sampling is that if the population size is too large then we need to spend a lot of time in listing the units of the population.
- 2) The simple random sampling procedure will be efficient only when we have a homogeneous population. Suppose we have a population with characteristics such as gender, age, social status, etc. Then, we need a larger sample size to accommodate a representative sample with all those characteristics of the population units. A better way to tackle this issue is to use stratified sampling procedure which you will learn later in this unit.

4.5 SELECTION OF SYSTEMATIC RANDOM SAMPLE

The systematic random sampling procedure is somewhat similar to the simple random sampling procedure. In this sampling procedure, we select a starting point at random and then systematically select the sample units from the population units at a specified sampling interval.

The starting point and the sampling interval are based on the required sample size. The sampling interval will be represented as K . The selection of a sample using systematic random sampling procedure is very simple. Suppose the population consists of n units and you have decided to select a sample of n units using systematic random sampling procedure. Follow the following steps.

1. Number the units in the population from 1 to N (suppose you have 1000 units).
2. Decide the sample size n you need (suppose you want to select 100 units).
3. Determine the sampling interval by dividing the population by the sample size. $K = N/n =$ the interval size (here $K=1000/100 = 10$).
4. Select a unit at random from the first K units (1 to K) (suppose you have selected unit number 5 as your first sample unit).
5. Then select the subsequent sample units by adding K to the previous unit (the subsequent samples will be 15 (5+10), 25 (15+10), ..., 995 (985+10)).

Example 4.3

From a population consisting of 500 units draw a sample of 60 units using systematic random sampling procedure.

To use systematic random sampling, the first thing we need to do is listing of the population units in a random order by giving numbers from 1 to 500. The sampling interval is $K=500/60 = 8.3$ or say 8. Now we select the first sample unit at random from the first 8 population units. Suppose the first unit selected is 5. The subsequent sample units selected are: 13, 21, 29, 37.....477. Therefore, following are the population units selected in the sample.

5	13	21	29	37	45	53	61	69	77
85	93	101	109	117	125	133	141	149	157
165	173	181	189	197	205	213	221	229	237
245	253	261	269	277	285	293	301	309	317
325	333	341	349	357	365	373	381	389	397
405	413	429	429	437	445	453	461	469	477

Thus, in the systematic random sampling procedure, the first sample unit is selected at random and this sample unit in turn determines the subsequent sample units to be selected. However, it is essential that the units in the population are randomly ordered. In certain cases we prefer using systematic random sampling procedure to simple random sampling procedure because it is easier to select sample units. For example, if you want to find out the yield of coconut trees in a field, select a tree at random. Other trees are automatically selected at a gap equivalent to sampling interval.

4.5.1 Advantages of Systematic Random Sampling

- a) The main advantage of using systematic random sample is that the time taken and work involved in systematic random sampling is less than simple random sampling procedure. It is frequently used in exit polls on voting behaviour and obtaining the opinions and views of consumers in marketing research.
- b) The other advantage of systematic random sampling procedure is that this method can be used even when no formal list of the population units is available. For example, suppose if we are interested in knowing the opinion of consumers on improving the services offered by a bank, we may simply choose every k^{th} account holder visiting a bank branch, provided that we know how many account holders are there : (1) For example, there are 2000 account holders in the population and we want to have 200 account holders as sample size. Then, $K=2000/200=10$) and we select every 10^{th} account holder visiting the bank.

4.5.2 Disadvantages of Systematic Random Sampling

- a) The main disadvantage of systematic random sampling procedure is that if there is a periodicity in the occurrence of units of a population, the use of systematic random sampling procedure gives a highly unrepresentative sample. For example, suppose you are interested in obtaining the views/opinions of consumers of a store in your locality. You may arrange all the consumers of the store according to their date of visit and start selecting a sample of customers who visit the store on 1st of every month. You know that the 1st day of every month cannot be representative of the whole month.
- b) The other disadvantage of systematic random sampling procedure is that every unit of the population does not have an equal chance of being selected. Rather the selection of population units in the sample depends on the initial unit of selection. Regardless of how we select the first unit of the sample, subsequent units are automatically determined. This lacks complete randomness.

4.6 SELECTION OF STRATIFIED RANDOM SAMPLE

In some cases the population may not be homogenous, that is, all the units may not be equal with respect to the characteristic we intent to survey. The characteristics of the population under study may be male/female, rural/urban, literate/illiterate, high income/low income groups, etc. In situations where these units vary widely, the simple random sampling procedure or the systematic random sampling procedure will not provide us with a representative sample. In such situations by using stratified random sampling we can obtain a representative sample.

In stratified sampling, we divide the population into different strata in such a way that units are homogenous within each stratum. Moreover, each stratum is different. Suppose we want to stratify the population on the basis of gender distribution then we list the population units separately according to males or females. Subsequently, we decide the sample size to be drawn from each stratum. There are two approaches to decide the sample size from each stratum. These are: (a) proportional stratified sample, and (b) disproportional stratified sample. We will discuss these two procedures below.

4.6.1 Proportional Stratified Sample

When we take a sample from a population with several strata, we are required to take samples from each stratum. Such sample could be in proportion of the stratum population size to the total population size. Suppose we divide the population (N) into K non-overlapping strata $N_1, N_2, N_3, \dots, N_K$ such that $N_1 + N_2 + N_3 + \dots + N_K = N$. We decide to draw a sample of the size n. Then the sample proportions of different strata are given by:

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n_3}{N_3} = \dots = \frac{n_K}{N_K}$$

Example 4.4

Suppose we want to draw a sample of 200 units from a population consisting of 1000 units. The population is heterogeneous in nature in terms of high income or low income and rural or urban. The strata population sizes are given as follows:

High income - urban	= 200
Low income - urban	= 400
High income- rural	= 100
Low income-rural	= 300

To have a representative sample each stratum in the sample should represent the corresponding stratum in the population. For this we should take a different sample size from each stratum depending upon the stratum size. The deciding factor in each of the stratum is same as the proportion of total sample to the

population. In our example, to have a sample of 200 units, the proportion of

sample to the population in each stratum is

$$\frac{n}{N} = \frac{200}{1000} = 0.2$$

Observe that we are considering the same proportion for each stratum. Then the sample from each stratum will be as follows:

Stratum Category	Stratum Population size(N_i)	Sample to population proportion	Stratum sample size
(1)	(2)	(3)	(4)=(2) × (3)
High income - urban	200	0.2	40
Low income -urban	400	0.2	80
High income-rural	100	0.2	20
Low income-rural	300	0.2	60
Total	1000	0.2	200

There are several advantages of stratified sampling over simple random sampling. The stratified sampling ensures sample representation of not only the entire population, but also each stratum. This is important where the stratum size is small. Moreover, stratified sampling generally has more statistical precision than simple random sampling.

4.6.2 Disproportional Stratified Sampling

In proportional stratified sampling, we assumed that each stratum in the population is homogeneous. Consequently, we expect that variability within stratum is lower than the variability for the population as a whole. On the other hand, if the variability within each stratum is not small then we use disproportional stratified sampling. In disproportional stratified sampling, the strata allocation is based on size and variability (that is, the standard deviation of the characteristic under study). In this procedure a larger sample is drawn from the stratum having higher variability. This procedure is also sometimes called double weighing scheme and provides the most efficient sample and most precise/reliable estimates for a given sample size. The only requirement is that we should have knowledge/estimate of the standard deviation of the characteristic under study within each stratum.

Follow the steps given on next page for using disproportional stratified sampling.

- 1) Divide the population into strata based on the chosen characteristic (example, Rural/Urban, Male/Female, etc.)
- 2) The number of units taken from each stratum is directly proportional to the relative size of the stratum and standard deviation σ_i of the characteristic under consideration. Suppose, if $\sigma_1, \sigma_2, \sigma_3, \dots, \sigma_k$ are the standard deviations of k strata and $P_1, P_2, P_3, \dots, P_k$ are the stratum proportions to the total population, and n ($= n_1 + n_2 + \dots + n_k$) is the sample size required. Then the stratum sample size using disproportional stratified sampling procedure is

$$n_i = \frac{P_i \times \sigma_i \times n}{\sum P_i \sigma_i}$$

- 3) Choose the sample from each stratum using either simple random sampling or systematic random sampling.

Let us go back to Example 4.4, where we have divided the population into 4 strata. We observe that there are small number of households in high income strata and large number of households in low income strata. Assume that the variance of income among higher income groups is higher than the variance among the lower income groups. Therefore, in order to avoid under-representation of higher income groups in the sample, a disproportional sample is taken in each stratum. That means, if the variability within the stratum is higher, we must have larger sample size of that stratum to increase the precision of the estimates. Similarly, if the variability within the stratum is lower, we must have smaller sample size of that stratum. That is, higher the stratum variance larger the stratum sample size and lower the stratum variance smaller the sample size. This is in addition to the fact that larger stratum size requires a larger sample size.

Example 4.5

Consider Example 4.4 again. Suppose the stratum variances are given as follows:

Stratum	Variance (σ^2)
High income urban	6.5
Low income urban	2.5
High income rural	4.5
Low income rural	2.0

Use the disproportional stratified sampling procedure to choose a sample of size 200 from the four strata.

For this example the disproportional sample size for each stratum is given on next page.

Stratum	Stratum population	Stratum population proportion (P_i)	Stratum Variance (σ_i^2)	Stratum standard deviation ($\sigma_i = \sqrt{\sigma_i^2}$)	$P_i \times \sigma_i$	Sample size $\frac{P_i \times \sigma_i \times n}{\sum P_i \sigma_i}$
High income-urban	200	0.20	6.5	2.5	0.50	56
Low income-urban	400	0.40	2.5	1.6	0.64	72
High income-rural	100	0.10	4.5	2.1	0.21	24
Low income-rural	300	0.30	2.0	1.4	0.42	47
Total	1000				1.77	200

4.6.3 Advantages of Stratified Sampling

- In stratified random sampling the sample is drawn from each stratum of the population. Therefore, the stratified random sampling procedure is more representative.
- The stratified random sampling procedure is more precise than simple random sampling. Therefore, to a great extent this procedure avoids sample selection bias.
- As we have seen in simple random sampling and systematic random sampling procedures, when there is heterogeneity of population we need to have a large sample size to have a fairly representative sample. However, in stratified random sampling this objective can be achieved with a smaller sample size. This saves a lot of time, money and other resources for data collection.

4.6.4 Disadvantages of Stratified Sampling

- The main disadvantage of stratified random sampling procedure is that we need a detailed knowledge of the distribution of the characteristics in the population. If we cannot accurately identify the homogeneous groups, it is better to use simple random sample since improper stratification can lead to serious error.
- The other disadvantage of stratified random sampling is that we need to prepare a list of population units for each stratum separately. As the list of population units may not be readily available for each characteristic, the preparation of lists may be a very difficult task.

4.7 SELECTION OF A CLUSTER SAMPLE

Very often population units are spread over a vast geographical area. In that case collection of data through simple random sampling requires a lot of time, money

and manpower as we have to cover the entire geographical area for collecting data on the selected units. Imagine taking a sample of respondents spread all over Uttar Pradesh in order to conduct personal interviews. Using simple random sample, the respondents will be spread all over the state and you have to travel and spend lot of money meeting the respondents. In such situation cluster sampling will be much useful.

The basic principles of cluster sampling are:

- i) The differences or variability within a cluster should be as large as possible. As far as possible the variability within each cluster should be the same as that of population.
- ii) The variability between clusters should be as small as possible. Once the clusters are selected, all the units in the selected clusters are included in the sample for obtaining data.

In cluster sampling we divide the population into groups called clusters. Then we select a sample of clusters using a simple random sampling.

The population units in each of the clusters are assumed to be as heterogeneous as those in the total population. That is, each cluster itself is a representative of the population.

4.7.1 Steps in Cluster Sampling

In cluster sampling, we follow the steps given below:

- 1) Divide the population into a number of clusters.
- 2) Determine the number of clusters needed for your sample.
- 3) Randomly select the sample of clusters.
- 4) Survey all units within the sampled clusters.

Suppose the division of clusters is based on the geographical boundaries of the population, then it is called *area sampling*. You have observed that in the case of cluster sampling the clusters are selected using random sampling method. Subsequently all the population units within each sampled cluster are included in the sample. Suppose instead of including all the population units within each selected cluster you chose to include only a sample of units within each cluster. Then you can clearly understand that there are two stages.

In the first stage you select the clusters and in the second stage you select the sample units within each sampled clusters. This sampling procedure is called *two-stage sampling*. Here, the clusters are called primary units and the units within the sampled clusters are called secondary units.

Example 4.6

Suppose we are interested in finding the options of ATM customers of a Bank in Uttar Pradesh state. We can divide the state into say 30 clusters (for example, we

can consider district as a unit and include one or two districts in one cluster). Here, we assume that each of these clusters will represent the opinions of the ATM customers of Uttar Pradesh as a whole. We then select a sample of clusters and obtain the opinion of all the ATM customers in each of the cluster.

4.7.2 Advantages of Cluster Sampling

- a) The main advantage of cluster sampling is that it takes less travel time and related data collection costs.
- b) Since the researcher need not cover all the clusters and only sample of clusters are covered, it is a more practical method which facilitates fieldwork.

4.7.3 Disadvantages of Cluster Sampling

- a) In cluster sampling we assume that each cluster represents the heterogeneity of the population units of all clusters. However, this assumption may not be true in many cases, because often the tendency is that the units in the clusters are more homogeneous than the units of the entire population. That means it is difficult to form heterogeneous clusters.
- b) The cluster sampling has a lower sampling efficiency for a given sample size than random sampling and stratified sampling. This method is cost effective but not statistically efficient.

4.8 MULTISTAGE SAMPLING

We have seen in cluster sampling that when we select a sample instead of covering all the units from each cluster, we call it two-stage sampling. The multistage sampling is an extension of two-stage sampling.

The four methods we have covered so far, namely, (a) simple random sampling, (b) systematic random sampling, (c) stratified sampling, and (d) cluster sampling are the simplest probability (or random) sampling procedures. However, in real-life, we use sampling methods that are more complex than the above four methods. The basic principle in multistage sampling is that we can combine these simple methods in a variety of useful ways to address our sampling needs. We call it multistage sampling when we combine two or more of the above sampling methods.

Example 4.7

Let us consider the case of interviewing school students in Haryana in order to grade the schools according to socio-economic background of the parents. For this problem, in the first stage we need to apply cluster sampling. We divide the state of Haryana into a number of clusters, say districts. Then we select a sample of districts (clusters) using simple random sampling method. In the second stage we

divide the schools using stratified sampling. Here the strata may be government schools, government-aided schools, central schools, and public schools. We select a sample of schools in each stratum using either a simple random sampling or a systematic random sampling. In the third stage we again use simple random sampling and select a sample of classes in each sampled school for face-to-face interviews with the students. In the fourth stage of sampling we consider selecting a sample of students from each sampled class using simple random sampling or systematic random sampling.

In multi stage sampling it is possible to consider as many stages as necessary to achieve a representative sample. In each stage a suitable method of sampling is used. Each stage results in a reduction of the sample size.

Advantages

- a) Multistage sampling procedure provides cost gains by reducing the data collection costs.
- b) Multistage sampling is more flexible and allows us to use different sampling procedures in different stages of sampling.
- c) If the population is spread over a very wide geographical area, multistage sampling is the most appropriate sampling method.

Disadvantages

If the sampling units selected at different stages are not representative, multistage sampling becomes less precise and less efficient.

Check Your Progress 1

- 1) Which of the following is a procedure of selecting samples from a population?
 - a) Random sampling
 - b) Non-random sampling
 - c) Stratified sampling
 - d) All the above
- 2) Suppose you are applying a stratified random sampling procedure on a population. How do you make your sample selection?
 - a) Select at random an equal number of units from each stratum
 - b) Draw equal number of units from each stratum and weigh the results
 - c) Select the sample at random from each stratum proportional to the population
 - d) b. and c. both
 - e) a. and c. both

- 3) State whether the following statements are 'True' or 'False'.
- A sampling procedure that selects units from a population at uniform intervals is called simple random sampling.
 - A sampling procedure that divides the population into well-defined groups from which random samples are drawn is known as stratified sampling.
- 4) A population is made up of groups that have wide variation between groups but little variation within each group. In this situation the appropriate type of sampling procedure to use is
- Cluster sampling
 - Systematic sampling
 - Stratified sampling
 - Multistage sampling

4.9 NON-PROBABILITY SAMPLING PROCEDURES

There are four different types of non-probability sampling such as

- Convenience Sampling
- Judgment Sampling
- Quota Sampling
- Snowball Sampling

We discuss the procedure of drawing a non-probability sampling below.

4.9.1 Convenience Sampling

This is one of the most commonly used methods of non-probability sampling. In this method the researcher's convenience forms the basis for selection of the sample. Especially for an exploratory research there is a pressing need for data. In such situations the selection of sampling units is left to the interviewer. The population units are included in the sample simply because they are in the right place at the right time. This method is often used during preliminary research efforts to get a gross estimate of the results, without incurring the cost or time required to select a sample. For example, during the budget session or when the price of a product is increased or a new government is formed, convenience samples are used by the researchers/journalists to reflect public opinion. Convenience samples are extensively used in marketing research.

The advantage of convenience sampling is that it is less expensive and less time-consuming. The limitations of convenience sampling are: (a) it involves sample

selection bias, and (b) it does not provide a representative sample of the population and therefore we cannot generalise the results.

4.9.2 Judgment Sampling

This is another commonly used non-probability sampling procedure. This procedure is often referred to as *purposive sampling*. In this procedure the researcher selects the sample based on his/her judgment. The researcher believes that the selected sample elements are representative of the population. For example, the calculation of consumer price index is based on judgment sampling. Here the sample consists of a basket of consumer items and other goods and services which are expected to reflect a representative sample.

The prices of these items are collected from selected cities that are viewed as typical cities with demographic profiles matching the national profile.

The advantage of judgment sampling is that it is low cost, convenient and quick. The disadvantage is that it does not allow direct generalisations to population. The quality of the sample depends upon the judgment of the researcher.

4.9.3 Quota Sampling

In this procedure the population is divided into groups based on some characteristics such as gender, age, education, religion, income group, etc. A quota of units from each group is determined. The quota may be either proportional or non-proportional. The proportional quota sampling is based on the proportion of each characteristic in the population so that the proportion in the sample represents the population proportion. For example, if you know that there are 80% of the households whose income is below say Rs.100000 per annum and 20% households have income above Rs.100000 per annum in a city. You want to take a sample of size 100 households. Then you include 80 households from below Rs.100000 income and 20 households from above Rs.100000 income. The objective here is to meet the proportional quota of sampling from each characteristic in the population.

The non-proportional quota sampling is a bit less restrictive. In this procedure, you specify the minimum number of sampled units from each group. You are not concerned with having proportions in the population. For instance, in the above example you may simply interview 50 households from each income group instead of 80 per cent and 20 per cent. The interviewer is instructed to fill the quota for each group based on convenience or judgment. The very purpose of quota sampling is that various groups in the population are represented to the extent the investigator desires.

Do not confuse the quota sampling with stratified sampling that you have learned earlier. In stratified sampling you select random samples from each stratum or group whereas in quota sampling the interviewer has a fixed quota. For example, in a city there are five market centres. A company wants to assess the demand for its new product and sends 5 investigators to assess the demand by interviewing

50 prospective customers from each market. It is left to the investigator whom he/she will interview at each market centre. If the product is targeted to women, this way you cannot elicit the information among various groups of women customers like housewives or employed women or young or old. In this sampling you are simply fixing a quota for each investigator.

The quota sampling has the advantage over others if the sample meets the characteristics of the population that you are looking into. In addition, the cost and time involved in collecting the data are greatly reduced. However, there are many disadvantages as well. In quota sampling, the samples are selected according to the convenience of the investigator instead of selecting random samples. Therefore, the selected samples may be biased.

If there are a large number of characteristics on the basis of which the quotas are fixed, then it becomes very difficult to fix the quotas/sub-quotas for each group/sub-group. Also the investigators have the tendency to collect information only from those who are willing to provide information and avoid unwilling respondents.

4.9.4 Snowball Sampling

In snowball sampling, we begin by identifying someone who meets the criteria for inclusion in our study. We then ask him/her to recommend others who also meets the criteria. Although this method would hardly lead to representative samples, there are times when it may be the best method available. Snowball sampling is especially useful when we are trying to reach populations that are inaccessible or hard to find. For example, if we are studying the homeless, we are not likely to find good lists of homeless people within a specific geographical area. However, if we go to that area and identify one or two, we may find that they know very well who the other homeless people in their vicinity are and how we can find them.

4.10 DETERMINATION OF THE SAMPLE SIZE

The use of appropriate sampling procedure is necessary for a representative sample. However, this condition is not sufficient. In addition to the above, we should determine the sample size. The question of how large a sample should be is a difficult one. Sample size can be determined by various considerations. The following are the some of the considerations in determining the sample size:

- (i) Sampling error
- (ii) Number of comparisons to be made
- (iii) Response rates
- (iv) Funds available

Sampling Error: We know that smaller samples have greater sampling error than large samples. On the other hand, larger samples have larger non-sampling errors than smaller samples. The sampling error is a number that describes the precision of an estimate of the sample. It is usually expressed as a margin of error

associated with a statistical level of confidence. For example, for a prime minister preferential poll you may say that the incumbent is favored by 65% of votes, with a margin of error (precision) of plus or minus 5 percentage points at a 95% confidence level. This means that if the same surveys were conducted with 100 different samples of voters, 95 of the surveys would be expected to show the incumbent favoured by between 60% and 70% of the voters ($65\% \pm 5\%$). Remember as you increase the precision level of your results you need larger sample size.

Number of Comparisons to Make: Sometimes we may be interested in making comparisons of two or more groups (strata) in the sample. For example, we may want to make the comparison between male and female respondents or between urban and rural respondents. Or we may want to compare the results for 4 geographical regions of the country say north, south, west and east. Then we need an adequate sample size in each region or stratum of the population. Therefore, the heterogeneity of population characteristics plays a significant role in deciding the sample size.

Response Rates: In mail surveys, we know that all those questionnaires mailed to the respondents may not reach us back after filling the questionnaires. As per the experiences on mail survey, the response rate ranges between 10% and 50%. Then, if you are expecting a 20% response rate, for example, you will have to mail 5 times the number of sample size required.

Funds Available: The funds available may influence the sample size. If the funds available for the study are limited then you may not be able to spend more than a certain amount of the total money available with you on collecting the data.

It is even more difficult to decide the sample size, when you use the non-probability sampling procedures. This is because there are no definite rules to be followed in non-probability sampling procedures. It all depends upon on what you want to know, the purpose of inquiry, what will be useful, what will have credibility and what can be done with available time and resources. In purposive sampling, the sample should be judged on the basis of purpose. In non-probability sampling procedures, the validity, meaningfulness, and insights generated have more to do with the information-richness of the sample units selected rather than the sample size.

Some Formulae to Determine the Sample Size

Technical considerations suggest that the required sample size is a function of the precision of the estimates you wish to achieve, the variance of the population and the confidence level you wish to use. If you want more precision and confidence level then you may need larger sample size. The more frequently used confidence levels are 95% and 99%. And the more frequently used precision levels are 1% and 5%. There are different formulae used to determine the sample size depending upon various considerations discussed above. In this section we will discuss three of them.

If we wish to report the results as percentages (proportions) of the sample responding, we use the following formula:

$$n_i = \frac{P_i(1-P_i)}{\frac{A^2}{Z^2} + \frac{P_i(1-P_i)}{N_i}}$$

Where, n_i = sample size of the i^{th} attribute required

P_i = estimated proportion of the population possessing i^{th} attribute of interest (for example, proportion of males, females, urban, rural, etc.)

A = precision required (0.01, 0.05 etc.)

Z = standardized value indicating the confidence level ($Z=1.96$ at 95% confidence level and $Z=2.58$ at 99% confidence level)

N_i = population size of the i^{th} attribute (known or estimated)

Example 4.8: A population consists 80% rural and 20% urban people. Given that the population size is 50000, determine the sample size required. Assume that the desired precision and confidence levels are 1% and 99% respectively.

In this example,

P_1 = proportion of rural people = 0.80

P_2 = proportion of urban people = 0.20

N_1 = rural population size = $50000 \times 0.80 = 40000$

N_2 = urban population size = $50000 \times 0.20 = 10000$

$A = 0.01$

$Z = 2.58$ (at 99% confidence level)

The required sample size is

$$\begin{aligned} n_1 = \text{rural sample} &= \frac{P_1(1-P_1)}{\frac{A^2}{Z^2} + \frac{P_1(1-P_1)}{N_1}} \\ &= \frac{0.80(1-0.80)}{\frac{0.01^2}{2.58^2} + \frac{0.80(1-0.80)}{40000}} \\ &= \frac{0.80(0.20)}{\frac{0.0001}{6.6564} + \frac{0.80(0.20)}{40000}} \\ &= \frac{0.16}{0.000019 + \frac{0.16}{40000}} \\ &= \frac{0.16}{0.000019 + 0.000004} \end{aligned}$$

$$= \frac{0.16}{0.000023} = 8410.8 \text{ or say } 8411$$

$$\begin{aligned} n_2 = \text{urban sample} &= \frac{P_2(1-P_2)}{\frac{A^2}{Z^2} + \frac{P_2(1-P_2)}{N_2}} \\ &= \frac{0.20(1-0.20)}{\frac{0.01^2}{2.58^2} + \frac{0.20(1-0.20)}{10000}} \\ &= \frac{0.20(0.80)}{\frac{0.0001}{6.6564} + \frac{0.20(0.80)}{10000}} \\ &= \frac{0.16}{0.000019 + \frac{0.16}{10000}} \\ &= \frac{0.16}{0.000019 + 0.000016} \\ &= \frac{0.16}{0.000035} = 4568.4 \text{ or say } 4568 \end{aligned}$$

Therefore, we need to have a sample of size $8411+4568 = 12979$ units.

If we wish to report the results as means (averages) of the sample responding, we use the following formula:

$$n_i = \frac{P_i^2}{\frac{A^2}{Z^2} + \frac{P_i^2}{N_i}}$$

Where, n_i = sample size of the i^{th} attribute required

P_i = estimated standard deviation of the i^{th} attribute of interest (for example, average income of high income group, low income group etc.)

A = precision required (0.01 or 0.05 as the case may be)

Z = standardized value indicating the confidence level ($Z=1.96$ at 95% confidence level and $Z=2.58$ at 99% confidence level)

N_i = population size of the i^{th} attribute (known or estimated)

Example 4.9: It is planned to conduct a study to know the average income of households. Given that the standard deviation of households is 2.5 and the population size is 10000, determine the sample size required. Assume that the desired precision and confidence levels are 5% and 95% respectively.

In this example,

P_1 = standard deviation of income = 2.5

N_1 = number of households = 10000

$$A = 0.05$$

$$Z = 1.96 \text{ (at 95\% confidence level)}$$

The required sample size is

$$\begin{aligned} n_1 &= \frac{P_1^2}{\frac{A^2}{Z^2} + \frac{P_1^2}{N_1}} \\ &= \frac{2.5^2}{\frac{0.05^2}{1.96^2} + \frac{2.5^2}{10000}} \\ &= \frac{6.25}{\frac{0.0025}{3.8416} + \frac{6.25}{10000}} \\ &= \frac{6.25}{0.000651 + 0.000625} \\ &= \frac{6.25}{0.001276} = 4898 \end{aligned}$$

If we wish to report the results in a variety of ways or we have the difficulty in estimating the proportion or standard deviation of the attribute of interest, we use the following formula:

$$n = \frac{0.25}{\frac{A^2}{Z^2} + \frac{0.25}{N}}$$

where, n = sample size required

A = precision required (0.01 or 0.05 as the case may be)

Z = standardized value indicating the confidence level (Z=1.96 at 95% confidence level and Z=2.56 at 99% confidence level)

N = population size (known or estimated)

Example 4.10: Given that the population size is 10000, determine the sample size required when desired precision and confidence levels are 5% and 99% respectively.

In this example,

$$N = 10000$$

$$A = 0.05$$

$$Z = 2.58 \text{ (at 99\% confidence level)}$$

The required sample size is

$$n = \frac{0.25}{\frac{0.05^2}{2.58^2} + \frac{0.25}{10000}}$$

$$n = \frac{0.25}{\frac{0.0025}{6.6564} + \frac{0.25}{10000}}$$

$$n = \frac{0.25}{0.0003756 + 0.000025} = \frac{0.25}{0.000401} = 624$$

Check Your Progress 2

- 1) Say whether the following statements are 'True' or 'False'.
 - a) When the units included in the sample are based on judgment of the investigator, the sampling is said to be random.
 - b) With increasing sample size the sampling error decreases.
 - c) Convenience sampling has the disadvantage that it may not be representative sample.
- 2) One of the major disadvantage of judgment sampling is
 - a) The procedure is very cumbersome
 - b) The sample selection depends on the individual judgment of the investigator
 - c) It gives small sample size
 - d) It is very expensive

4.11 LET US SUM UP

The most commonly used probability sampling procedure is the simple random sampling which allows a chance to all population units to be included in the sample. The sample units are chosen using random number tables. A systematic random sample uses the first sample unit at random as a starting point and the subsequent sample units are chosen systematically. A stratified sample guarantees inclusion of units from each stratum. A cluster sample involves complete enumeration of one or more randomly selected clusters.

The non-probability sampling procedures include convenience sampling, judgment sampling, quota sampling and snowball sampling. These sampling procedures are not independent from sampling bias but still popular in some situations particularly marketing research.

A number of factors decide the sample size. It may be the number of groups in the population, the heterogeneity of population, funds and time available, etc.

Using a sample saves a lot of money, time and manpower. If a suitable sampling procedure is used in selecting units, appropriate sample size is selected and necessary precautions are taken to reduce sampling errors, then a sample should yield a valid and reliable information about the population.

4.12 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

- 1) d
- 2) c
- 3) a) False
b) True
- 4) c

Check Your Progress 2

- 1) a) False
b) True
c) True
- 2) b

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 5 CENSUS AND SURVEY*

Structure

5.0 Objectives

5.1 Introduction

5.2 Census

5.2.1 What is Census?

5.2.2 Uses of Census Data

5.3 Survey

5.3.1 What is Survey Research?

5.3.2 Advantages and Disadvantages of Survey Research Method

5.4 Comparison between Census and Survey

5.5 Let Us Sum Up

5.6 References

5.7 Answers to Check Your Progress

5.0 OBJECTIVES

After going through this unit, you should be able to:

- explain what census and survey are;
- discuss advantages and disadvantages of census;
- discuss advantages and disadvantages of survey; and
- compare census with survey.

5.1 INTRODUCTION

Census and survey are both research methods which outline the process of data collection (for analysis) in a research. In the beginning of this Unit, we will learn what census is and how census data are utilized. This will be followed with an understanding of what and how survey is and how it is utilized in social research. According to United Nations, the census of population is "the total process of collecting, compiling and publishing demographic, economic and social data pertaining, at a specified time or times to all persons in a country or delimited territory. United Nation document–Principles and Recommendations for National Population Censuses (ST/STAT/SERM/27). When translated in terms of

* Written by Indrani Mukherjee, Research Scholar, Department of Anthropology, University of Delhi, Delhi

everyday research, the ‘population’ of study is defined by the research interest/research question. A population is any complete group with at least one characteristic in common. Once the population of the study has been defined, data is systematically obtained every element/individuals of the population across common predetermined parameters. A survey, on the other hand is a method of gathering information from a sample of elements/people, considered to be a representative of the research population, traditionally with the intention of generalizing the results to the larger population. The sampling in a survey is often based on a census of the research population, as a baseline for drawing a representative sample.

Both census and survey recognize the universe of study which defines the population under consideration. They simultaneously use various tools and techniques of data collection like questionnaires, interview schedules, interviews, and non response follow-up techniques etc., both need to process and edit the data, and the susceptibility to various sources of error.

5.2 CENSUS

Let us discuss certain basic issues related to census.

5.2.1 What is Census?

Census refers to the complete enumeration of a universe. A universe may be a place or a specific locality, a group of people, or objects/ elements of interest from which data is collected. The population of a census can vary according to the focus or need of the research. A census is an attempt to list all elements/units/members in a group and to measure one or more characteristics of those elements/units/members. The most common are census of with national population, housing census other types of census include agriculture, animal, business, farms, cars from an assembly line, traffic censuses and so on. Census is a quantitative research method and the data collected is analyzed with the use of statistics.

Box 5.1 Census Moment

Census in most countries of the world is conducted in years ending in 0 or 1. In India, the census is conducted in years ending in 1. International comparability is thus maintained. Usually, a day as well as “census moment” is fixed for the census. In India, the “census moment” now is the sunrise of 1st March, of the census year. The fixing of the “census moment” helps the enumerators to decide which persons are to be included in the census. The persons born after the census moment or the persons dying before the census moment are to be excluded from the census.

Source: https://censusindia.gov.in/Data_Products/Library/Indian_perceptive_link/Census_Objectives_link/censusobjectives.htm

While the size of the population under consideration is dependent on area of research and can thus vary (in size), being a quantitative method census usually

deals with large numbers. One of the major advantages of census method is accuracy as each and every unit of the population is studied before drawing any conclusions of the research. When more and more data are collected the degree of correctness of the information also increases. Thus the results based on this method are also expected to be less biased. Data collected through census method allows opportunity for intensive study and analysis due to the enormity of the data collected.

Today, census is well recognised for its advantages for maintaining national level demographic figures, which are comparable on a periodic scale, regionally, nationally as well as internationally. Let us try to understand census better using the example of population census. A population census covers each person present and/or residing within its scope, without omission or duplication in the order to ensure completeness and accuracy of census data. The total population enumerated refers to one well defined point or period of time. The population figures have no meaning unless they refer to a well-defined territory. The territory covered, along with any changes in its area in successive censuses thus needs to be clearly and explicitly stated. A census can provide detailed information on all or most elements in the population, thereby enabling totals for rare population groups or small geographic areas. Thus a population census is the process of collecting, compiling, analyzing and disseminating demographic, social, cultural and economic data relating to all persons in the country, at a particular time in regularised intervals.

Box 5.2 History of Census in India

The history of census in India is compiled in Drop-in Article on Census- No. 5. The following account is drawn from the above mentioned source available at: [censusindia.gov.in/Ad-Campaign/drop-in-articles/05-History of Census-in-India/pdf](https://censusindia.gov.in/Ad-Campaign/drop-in-articles/05-History%20of%20Census-in-India/pdf))

India has a long tradition of conducting census every ten years. The mention of some kind of population count during 800-600 BC is found in the Rig Veda. Written about 321-296 BC, Kautilya's Arthashastra laid stress on Census taking as a measure of state policy for the taxation purpose. The administrative report 'Ain-e-Akbari' offered extensive data on population, industry, wealth etc. during the reign of Mughal ruler Akber the Great. In the year 1830, Henry Walter in Dacca conducted the first complete census of an Indian city. In this census following characteristics were included: sex, broad age group and houses with their amenities. The census of 1881 undertaken on February 17, 1881 by W.C. Plowden, Census Commissioner of India was a major step towards modern census. The first census of India after Independence was conducted in 1951 which was seventh in the series. In the present day, Census is conducted by the office of the Registrar General and Census Commissioner, India, Ministry of Home Affairs, Government of India.

To conduct a population census, vast organisation and considerable resources (financial, logistical, managerial etc.) are needed. The census organisation has to

mobilize extensive administrative machinery with adequate legislative authority. This can be done only by National Government with the cooperation of States and Local Governments.

5.2.2 Uses of Census Data

Modern census data are commonly used for research and future projections; administrative policy and planning; business and marketing, and planning as a baseline for designing sample surveys by providing a sampling frame. Census also brings out the heterogeneity in a population this helps in understanding the stratification of the population into subgroups, as well as recognizing the weightage that each group deserves based on related parameters.

Let us take a look at how census information can be utilized by taking the example of the Indian population census. As discussed earlier, the census in India is conducted every 10 years, Office of the registrar general and the Census commissioner of India, (https://censusindia.gov.in/Data_Products/Library/Indian_perceptive_link/Census_Objectives_link/censusobjectives.htm), provides a detailed description of the objectives, essentials and utility of census, let us look at some of these in terms of the ways in which the information collected is utilized.

- (i) **Projections and Research:** Country-wide data on in terms of the urban-rural areas, geographic dispersion according to occupation, education, sex and age, economic characteristics and, social structure of population etc. analysis of the composition, distribution, and growth of the population accurate (both over a period of time as also projection for future). The changing patterns of various population distributions are provided by population census. Such appraisal is important for research and practical problems of national, industrial and commercial growth and management.

Census data serve as the baseline for computing various indicators. For example, census results, time-adjusted by migration statistics, can provide estimates of the future size, distribution and other characteristics of the population of the total country and sub-national areas. Or census data on fertility can provide a bench-mark check on the reliability of current birth statistics. Thus, census data can be utilized for making projections for the future as well as crosschecking the previous projections with factual reality.

- (ii) **Administration, Policy and Planning:** One of the most basic of the administrative uses of census data is in the demarcation of constituencies and the allocation of representation on governing bodies. Detailed information on the geographic distribution of the population is indispensable for this purpose.

Demographic and economic characteristics of the population collected through census are essential for evaluation of economic and social problems, which must precede the determination of policy affecting economic and social development.

Businesses and industries can effectively use census data to make reliable estimates of consumer demand for variety of goods, commodities and services, based on the population or social group that they cater to. The information from census feeds into operational research providing for demand related to housing, furnishing, clothing, recreational facilities, medical supplies and so forth. Further, the local availability of labour for production and distribution of commodities can be easily determined by the use of census data.

- (iii) **Recognising the heterogeneity of the population:** Census provides the official counts used to assign the number of elected representatives to regions. The social and cultural data collected in the census is employed to determine the total number of seats to be reserved for women, members of scheduled castes, scheduled tribes and other socially marginal groups in the House of People and the Legislative Assemblies of the States.
- (iv) **Census as frame for Sample Surveys:** Size of populations as also other characteristics of the population. This makes its necessary to conduct regular inter-census sample survey. These surveys provide updated and detailed data on many characteristics that are usually investigated in population census conducted intervals of ten years.

Box 5.3 Essential features of Census

- It is a quantitative research method.
- It advocates a systematic, standardized data collection based on predetermined parameters.
- It is defined by a specific universe (which can be a territorial region, a group of people, or objects or elements like crops, vehicles etc.).
- Census is defined by a point or period of time.
- It covers every element of the universe or population of study, without omission or duplication.

The census can provide the frame for scientific design of surveys, by pointing towards the size of a representative sample, the strata that might be considered for the sample (economic categories like household income, property and asset, land holding or social categories like gender, age, caste, class, identity and so on,) etc. The results from a survey can be cross checked and compared with census computations provided the parameters and definitions of enquiry are in alignment with each other. A sample might utilize the same units of measure as utilized by the census for the purposes of uniformity.

Check Your Progress 1

- 1) What is meant by census?

.....
.....

2) Outline four ways in which population census is utilized in India.

Census and Survey

.....
.....
.....
.....

5.3 SURVEY

Survey is a method of collecting information from a sample of the population. A sample is a subset of units in a population, selected to represent all units in a population of interest. It is a partial enumeration because it is a count from part of the population. Information from the sampled units is used to estimate the characteristics for the entire population of interest. Surveys come in many different forms and have a wide variety of purposes. Unlike a census, they gather information from only a small sample of people (or farms, businesses or other units, depending on the purpose of the study). Information is collected by means of standardized questions so that every individual/unit surveyed responds to and is evaluated for exactly the same question. Data is gathered to study demographic information, knowledge, opinions, behavior, attitudes, habits, desires, values and beliefs etc. Survey as a method seeks to identify principles on the sample design, data collection instruments, statistical adjustment of data and data processing. A survey can be used for both quantitative and qualitative research.

5.3.1 What is Survey Research?

Traditionally, surveys are recognised as a part of quantitative research, and are vastly used as such till date. In quantitative research the sample is scientifically chosen so that each individual in the population has a known chance of selection. In this way, the results can be reliably projected to the larger public. Thus the sample must be comprehensive in its design and extensive enough to provide a reliable representation of the whole population.

The objective is not to describe the particular individuals who by chance are part of the sample, but to obtain a statistical profile of the population. Individual respondents are not identified. Survey's results are presented in the form of summaries, such as statistical tables and charts.

A survey can be used to acquire demographic and socio-economic profile of a population, however it can also be used for knowledge, opinions, behavior, and attitudes etc., however it misses out on personal feelings or thoughts on the situation.

5.3.2 Advantages and Disadvantages of Survey Research

Survey research has the following advantages:

- (i) It is inexpensive and less time consuming with in-depth fieldwork-based research.

- (ii) It has a wide coverage (i.e. a large section of population can be covered through survey research) Also, survey research can be carried out in remote locations through e-mails, telephone etc.
- (iii) Many questions can be asked through a survey.
- (iv) Since the same set of question is asked to the whole sample, it is possible to draw comparisons.

Survey research has the following disadvantages:

- (i) Sometimes, it is difficult to frame question that are relevant to/ can be answered by everyone included in the sample.
- (ii) It is in- flexible in the sample.
- (iii) These is a likelihood that respondents do not answer all the question or not provide reliable data.
- (iv) Socio-economic context is often ignored.

5.4 COMPARISON BETWEEN CENSUS AND SURVEY

Census and Survey are both research methods that guide the collection of data in a research process. While census enumerates every unit of the population, survey is based on a sample or subset of the population (based on the parameters of the research). As survey (especially in quantitative research) works through a representative sample of the population, census often forms the baseline for designing sample surveys by providing a sampling frame. They use standardized patterns of information collection to facilitate comparative analysis; future projections; administrative policy and planning; business marketing, and planning. Surveys can also provide information on opinions, behavior, attitudes, habits, desires, values and beliefs etc, and are used in both quantitative and qualitative research. Census usually needs a large amount of resources in terms of administration, finance, logistics, manpower and time. In comparison with census, a survey based on sample collection (for the same research) is less resource intensive. Both census and survey are valuable research methods, and their use needs to be dictated by the need of the research.

Sl. No.	Census	Survey
1	Census is used for quantitative research.	Survey is used in both quantitative and qualitative research.
2	Complete enumeration: Census is a systematic method that requires collecting recording of data about all the members of the population.	Partial enumeration: Survey considers a subset of the population selected to represent the entire group, in all its characteristics.

3	Census is best suited for a population of heterogeneous nature.	Survey is better suited for homogeneous social groups.
4	As census accounts for every unit of the population, the data when analyzed has greater accuracy and reliability.	There are predictable chances of errors in the results drawn from the sample
5	Benchmark data may be obtained for future studies and baseline for drawing sample for surveys.	Interim surveys can be utilized to draw enhanced statistical projections from census data
6	Census leads to numerical enumerations of the population.	Survey can gather information on opinions, behaviour, attitudes, habits, desires, values and beliefs
7	Census requires considerable resources (administrative, financial, logistical, managerial etc.)	A survey requires comparatively less resources as the enumeration of a sample is easier than the whole.
8	Census is time consuming.	A survey requires comparatively less time.

Check Your Progress 2

1) What is meant by a survey?

.....

.....

.....

.....

.....

2) Fill in the blanks

- (i) Survey considers a of the population.
- (ii) Survey is better suited for population.
- (iii) Voting behaviour can be studied through method.

5.5 LET US SUM UP

In this Unit we have learnt about the research methods ‘census’ and ‘survey’. We looked at how census is utilized in the Indian context to understand its operationalisation. We consolidated this information to understand the properties, advantages and disadvantages of census. We continued on our journey to learn

about survey and how it can be used as for both quantitative and qualitative research, finally closing with a comparison between census and survey.

5.6 REFERENCES

- Creswell, John W. 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd edition Sage: Los Angeles.
- Freedman, David, Robert Pisani, and Roger Purves. 2007. *Statistics*. 4th edition. W. W. Norton & Company: New York.
- Given, Lisa M. 2008. *The Sage Encyclopedia of Quantitative Research Methods*. Sage Publications: Los Angeles.
- Jansen, Harrie. 2010. The Logic of Qualitative Survey Research and its Position in the Field of Social Research Methods. *Forum Qualitative Social Research*, Vol. 11(2).
- Office of the registrar general and the Census commissioner of India, GOI (https://censusindia.gov.in/Data_Products/Library/Indian_perceptive_link/Census_Objectives_link/censusobjectives.htm)
- Triola, Mario F. 2001. *Elementary statistics*. 8th edition Addison-Wesley: Boston.
- Weeks, John R. 2008. *Population: An introduction to concepts and issues*. 4th edition Thomson Wadsworth: Belmont, CA.
- Drop-in Article on Census-No.5 available at: www.censusindia.gov.in/Ad-Campaign/drop-in-articles/05-History-of-Census-in-India.pdf

5.7 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

- 1) Census or complete enumeration is a systematic (research) method of collecting and recording data of all units/members of the (research) population. It is best suited for a population of heterogeneous nature and uses statistics to provide quantitative or numerical enumerations of the population. As census accounts for every unit of the population, the data when analyzed has greater accuracy and reliability.
- 2) Population census is utilized in India in the following ways:
 - a. Projections and Research: The population census provides indispensable data for scientific analysis and appraisal of the composition, distribution and past and prospective growth of the population.
 - b. Administration, Policy and Planning: Detailed information on the geographic distribution of the population is used for demarcation of constituencies and the allocation of representation on governing bodies

- c. Recognising the heterogeneity of the population: Social and cultural data collected in the census is employed to determine the total number of seats to be reserved for women, members of scheduled castes, scheduled tribes and other socially marginal groups in the House of People and the Legislative Assemblies of the States.
- d. Census as frame for Sample Surveys: Census forms the basis for sampling to inter-census informational updates.

Check Your Progress 2

- 1) Survey is a method of collecting information from a sample/subset of the population. It is a partial enumeration where information from the sampled units is used to estimate the characteristics for the entire population of interest. Survey can be used for both quantitative and qualitative research.
- 2) (i) sample
(ii) homogeneous
(iii) survey

UNIT 6 OBSERVATION*

Structure

- 6.0 Objectives
- 6.1 Introduction
- 6.2 What is Observation?
- 6.3 Types of Observation
- 6.4 Studies Based On Participant Observation
 - 6.4.1 Argonauts of the Western Pacific
 - 6.4.2 Street Corner Society
- 6.5 Merits and Demerits of Observation Technique
- 6.6 Let Us Sum Up
- 6.7 References
- 6.8 Answers to Check Your Progress

6.0 OBJECTIVES

After going through this Unit, you should be able to:

- explain observation as a technique of data collection;
- outline the types of observation , their uses and limitations; and
- discuss the advantages and disadvantages of observation technique.

6.1 INTRODUCTION

Whenever we visit the marketplace or workplace or travel around, we typically see objects, people, events and much more. Most of the time, we must see them and carry on with our work. Sometimes, however, we look closely and carefully with a purpose in mind. This could be because of our interest or some other reason. When we see something closely and carefully with a purpose in mind, we are not just seeing but observing an object, person, group or event. In this Unit, we will see how the technique of observation is used in the social science research. We will begin with understanding the nature and scope of observation in social research and then look into the types of observation. We will also discuss the advantages and limitations of observation as a technique in research.

* Written by Shaily Bhashanjaly, North Cap University, Gurugram

6.2 WHAT IS OBSERVATION?

Observation has to be distinguished from everyday seeing. Observation is a significant technique of social research. It which enables researchers to understand people and their behavior in their natural setting and arrive at meaningful inferences. Unlike other research techniques, in observation the researcher witnesses events or phenomena first hand i.e., he/she as it exists in real life situation. . The researcher relies upon his/her sense impressions in interpreting the event or phenomena. Observation may serve a variety of research purposes, it may be used to explore the given area of subject matter or to gain insight into the research problem.

Far from being simple and easy to use, the technique of observation comprises standardized procedures and established practices failing which inferences could be erroneous. The hypothesis could guide observation. For example, if we are interested in the causes of highway accidents and have tentatively formulated an exploratory hypothesis that accidents are caused by over speeding, we focus our attention on our concepts of speeding vehicles and the consequences. Research objectives determine the direction and stresses the facts upon which to concentrate. However, the observer must be sensitive to wholly unanticipated and chance phenomena which may come to play an important role in the research process. Let us consider the example of highway accidents again. Now, while observing speeding vehicles, the researcher could also observe that most of the speeding vehicles are driven by youth in the age group of 20-25 years. The additional information regarding age of drivers of speeding vehicles could enrich the analysis.

Observation in research process differs from other techniques of data collection in four ways: (i) observation is always direct while other methods could be direct or indirect; (ii) field observation takes place in a natural setting; (iii) observation tends to be less structured; and (iv) it makes only the qualitative (and not the quantitative) study which aims at discovering subjects' experiences and how subjects make sense of them (phenomenology) or how subjects understand their life (interpretivism).

The observer must apply a range of tools for systematizing and recording the data which are part of his/her experience. This helps to make the process of observation effective and reduce the errors of the observer. Sometimes, the researcher looks for a guideline to direct the observation. In order to deal with such a requirement, he/she prepares a form which lists down the behaviour, situation, and context to be observed and recorded in the course of observation. This is called an observation schedule.

It is important to note down what one observes in the field situation. It could be in the form of a diary, or it may be a daily record of each item, written under appropriate subheadings. Original notes are later re-analysed and placed under the appropriate categories. Field notes aid an orderly and purposeful study and to grasp certain relationship of events.

Box 6.1 Observation Record of Children's Behaviour in Test Situations by Charlotte Buhler & Jean Cross									
Name:	No.	Adequate	Immature	Dependent	Timid	Withdrawn	Aggressive or Hostile	Disorganized	Other
Age: 4 ¼ year	Date:								
School :	Grade:								
<u>Behavior in waiting room.</u>									
1. Quiet absorption in play material	1.								
2. Clinging to mother	2.	✓	✓	✓					
3. Restless running around	3.								
<u>Behavior on way to examination room.</u>									
4. Willingly leaving mother	4.								
5. Takes examiner's hand	5.	✓							
6. Lets mother go after starting to use material in examination room	6.	✓		✓					
7. Unable to stay alone with examiner	7.								
8. Initiates conversation with desire to communicate	8.								
9. Initiates conversation with noisy chattering, not waiting for responses	9.	✓					✓		
10. Answers direct questions only	10.								
11. Remains silent, shy	11.								
12. Remains silent, hostile	12.								Anxious
<u>Behavior during test.</u>									
13. Enters tasks with pleasure and pursues with interest	13.	✓							Short span
14. Accepts tasks as duty	14.								
15. Asks for help and reassurance	15.	✓		✓					
16. Accepts tasks but is easily distracted	16.								
17. Cannot accept tasks for distractibility	17.								
18. Rejects some tasks as too hard "I Can't"	18.								
19. Rejects tasks with hostility	19.								
20. Shares responsibility of putting away, cleaning up	20.	✓							
21. Refuses to help	21.								Over Conforming Anxious

Source: Goode and Hatt (1952), page No. 172

Observation can be effectively supplemented with photographs. Use of photographs as supplement to observation overcomes the possibility of some important aspect escaping the eye of the researcher. Also, photographs are free of bias of the researcher. They bring in detail, authenticity, and impartiality to the research work. According to Pauline V. Young (1966:178), "Whenever possible, photographs should be introduced in a series which might illustrate various aspects of situations...For example, when the photographs of three Molokan generations are compared, we can gain considerable understanding of the degree of cultural change which occurred in the assimilation process of the groups portrayed, judging at least by outward appearances."

We must remember that observation of social phenomena becomes scientific when it is pursued with stated objectives, is systematically organized, employs

certain mechanisms to resist bias and uncontrolled observation. A researcher could observe interactions, processes, or behaviours as they occur; for example, observing student- teacher interaction or he/she could watch the results of interactions, processes, or behaviours; for example, measuring the amount of plate waste left by students in a school cafeteria to determine whether a new food item is liked by them.

Box 6.2 Recording Observations

“The basic document will be, of course, some type of field-experience log. This may take the form of a diary, or it may be a daily record of each item, written under appropriate subheadings. Thus, there might be a set of subheadings dealing with "socialization": crisis situations involving mother and child, scolding by mother, aggression by siblings of various ages, weaning, toilet training, etc. Since social action is swift and the day long, in many cases it will be profitable to keep running notes. These may be scribbles on small cards, key words written in a notebook, or typed notes written at odd moments during the day.... It is particularly important to write out details during the early phases of the field work, for later many of these details will have faded into the expected and taken for granted.

Whether or not the original notes are recorded under subheadings, they must later be reanalyzed and placed under the appropriate categories.... Thus, the presumed purpose may be to study socialization, but the notes may be concentrated upon other "interesting" occurrences, such as conflicts and gossip, extramarital sexual behavior, or food habits. These may be relevant to socialization, but whether or not they have been chosen for that purpose, or for an extraneous purpose, can be seen when the notes are organized. Further, it is then possible to correct the error while field work is in progress.....

The investigator may find it profitable to record both observation and interpretation of the observation.” (Goode and Hatt 1952 124-125)

6.3 TYPES OF OBSERVATION

Observation is of three types. Let us discuss each one of them.

- (i) **Non-Participant Observation:** Here the researcher observes the group without interfering with what is going on. He/she remains an outsider throughout the period of observation. This enables him/her to be objective in recording observations. There is a possibility that those who are being studied know that they are being observed. This is overt observation. The other possibility is that those who are being studied do not know that they are being observed. This is covert observation.

The advantage of covert observation is that people are more likely to behave naturally if they do not know they are being observed. However, covert observation could be ethically improper. This technique, however, offers limited scope to understand the research situations. Insightful socio-cultural accounts of events, actions of the members of the group or community will remain elusive with non-participant method.

- (ii) **Participant Observation:** A researcher is said to be carrying out participant observation when he/she participates in the social and cultural life of a people and understand a social phenomenon or problem. Since doing participant observation means being embedded in the action and context of a social setting, it involves: (i) getting into the location of whatever aspect of the human experience one wishes to study; (ii) building rapport with the participants; and (iii) spending enough time interacting to get the needed data.

Thus participant observation enables the researcher to gain a deep understanding and familiarity with a certain group of individuals, their values, beliefs, and way of life. By living with the group, becoming a part of it, for an extended period of time enables to gain a sufficient range of experiences, conversations, and relatively unstructured interviews for analysis. The researchers are able to formulate first-hand accounts of their lives and gain novel insights.

This research method was pioneered by anthropologists Bronislaw Malinowski and Franz Boas but was adopted as a primary research method by many sociologists later. Today, participant observation, is a primary research method practiced by qualitative sociologists around the world. Participant observation calls for the researcher going to the site where the people he/she wants to study reside. The researcher is required to preferably live with them (usually between 12 and 18 months), speak their language, participate in their day-to-day activities and collect information using the anthropological techniques (e.g., census surveys, questionnaires, interviews). The advantage of participant observation is that the researcher obtains an insight into the social and cultural life of the people from their own perspective. Any question arising in the mind of the researcher pertaining to, for example, why they are performing a particular ritual, is answered by the people there and then.

- (iii) **Quasi-Participant Observation:** It is also important to note that purely non-participant observation is extremely difficult. One cannot understand or comprehend the social life without any participation in it. The term, quasi-participant observation refers to a situation when the researcher participates selectively in the life and culture he/she sets out to study. The observer actively participates in some of the ordinary activities and observes passively from distance in others.

- 1) Point out the characteristics of observation as a technique of data collection.

.....

.....

.....

.....

.....

- 2) In what way do photographs supplement observation to observation?

.....

.....

.....

.....

.....

.....

6.4 STUDIES BASED ON PARTICIPANT OBSERVATION

Participant observation, for many years, has been a hallmark of both anthropological and sociological studies. Here let us refer to two important studies that are based on participant observation.

6.4.1 Argonauts of the Western Pacific

Malinowski conducted field work in the Trobriand Islands, off the eastern coast of New Guinea. His fieldwork centred around the idea that the researcher should fully participate in the culture that is being studied as much as possible, in order to best observe and record the practices that are occurring. He spent 23 months in the Islands in two phases. His famous work is *Argonauts of the Western Pacific* which documents his stay in the Trobriand Islands, observing social relationships, economic matters, as well as the physical and psychological well-being of its people. In his studies of the Trobriand Islands, Malinowski attempted “to grasp the native’s point of view, his relation to life, to realize his vision of his world”.

His ethnography of the Trobriand Islands described the complex institution of the Kula ring (a complex inter-island exchange of arm shell bracelets and necklaces), and became foundational for subsequent theories of reciprocity and exchange. In Melanesia, every year there was a ceremonial exchange of a bracelet and a necklace that involved a ring of many islands. The items were not made of particularly expensive materials nor represented excellent craftsmanship. This led Malinowski to investigate what function this could provide for the islanders. Malinowski's work brought the value of such previously unappreciated cultures to the attention of those in developed societies, showing that there is much to learn about human nature and human society from all people, no matter how "primitive" or strange they might initially appear.

6.4.2 Street Corner Society

The pioneering work of W.H. Whyte on street corner society is based on participant observation of an Italian slum in the suburbs of Boston. It is perhaps the most referred work on participant observation in sociology. During the late 1930s, William F Whyte moved into a low income Italian neighbourhood in Boston. For nearly four years he was a member of the social circle of “corner boys” that he describes in *Street Corner Society*. Whyte revealed his identity to these men and joined in their conversations, bowling and other leisure time activities. His goal was to gain greater insight into the community that these men had established. As Whyte listened to the leader of the group, he “learned the answers to questions I would not even have had the sense to ask if I had been getting my information solely on an interviewing basis”. Whyte’s work was especially valuable, since at the time the academic world had little direct knowledge of the poor, and tended to rely for information on the records of social service agencies, hospitals, and courts (P Adler et al. 1992)

Box 6.3 Participant Observation: Some Difficulties Foote

Compared to the anthropologist who studies a primitive tribe in a remote part of the world, the student of a modern American community faces distinctly different problems. In the first place, he is dealing with a literate people. It is certain that some of these people, and perhaps many of them, will read his research report. If he disguises the name of the district as I have done, many outsiders apparently will not discover where the study was actually located... The people in the district, of course, know it is about them, and even the changed names do not disguise the individuals for them. They remember the researcher and know the people with whom he associated and know enough about the various groups to place the individuals with little chance of error. In such a situation the researcher carries a heavy responsibility. He would like his book to be of some help to the people of the district; at least, he wants to take steps to minimise the chances of it doing any harm, fully recognising the possibility that certain individuals may suffer through the publication. (Foote 2012:342)

Activity 1

Choose an accessible social or cultural venue or event that is not familiar to you. Go to the location and do a purely visual observation—note who is there (demographically); note signs of social differences (deference, use of space, identifying clothing, badges, etc.) and watch how people interact with one another. Note what parts of the scene are easy to understand and which parts are confusing to you as an outside observer. What sorts of additional data—interviews, secondary sources, or additional observation— would you need to conduct to get clarity on these issues?

6.5 MERITS AND DEMERITS OF OBSERVATION TECHNIQUE

Following are some of the merits of observation method:

- i) It is a simple method of data collection.
- ii) Observation is the most direct means of studying a wide variety of phenomena based on actual and first-hand experience.
- iii) It enables the observer to study and record behaviour at the time of its occurrence.
- iv) Data collected under this method is more accurate and reliable, as it is based on the first hand perception of the eyes. Non-participant observation, however, may provide limited insight into the meaning of the social context studied. If this contextual understanding is important, participant observation might be needed.
- v) It enables data collection on a large size of the population at one time.
- vi) Non-participant observation does not depend on people's cooperation with the researcher.
- vii) Participant observation, in which researchers join people and participate in a group's routine activities for the purpose of observing them within that context, lets researchers study a naturally occurring social activity. This fosters an in depth and rich understanding of a phenomenon, situation and/or setting and the behaviour of the participants within that setting.

Observation, as method of study and data collection, however raises many ethical questions. Particularly, participant observation is a much debated method among all social science research methods.

- i) The researcher lives with the people being studied for relatively a longer period of time, participates in their day-to-day activities, observes them from close quarters, gains access to the intimate aspects of social, cultural life of people. At the same time s/he has the task of recording of the observed events, phenomena, activities in a faithful manner. Bernard (1988) suggests that what could clearly be considered as 'gossip' during informal conversation can become data for the researcher. Such kind of issues becomes part of the ethical concerns of the social science research. It is often questioned that whether the researcher informs the people she interacts with for data collection about the usage of such information.
- ii) Further, there are ethical questions regarding the publication of information collected through field notes. Related to this is the preservation of anonymity of participants identified in field notes. It is suggested that with computerization maintenance of anonymity becomes easier than otherwise.

- iii) Another concern raised about the limitation of participant observation is the extent of attachment developed by the researcher with the respondents or informants during field work. During participant observation the researchers often get so involved that they tend to lose objectivity in their observation. Critics go to the extent of questioning how valid participant observation really is. They argue the method lacks objectivity. It can be very difficult for the researcher to avoid subjectivity and forming biased views of the group being studied.
- iv) Also researchers decide what is significant and worth recording and what's not, therefore, it depends on the values of the researcher. In extreme cases, researchers might 'go native', where they become sympathetic with the respondents and omit any negative analysis of their way of life.
- v) Personal factors of the researcher become critical in participant observation. Factors like age, sex, race, ethnic background, presentation of self may influence the process of observation. For example, Dewatt and Dewatt (2002) note that male and female researchers have access to different information, as they have access to different people, settings, and bodies of knowledge.
- vi) Basic understanding of cultural aspects of the group or community studied on the part of the researcher may be essential for a good observation. Researcher proceeds for observation with theoretical framework and pre-fixed hypotheses with an assumption to test them. Sometimes these may become so fixed that the researcher ignores the other equally significant events in the field. It is a widely held notion that social scientists, how much ever they try to restrain from imposing their values, ideas, beliefs in observation, fail to capture social reality objectively.
- vii) Participant observation method tends to be time consuming and expensive. It can take time to gain trust and build rapport, and so for this reason, it may take several days, weeks or even months, before the respondents really start to relax in the presence of the researcher.
- viii) Gaining access in their area/community can also be a problem as many people may not be comfortable in being researched this way, and where covert research is concerned, researchers are limited by their own characteristics. The challenge, however, is gaining access to a setting without disrupting the pattern of others' behaviour. Becoming an inside member of a group, organization, or subculture takes time and effort. Researchers must pretend to be something they are not. The process could involve role playing, making contacts, networking, or applying for a job. A researcher might work as a waitress in a diner, or live as a homeless person for several weeks, or ride along with police officers as they patrol their regular beat. Often, these researchers try to blend in seamlessly with the population they study, and they may not disclose their true identity or

purpose if they feel it would compromise the results of their research. Once inside a group, some researchers spend months or even years pretending to be one of the people they are observing. Information gathered during participant observation is mostly qualitative, rather than quantitative, the end results are often descriptive or interpretive.

- ix) Fieldwork by its very nature involves very long drawn out and intensive research usually by a single scholar working alone. As such, it can only cover a very small part of the world — generally a single village or small community. We can never be sure whether what the anthropologist or sociologist observed during fieldwork is really very common in the larger community (i.e. in other villages, region, or in the country) or whether it is exceptional.
- x) Another important limitation of field work method is that we are never sure whether it is the voice of the researcher we are hearing or that of the people being studied. Of course, the aim is to represent the views of the people being studied, but it is always possible that the researcher — whether consciously or unconsciously — is selecting what will be written down in his/her notes, and how it will be presented to the readers of his/her books or articles. Because there is no other version available to us except that of the researcher, there is always the chance of bias or error.

Check Your Progress 2

- 1) List some ethical concerns that may emerge during participant observation. Use the space below to write your answer.

.....

.....

.....

.....

.....

- 2) In what way do photographs supplement observation to observation?

.....

.....

.....

.....

.....

6.6 LET US SUM UP

Observation has been used in a variety of disciplines as a tool for collecting data about people, processes, and cultures in qualitative research. A vital and useful tool for the understanding of people 's behaviour within the context of their everyday social setting, Observation, particularly participant observation, as a method of social research, enables understanding of many unexplored and

unexplained social phenomena and foundation for theory and hypothesis development.

Observation method of data collection employs other methods like interview and enumeration or sampling during the course of field work. Informal interviewing is a widely used technique by the researcher while interacting with the key informants. This is to suggest that observation method is not always employed in isolation, rather, when employed with other techniques it yields comprehensive results.

The drawbacks, however cannot be undermined such as the probable lack of reliability and representativeness and the problem of not being able to replicate the study. It may be worthwhile to back up with the quantitative information to reach more objective conclusion.

6.7 REFERENCES

- Adler, P. A. and P. Adler. 1987. *Membership Roles in Field Research*. Newbury Park, CA: Sage Publications.
- Bogdewic, S.P. 1999. "Participant Observation" In BF Crabtree and WL Miller (Eds.) *Doing Qualitative Research* (2nd Edition). Thousand Oaks, CA: Sage Publications.
- Goode & Hatt. 1952. *Methods in Social Research*. Mc Graw Hill Book Company.
- Lofland, J. & Lofland, L.H. 1995. *Analyzing Social Settings*. (3rd Edition). NY: Wadsworth Publication Company.
- Whyte, William Foote. 2012. *Street Corner Society: Social Structure of an Italian Slum*. The University of Chicago Press.
- Young, Pauline. 1966. *Scientific Social Surveys and Research*. Englewood Cliffs: Prentice Hall.

6.8 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

- 1) Some of the characteristics of observation are: (i) behaviour is observed in natural surroundings; (ii) It enables understanding significant events affecting social relations of the participants; (iii) It determines reality from the perspective of observed person himself; and (iv) it identifies regularities and recurrences in social life by comparing data in the study with that of other studies.
- 2) Use of photographs as supplement to observation overcomes the possibility of some important aspect escaping the eye of the researcher. Also, photographs are free of bias of the researcher. They bring in detail, authenticity, and impartiality to the research work. They are used to make the observation process as accurate as possible. They help to systematize and record data which are part of his experience. By doing so the process

of observation becomes more effective and errors are considerably reduced.

Observation

Check Your Progress 2

- 1) The ethical concerns which emerge during participant observation can be the –a) the researchers may get so involved that they tend to lose objectivity in their observation b) due to the publication of information collected through field notes the participant's identity is no longer anonymous. c) It is often questioned that whether the researcher informs the people she interacts with for data collection about the usage of such information). Personal factors like age, sex, race, ethnic background of the researcher influence/become critical in participant observation.



UNIT 7 QUESTIONNAIRE, SCHEDULE AND INTERVIEW*

Structure

7.0 Objectives

7.1 Introduction

7.2 Questionnaire

7.3 Schedule

7.4 Comparison between Questionnaire and Schedule

7.5 Interview

7.5.1 Merits and Demerits of Interview Technique

7.6 Let Us Sum Up

7.7 References

7.8 Answers to Check Your Progress

7.0 OBJECTIVES

After going through this unit, you should be able to:

- explain main features of questionnaire, schedule and interview;
- distinguish between questionnaire and schedule; and
- explain the significance of interview guide.

7.1 INTRODUCTION

Basic to any research design is the collection of data, either from primary sources or secondary sources or both primary and secondary sources. Primary data is directly collected through the process of fieldwork, whereas secondary data is obtained from published and unpublished documents. In this unit, we shall discuss the main techniques of primary data collection - questionnaire, interview and schedule. The choice of the research tool to be employed, its specific preparation, construction and use largely depends on the aim, nature, and scope of the study.

In following sections, each of these research techniques will be discussed, along with their merits and demerits.

* Written by Shailey Bhashanjali, North Cap University, Gurugram

7.2 QUESTIONNAIRE

Questionnaire refers to a technique of data collection by using a set of questions which is sent to respondents by hand, by post or email. It could be web-based too. A respondent fills up the answers of the questions in the space provided for the purpose in the questionnaire itself.

Questionnaire is a technique of data collection which is widely used not only by those in academics but also people engaged with marketing, advertising and other fields. It is designed to collect data from large, diverse and widely scattered groups of people. A questionnaire consists of number of questions that are printed in a definite order. The questionnaire is sent to respondents either by hand, post or email. At the beginning of the questionnaire itself, the researcher explains his/her identity and the purpose of the research for which data is being collected through the questionnaire. Often, the researcher adds that the information collected through the questionnaire will be kept confidential, will not be shared with anyone, and will be used only for the purpose for which it is being collected. This provides a sense of confidence among respondents. It is ethically improper to break the confidence of respondents.

Box 7.0 Guidelines for Construction of a Questionnaire

Following are Guidelines for Construction of a Questionnaire:

- (i) A questionnaire should clearly mention the purpose for which it will be used.
- (ii) Clear instructions should be provided to respondents for filling-up the questionnaire.
- (iii) All questions should be in simple language, easy to understand and self-explanatory.
- (iv) All the questions should relate with the objectives of the study. Unnecessary questions should not be included in the questionnaire.
- (v) The number of questions should not be too large. It has been found that respondents get bored or do not have the time to fill-up long questionnaires.
- (vi) Questions should be arranged judiciously. Simple questions should be placed in the beginning and complex ones towards the end.
- (vii) Questions should be logically arranged. In fact, one question should be related with the other.

Questions in a questionnaire can be closed-ended, open-ended or a combination of both. Closed-ended questions are those that for which respondents are required to choose a response from a set of given options. Let us consider an example:

What is your marital status?

- (i) Single (ii) Married (iii) Divorced (iv) Separated (v) Widowed

Main advantage of closed-ended questions is that answers of respondents can be compared easily. Also, they are easy and can be answered quickly. However, they curtail the free-will and creativity of respondents.

Open-ended questions are used to get free and spontaneous expressions from the respondents. These can be used to as an exploratory tool to gain insight to the viewpoints and attitude, describe relationships, and indicate problems, details and events, without restrictions imposed as in the case of closed questions. Let us consider an example:

Please suggest some ways to enhance students' participation in online classes.

.....
.....
.....

Open-ended questions could fetch responses that are completely unanticipated by the researcher. The response could open up new ways of looking an issue and reveal the logic, understanding, and perspective of respondents. The main disadvantage, however, is the possibility of vague responses that do not relate with the question asked. Also, answers of different respondents cannot be compared; space provided for answers may be insufficient as some respondents tend to write long answers which will make the questionnaire bulky.

It is important that a questionnaire is pre-tested before being administered to respondents. For pre-testing, a questionnaire is first administered to a limited number of respondents. The purpose is to make an assessment of each question from the perspective of respondents. In the course of pre-testing, a researcher may find that respondents are not able to answer certain questions because of difficult language, way in which they are framed, or the manner in which they are arranged. The researcher makes necessary changes in the questionnaire based on the response received. It is only after a questionnaire is pre-tested that it is finalized.

The merits of the questionnaire method are:

- It can cover far-flung and inaccessible regions where a researcher may not be able to go.
- It is less expensive to administer. This is because questionnaires are most often mailed or handed over to a large number of respondents simultaneously.
- Data from a large sample can be collected at the same time. Large samples can be made use of and thus the results can be made more dependable and reliable.
- Questionnaire minimizes the bias of the researcher. . It offers greater anonymity. Since the researcher is interested in the answers to the questions, rather than the persons, respondents may feel free to express

their opinion without ambiguity or fear. It also puts less pressure on the respondents for immediate responses.

The main demerits of collecting data through a questionnaire are:

- Can be used only in literate population
- Response rate could be low
- Researcher is not available to explain a question or ask a question based on the response on an open-ended question. There is a possibility of ambiguous replies or omission of replies altogether to certain questions.

Activity1

Construct a questionnaire to be administered to the millennial belonging to higher middle class with an objective to understand their spending behaviour. Itemise the areas of expenditure as being food, entertainment, investment, mobile and electronic items purchase and frame closed and open ended questions to understand the percentage spending on each item and what priorities do they have in their spending behaviour.

7.3 SCHEDULE

Like the questionnaire, schedule is technique of data collection, which contains a list of questions. The difference between a questionnaire and schedule, however, is that while the former is filled by respondents, the latter is filled by the researcher. The researcher goes to the informants with the schedule, and asks them the questions. Researcher plays an important role in the collection of data, through schedules. They explain the aims and objects of the research to the respondents and interpret the questions to them when required. Most common example of data collection through schedule is population census.

The main advantage of schedule is the presence of the researcher. In simple terms, the researcher could explain the question in detail, seek additional information (i.e., information beyond the questions listed in the schedule), obtain clarification on the response, may change the sequence, language and style of questions.

While framing a schedule, the researcher has to take many aspects into consideration. In fact, it is appropriate to identify the aspects on which the schedule needs to be prepared. These aspects are logically arranged and relevant questions are framed. It is likely that more than one question is asked on an aspect with the purpose of obtaining complete information.

7.4 COMPARISON BETWEEN QUESTIONNAIRE AND SCHEDULE

There are certain differences between questionnaire and schedule as listed below.

- A questionnaire is filled by the respondents, while the researcher fills the schedule.

- Questionnaire is more rigid in structure than schedule. Researchers have no control over response rate in case of questionnaires as many people do not respond and/or often return them without answering all the questions. On the contrary, researchers have control over the response rate of schedules since they collect data themselves.
- While questionnaire has a larger reach since it can be distributed to a large number of people at the same time, schedule has a limited reach.
- Identity of respondents is protected when data is collected using questionnaire technique while identity of informants is revealed when data is collected using schedule technique of data collection.
- The success of the questionnaire depends much on the quality of the questionnaire while the research acumen and experience of the researcher determines the success of a schedule.
- The questionnaire can be employed only when the respondents are literate while schedule can be used for data collection from both literate and illiterate informants.

Possibility of obtaining incomplete and imprecise information is relatively more when data is collected through questionnaires than through schedules since the researcher is present in the field situation to verify and corroborate data there-and-then.

Check your progress 1

- 1) What are closed-ended questions?

.....
.....
.....
.....

- 2) What is a schedule?

.....
.....
.....
.....

7.5 INTERVIEW

Interviews are different from questionnaires as they involve close and sometimes prolonged interaction between researcher and informants. The objective of this interaction is collection of data relevant to the study. . There is a difference between “everyday conversation” and “interview”. While conversation is casual and could cover any topic of common interest, interview is based on the theme(s), ideas and issues pertaining to the study. Here, the researcher controls the direction in which the interaction proceeds. Taking interviews is preceded by preparing a list/outline of questions on the topics relating to the research problem.

This list is known as the interview guide. It is used as a ready reference by the interviewer to cover as many issues as possible in a limited period of time.

Box 7.1 Types of Interviews

Interviews may be broadly divided into two types. In the first, the interviewer has a basic idea of the areas about which he would be interviewing the respondent. He carries with him a list of the topics-what may be technically called the 'interview guide'- that would guide him during the interview, but he does not structure specific questions in advance. In such situations, an interview is more or less a 'free-floating' conversation. Often, the interviewer begins with what seems interesting to him or what he thinks would be interesting to the respondent. The interview may cover a number of areas. Called unstructured, it is time consuming, but yields a lot of information apart from the topics in which the interviewer is particularly interested. Therefore, the fieldworker has to sluice away the grit to index what is especially relevant to him.

By comparison, the other kind of interview is structured, in which the investigator prepares a set of questions in advance and is specifically committed to seeking their answer. If the respondent deviates from the topic, the interviewer requests him to return to the point. Such an intervention might annoy the respondent, making him lose interest in the interview, but much depends on how the interviewer handles such situations. Unstructured and structured interviews may also be combined. The interviewer prepares an inventory of the topics of his interest in advance and then starts the interview in a conversational manner. Once rapport has been established, he may gradually funnel down his enquiry, making it structured. The way in which the interview develops depends greatly on the respondent's perception of the interviewer's role.

More elaborately, interviews have been classified into focused, depth, clinical, repeated, and group interviews (Young 1968: 219-22). The focused interview takes place with an individual known to have been involved in a particular situation. In a depth interview, the interviewee is encouraged free expression for an understanding about his/her subjective dimension. Clinical interviews are conducted for gauging the health and psychological status of a person. Social workers, counsellors, and prison workers carry out personal history interviews of the subjects with whom they deal. In a depth interview, the investigator may use projective techniques, such as picture interpretation and sentence and story completion. When the same respondent is interviewed again, it is known as a repeated interview, and when it is an interview with a group, rather than an individual, it is termed a group interview. (Srivastava 2004: 29-30)

An interview guide helps in focussing attention on the salient points in the study; securing comparable data in different interviews; gathering the same range of items essential in the analysis of data or testing the hypothesis formulated. The interview guide serves as an aid or a helpful tool and should not lead to strict adherence to a set of pre-determined questions. It should not hamper the logical flow of the interview and serve as a flexible tool.

7.5.1 Merits and Demerits of Interview Technique

The main merits of the interview technique are as follows:

- Interview technique can be used in both literate and illiterate populations.
- The researcher can make a selection of informants he/she would interview. This selection can be based on researcher's personal choice or sampling procedure.
- Interview enables the researcher to obtain data in detail.
- Interview provides an opportunity to re-frame questions, add and/or skip some questions. The way questions are asked can be adapted to the nature of an informant.
- In the course of an interview the researcher is able to understand the social context of the informant as well as his/her willingness, emotions and personal characteristics that may enhance the quality of data.

However there are certain demerits of this interview method:

- Interviews are difficult to conduct in large heterogeneous population and in inaccessible geographical areas
- There is likelihood that the bias of the researcher would influence the interview.
- Interviews are more-time-consuming, especially when the sample is large and largely depend on the willingness of informants when the researcher goes prepared to conduct the interview. .

It is likely that the presence of the researcher influences the response of the informant. Researcher's personal nature and behaviour could influence the response. Also, the researcher needs to be trained to conduct the interview.

Check Your Progress 2

1) What is 'interview guide'?

.....
.....
.....
.....

2) What is the significance of an interview guide?

.....
.....
.....
.....

7.6 LET US SUM UP

To conduct any sociological research, the choice of a particular instrument of data collection will depend on the objective, nature and scope of the study. The characteristics of the respondents as well as the field situation, along with the constraints of resources-trained personnel, funds and time are other important considerations in the selection of the technique of data collection.

As we earlier understood, questionnaire is likely to be a less expensive procedure than the interview as it is simply mailed to the respondents with a minimum of explanation. Further, the questionnaire can often be sent to a large number of individuals simultaneously; an interview, on the other hand, usually calls for questioning each individual separately. With a given amount of funds, it is usually possible to cover a wider area and obtain information from more people by means of questionnaires than by personally interviewing each respondent. Another advantage, as discussed, is that of a greater confidence that respondents may have because of their anonymity, and thus feel freer to express view which they fear might be disapproved of or might land them into trouble. Interview is a more appropriate technique for revealing information about complex, emotionally laden subjects or for probing the sentiments that may underlie an expressed opinion.

Learning how to design and use interviews, questionnaires and schedule is an important skill for researchers. Such survey instruments can be used in many types of research, from case study, to cross-sectional survey, to experiment. The decision to use one method of collecting data over another method is a complex one and must take many factors into consideration.

7.7 REFERENCES

- Fred N. Kerlinger. 1973 *Foundations of Behavioural Research*. Delhi: Surjeet Publication. Delhi, 1973.
- Goode, William J. and Paul K. Hatt. 1952. *Methods in Social Research*. New Delhi: McGraw-Hill Book Company
- Pauline V. Young. 1988 *Scientific social surveys and research*. New Delhi: Prentice Hall of India Pvt Ltd.
- Srivastava, V. K. 2004. "Introduction" in *Methodology and Fieldwork*. Edited by V K Srivastava, Delhi: Oxford University Press.

7.8 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

- 1) Closed-ended questions are those that for which respondents are required to choose a response from a set of given options.

- 2) Schedule is technique of data collection, which contains a list of questions. The researcher fills answers of these questions on behalf of respondents.

**Questionnaire, Schedule
and Interview**

Check Your Progress 2

- 1) Taking interviews is preceded by preparing a list/outline of questions on the topics relating to the research problem. This list is known as the interview guide.
- 2) An interview guide helps in focussing attention on the salient points in the study; securing comparable data in different interviews; gathering the same range of items essential in the analysis of data or testing the hypothesis formulated. The interview guide serves as an aid or a helpful tool and should not lead to strict adherence to a set of pre-determined questions.



UNIT 8 CASE STUDY AND LIFE HISTORY*

Structure

8.0 Objectives

8.1 Introduction

8.2 Extensive Research vs. Intensive Research

8.3 Case Study Method

8.3.1 What is a Case Study?

8.3.2 Significance of Case Study Method in Social Science Research

8.4 Life History Method

8.4.1 What is Life History?

8.4.2 Significance of Life History Method in Social Science Research

8.5 Let Us Sum Up

8.6 References

8.7 Answers to Check Your Progress

8.0 OBJECTIVES

After going through this Unit, you should be able to:

- distinguish between extensive research and intensive research;
- explain the nature and scope of case study method and life history;
- discuss the significance of case study method and life history in social science research; and
- differentiate between case study method and life history

8.1 INTRODUCTION

We have heard the words, 'case study' and 'life history' on several occasions. Now-a-days people use the two terms in different senses and refer to detailed study about something. People often find it difficult to distinguish between case study and life history. In fact, most of the time people maintain that life history has something to do with the discipline of history rather than with sociology.

In this Unit, let us learn about case study method and life history closely. Let us find out in what way case study method and life history are used by sociologists to collect data. We begin by developing an understanding of what case study is and its significance in social science research. We will answer the question: what

*Written by Nita Mathur, School of Social Sciences, IGNOU, New Delhi

is a case? How many cases should be included in a study? Following this, we will discuss the technique of life history in terms of its significance. Towards the end of the Unit, we will differentiate between case study and life history.

8.2 EXTENSIVE RESEARCH VS. INTENSIVE RESEARCH

As you are aware, social research is based on study of individuals, society, culture. In a broad sense, this entails study of people, groups, organizations, institutions, villages, towns, and cities. Two options are available to researchers. First option is to carry out extensive research. Here, data is collected from a large number of informants. This could be through census or survey methods about which we learnt in Unit 5. When data is collected from a large number of informants, it is possible to draw comparison between units of observation. Often, extensive research uses quantitative techniques of data analysis. We will learn quantitative methods of data analysis in Units 10, 11, and 12. Extensive research brings large data that can be subjected to advanced mathematical models. Second option before researchers is to carry out intensive research. Here, detailed and in-depth data is collected from a small number of informants. Analysis of such data is carried out using qualitative methods. It is difficult to draw comparisons and arrive at generalizations.

Let us consider the example of a study of women's participation in elections. If a researcher seeks to carry out extensive research, she/he would obtain data on extent of women's participation using survey method in all the polling stations in a district. Of course, the questions in the survey would depend on the objectives and variables of the study. An intensive study on the same topic would focus on smaller sample size. Data collection would be limited to select polling stations and select informants. The researcher would use the techniques of, among others, participant observation, interview, and case study method. The choice between extensive and intensive research depends on the scope of the study. It is, however, possible to combine both extensive and intensive research in one study. Such a study would be voluminous but rich in content and reliable. Let us see how this is possible in the case of the example of participation of women in elections. The researcher could carry out extensive research covering all the polling stations in a district. Data from different polling stations could be combined leading to framing of generalizations. Further, she/he could use some of the generalizations to carry out intensive research. For example, the generalization in this case could be that participation of literate women in elections is higher than that of illiterate women. The researcher could carry out intensive research based on this generalization. She/he could bring in variables such as age, caste, religion etc and examine this aspect closely by way of intensive research. The sample size could be small but in-depth data could be collected on limited number of variables using various techniques about which we have discussed in earlier Units.

In this Unit, we will learn about two important methods of data collection used in intensive research. These are: case study and life history.

8.3 CASE STUDY METHOD

There are two aspects that social scientists can study. The first aspect comprises norms, beliefs, values and customs. The second aspect is the way in which these norms, beliefs, values and customs are negotiated or dealt with by people in day-to-day life. Consider the example of a fight on the road. Personal experiences, motivations and other aspects of people's lives are of tremendous importance to social scientists.

8.3.1 What is Case Study?

A case study is a holistic study of a social phenomena carried out in its natural context i.e., in the situation or framework within which it exists. It is carried out over a period of time/during a certain period of time. Yin (2003:4) explains that in a case study the 'case', i.e., phenomenon being studied, is not separated or kept apart from its context. In other words, a case is studied in totality and in its own or naturally occurring context rather than in isolation. Often students use the term 'case' erroneously. Let us go back to the earlier example of study on participation of women in elections. A student who interviews five women and reports them as case studies could be making a mistake. The student can be said to have collected case studies if she/he has carried out detailed study on each woman incorporating the socio-cultural context in which each woman lives. This would involve not one but a few in-depth interviews along with other techniques of data collection.

Box 8.1 Overall Understanding of Case Study Method

Social scientists study many culture groups, small social groups (a family, political party, a gang, a leadership group...They also study large groups as social units (sects, national and racial groups). Social institutions (courts, hospitals, churches, industrial organizations, governmental divisions have also been studied comprehensively as social units. Studies of neighbourhoods, large and small communities (Russian-Town, China-town, the Black Belt, the ghetto, the Gold Coast) have been frequent topics of studies as social units. A comprehensive study of a social unit, or a community is called a case study....

Case data may be gathered exhaustively of an entire life cycle of a social unit or a definite section of it. Whether a section or the whole life is studied, the aim is to ascertain the *natural history*, that is, an account of the generic development of a person or group, or whatever constitutes the social unit in a particular study, revealing the factors that modeled the life of the unit within its cultural setting. Because of its aid in studying behavior in specific, precise detail, Burgess termed the case study method, 'the social microscope' (Young 1996: 246-247).

You may be wondering, what a 'case' is? What can be considered as a 'case' depends on the objectives of research, the way research questions are framed, and the scope of the research. A case could be an individual, some event, an entity, an organization, or an institution. The next question which comes to mind is how

should cases be selected? It is appropriate to choose case(s) that will most likely address the research question(s). Case could be one or more in number depending on the way the research is planned. Only one actor may constitute the totality of case study or many actors could constitute the totality of case study. Selection of case(s) is guided by theoretical propositions—they suggest what needs to be focused. Researchers should choose case(s) that are most likely to shed light on research questions.

Research studies could be based on a single case or multiple cases. Each of the case in a study based on multiple case studies would be of the same type so that they replicate each other. For example if I want to do a case study of what happens when washing machine is introduced for the first time in a small town. I may take one single household using the washing machine and do an intensive case study. The choice of the household would be such that it represents all other households. Such a case can be referred to as a typical case. Alternatively, I could choose to study more cases of the same type and in same situation to make sure that nothing is missing. Two options are available. The first option is to take more cases of the same type. The expectation is that the same kind of findings will emerge. The second option is to take some case studies of households using a washing machine and some case studies of households that do not use the washing machine. Here we can expect contrasting results. Households can be opposed in terms of composition or they may be similar, but one has a washing machine and the other does not.

8.3.2 Significance of Case Study Method in Social Science Research

One of the distinguishing features of case study method is that it incorporates all aspects of a case. Consider the example of construction of road passing through a village in a hilly terrain. Now there are different lobbies of stakeholders emphasizing or opposing the construction. One lobby takes the development view, another takes the environmental stand and a third takes a humanitarian stand supporting the villagers' right to their way of life. I can undertake a case study of the village incorporating environmental and development variables, all at the same time. The result of my study need not be positive, but it should let me see what exactly would happen considering all dimensions: environmental, developmental, and humanitarian if road construction did take place. Depending upon the results, such a study can influence further policy making and also initiate debates and raise issues.

What a case study can demonstrate is the relative value of the variables and the way in which each stakeholder can be situated in relation to the others. For example, if we look at a project from one point of view only we are likely to prioritize that particular point of view, it could be only development or only environment. But a study that covers all the aspects will let us see what happens to environment if development is pushed to a certain level and how we can hold development at a certain level so that the society and culture undergo least trauma and also how to prioritize one variable over the other.

Case study method enables researchers to overcome the inability of individuals to explain norms and actual outcomes of those norms. Adamson Hoebel and Karl Llewellyn carried out a collaborative study on how disputes are resolved among the Cheyenne community. They made use of the case study method in the course of the study. They are credited with introducing the case study method in anthropological studies. They examined the processes by way of which dispute cases were settled. This included examining the situation, what each participant did in relation to the dispute, what steps were undertaken by other people who had something to do with the dispute, who intervened, what was the reasoning pattern, final outcome of attempts to resolve the dispute, effect of the dispute on the parties involved in the dispute and on the future of similar disputes, and on general life of the community. They found that people did explain the norms and rules to deal with disputes. It was possible to identify ideal patterns against which researchers could assess actual behaviour in real life situations. This was, however, not enough to arrive at an understanding of dispute resolution in practical situations. There was a divergence between what was expected and what actually happened when disputes arose. Use of the case study method resolved these issues. By using the case study method researchers were able to overcome the inability of individuals to explain how norms and rules are interpreted and made use of in dispute situations.

Check Your Progress 1

- 1) Define case study.

.....
.....
.....
.....
.....

- 2) How should cases be selected?

.....
.....
.....
.....
.....

8.4 LIFE HISTORY METHOD

Case Study and life history are closely related in the sense that both examine the subject under study in detail and holistically. When a case study is focused on the life of an individual person, it takes the form of what we refer to as life history. Let us understand what life history is.

8.4.1 What is Life History?

One of the main objectives of social scientists is to obtain the real and personalized account of individual's experiences, expectations, aspirations,

tensions, motivations and all that which influences her/his action, behavior and philosophy of life. As Young (1996:246) mentions, 'To understand his behavior fully and intimately, he must supply a detailed and penetrating account of what he does and has done, what he thinks he does and has does, what he expects to do, and says he ought to do. A fairly exhaustive study of a person or group is called a life or case history.'

For long anthropologists and sociologists have used the life history method to supplement their ethnographies and other accounts. One of the earliest studies in which life history method was used extensively was Thomas and Znaniecki's *The Polish Peasant in Europe and America*. This is a voluminous work. The book runs into five volumes. It is based on life history data which supplements their findings and generalizations. Other sources of data are: newspaper articles, letters, records from courts and welfare agencies. Life history method has been used, among others, in studies on race, migration, delinquency. Additionally, the method has been used in dalit studies. Here, life histories of dalits shed light on their social and economic conditions, struggles and other aspects of their life. These life histories are well supplemented with their written accounts, narratives, stories, poems, etc.

One of the most frequently referred to studies using life history method is that on the people of Alor. In this study, eight life histories of Alorese culture and personality (1944) are presented and analyzed by Cora Du Bois. She used life history method to construct typical personality of the Alorese. Cora Du Bois was, however, criticized for selecting atypical individuals i.e., those individuals who did not represent Alorese personality adequately. The main criticism was that the very basis on which Cora Du Bois developed typical personality of Alorese was ill-founded. It was said that a typical Alorese men and women remained too busy to spend time with Cora Du Bois. The men and women she selected for life history were the ones who were free and had nothing better to do than talk to Cora Du Bois. Such people were an exception in society and should not have been used for projecting typical personality of Alorese.

As social scientists we have to decide which part of the life history we are interested in. Our selection should be such that the individual we select represents others of her/his kind. The choice of whose life history we are taking involves some kind of a logic like age, gender etc. We would not like the life history of a too young person, or someone who is too typical and does not represent the culture is also not desirable. We do not want delinquents, drop outs etc, for they represent the anomalous unless of course our goals of research demand it. If we are doing work on drug addicts or criminals, we will need life histories of such people.

The life history is a tool of research which places individual at the centre of research and provides much depth to analysis. It appears in the narrative form that personalizes the data and makes it more human and subjective.

8.4.2 Significance of Life History Method in Social Science Research

The usefulness of a life history is that it provides an insight of a culture through intensive and holistic understanding of life histories of individuals selected for the reason that they represent the majority of the people. Importantly life histories incorporate multiple dimensions of an individual's life. It contextualizes the individual in the historical and cultural transformations of society. Since it is in the form of a narrative, it brings together the personal experiences and the social context of an individual's life. The theoretical basis of the life history method is to be sought in the link between culture and personality. This is based on the understanding that to some extent, people are shaped by culture and also that individual life experiences reflect to a large extent the collective experience.

Life history as a tool also depends a lot on memory. It could be argued that people do not remember everything and that only those aspects that influence her/him the most are usually retained in the memory. Well, while this is true, the life history method continues to remain valid and important for the reason that people usually remember what is important to them and since people are embedded in their society and culture, what they remember would be socially important.

As a tool of research, life history is an interaction and it brings together two people (the researcher and the informant) in close interaction. Tonkin (1922) has pointed out, that both, the researcher and informant are able to understand each other better since both get an opportunity to discuss their own life histories. This enriches the data further.

Check Your Progress 2

- 1) What is the main criticism of Cora Du Bois's case studies of the Alorese?

.....
.....
.....
.....

- 2) What is life history method in social science research?

.....
.....
.....
.....

8.5 LET US SUM UP

In this Unit, we learnt two important methods of data collection: case study method and life history method. We discussed what a case is, how many cases should be included in a study and the importance of case study method. We also learnt the difference between case study and life history. We discussed the importance of life history method in social science research. We found that in collecting case studies and life histories the techniques of interview and

observation are most commonly used. It is also important to remember that as we deal with personal accounts of people, it is important to remain sensitive to their concerns. Also, we should not be judgmental in data collection and analysis. Social and cultural context should provide the backdrop against which data is interpreted.

8.6 REFERENCES

Du Bois Cora. 1994. *The People of Alor*. Minneapolis: The University of Minnesota Press.

Thomas, William and Florian Znaniecki. 1958 (originally published in 1920). *The Polish Peasant in Europe and America*. Dover: New York.

Tonkin, Elizabeth. 1992. *Narrating our Pasts : The Social Construction of Oral History*. Cambridge: Cambridge University Press.

Yin, Robert K. 2003. *Applications of Case Study Research*. New Delhi: Sage.

Young, Pauline. 1996. *Scientific Social Surveys and Research*. Englewood Cliffs: Prentice hall International, Inc.

8.7 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

- 1) A case study is a holistic study of a social phenomena carried out in its natural context i.e., in the situation or framework within which it exists.
- 2) It is appropriate to choose case(s) that will most likely address the research question(s). Selection of case(s) is guided by theoretical propositions—they suggest what needs to be focused. Researchers should choose case(s) that are most likely to shed light on research questions.

Check Your Progress 2

- 1) Cora Du Bois was criticized for selecting atypical individuals i.e., those individuals who did not represent Alor personality adequately. The main criticism was that the very basis on which Cora Du Bois developed typical personality of Alor was ill-founded. It was said that a typical Alor men and women remained too busy to spend time with Cora Du Bois. The men and women she selected for life history were the ones who were free and had nothing better to do than talk to Cora Du Bois. Such people were an exception in society and should not have been used for projecting typical personality of Alor.
- 2) Life history is the real and personalized account of individual's experiences, expectations, aspirations, tensions, motivations and all that which influences her/his action, behavior and philosophy of life.

UNIT 9 ETHNOGRAPHIC METHOD*

Structure

9.0 Objectives

9.1 Introduction

9.2 What is Ethnography?

9.3 Use of Ethnographic Method in Fieldwork

9.4 Emerging Trends in Use of Ethnographic Method

9.5 Merits and Demerits of Ethnographic Method

9.6 Let us Sum Up

9.7 References

9.8 Answers to Check Your Progress

9.0 OBJECTIVES

After going through this Unit, you should be able to:

- explain what ethnography is,
- discuss how ethnography developed as a method of social research,
- explain Malinowski's use of ethnographic method, and
- discuss emerging trends in use of ethnographic method.

9.1 INTRODUCTION

In this Unit, we are concerned with ethnography. Interestingly, ethnography is both a method of research and product of research. Let us understand this aspect of ethnography more closely. Ethnography as a method of research or let us say the ethnographic method of research, refers to a composite range of techniques of data connection that are used in combinations to develop a holistic understanding of the way of life of a group. Ethnography as a product of research refers to the descriptive account presented by the researcher as an outcome of research.

We begin this Unit with developing an understanding of what ethnography is. Thereafter, we will outline how ethnography developed as a method over a period of time, how ethnographic method is employed in research. Finally, we will discuss the new trends in use of ethnographic method.

* Written by Nita Mathur, School of Social Sciences, IGNOU, New Delhi

9.2 WHAT IS ETHNOGRAPHY?

There is no standard or simple definition of ethnography which is accepted by all social science researchers. Broadly, however, ethnography is understood as an in-depth, detailed account of a social or cultural group. Typically, ethnography ‘involves the ethnographer participating in people’s daily lives for an extended period of time, watching what happens, listening to what is said, asking questions- in fact, collecting whatever data are available to throw light on the issues that are the focus of research’ (Hammersley and Atkinson 1995:1). In a conventional sense, ethnography comprises holistic description of a group or community.

Box 9.1 Ethnography as Method and Product (Berreman 2004)

‘An ethnography’ is a written report summarizing the behaviours, and the beliefs, understandings, attitudes, and values they imply, of a group of interacting people. Thus, an ethnography is a description of the way of life, or culture, of a society. Everyone who does empirical research in cultural or social anthropology engages in ethnography as a process, though by no means does everyone report his research in the form of an ethnography. An ethnographic study is generally expected to give an overall view of the culture of the people about whom it is written, within the limits inherent in ethnographic research in prose exposition. As such, it attempts, to cover all aspects of the culture of a given society. In practice, it covers those aspects which the ethnographer considers relevant to an understanding of the main features of the culture he studied (p.157).....

Whatever the method of presentation and analysis, the aim of ethnography is to report the culture studied in sufficient depth and breadth to enable one who has not experienced it to understand it. Understanding may be assumed to have been achieved if the reader learns how participants in the culture see themselves, others, and their environment, and how they deal with each. Ideally, this requires that the ethnographer convey information sufficient in quantity, variety, and quality, and so organized and analysed that it would enable the reader to understand events of the culture if he were to experience them, to anticipate the reactions of members of the society to the events they experience, and, ultimately, to behave appropriately in that society in a manner similar to that of a person who has lived the culture. In this respect, the ethnography is analogous to a linguist’s description of a language. It may not enable the reader to speak the language easily, but it enables him to understand how it works. An ethnographic study comprises, essentially, a statement of a set of rules which describe how people act in their culture. Given time, the reader of an adequate ethnographic (or linguistic) description can understand what is done (or said), can act (or speak) correctly and meaningfully in the culture (or language), can anticipate the responses he will evoke in others, and can respond to them in ways they deem appropriate (p. 159).

During the colonial period, description of natives was often used as a tool of the colonial agenda of administration. Many of the earliest ethnographies such as *The Angami Nagas*, *The Sema Nagas*; have been written by administrators like J.H.Hutton. They were impersonal, detailed accounts of people, prepared largely for benefit of administrators. Such records providing exhaustive information including minute details about natives were used by administrators to govern them. Such descriptions presented the people being studied as 'interesting specimens'. Many anthropologists such as Meyer Fortes and E.E. Evans-Pritchard were sent to the field as employees of the state to collect data for the benefit of the administrators. They, however, were compassionate to the people and no longer felt that they had to support their respective governments. In fact in the post colonial period, many anthropologists have given up the responsibility of providing information to the state which could be used to administer the people they study. Interestingly, rather than fostering the ambition of the state, many of them have emerged as spokespersons for the people they study and take up their case before the state.

It is important to note that in the post-colonial era, natives are no longer treated impersonally as 'interesting specimen'. This has brought about a change in the way ethnography is understood. While earlier ethnography was treated as a product in the form of a physical document containing detailed description, it is now treated as a method. From external and uninvolved 'scientific' description, ethnography is a more engaging endeavour in which the researcher gets involved in the life and activities of the people. In the course of close engagement with the life of people, she/he comes to appreciate and develop an understanding of the other's point of view. The process has enabled a more humanizing approach to the understanding of other cultures.

Some of the distinguishing features of studies employing the ethnographic method are the following. First, they are based on data collected by researcher by spending as much time as possible with the people. Ideally, this duration should be at least one year. Researcher uses participant observation technique to collect data. She/he participates in the daily life of people (establishes continuous interaction with them and participant observation and interview. These are combined with other techniques of data collection such as survey. The ethnographic account usually provides a rich description of how fieldwork was carried out and how data was collected. Second, ethnographic accounts are based on people's point of view. The interpretation of data is situated in the local context of the people under study. The researcher refrains from making judgments.

The credit of fieldwork, more specifically use of ethnographic method, is ascribed to Bronislaw Malinowski. Let us find out how Malinowski used the ethnographic method in research. Before we do that, however, let us complete check your progress exercise below.

- 1) What do you understand by ‘ethnography as method’ and ‘ethnography as product’?

.....

.....

.....

.....

.....

- 2) Outline three distinguishing features of ethnographic research.

.....

.....

.....

.....

.....

9.3 USE OF ETHNOGRAPHIC METHOD IN FIELDWORK

As we have learnt earlier, prior to the early part of 1900s, information about different cultures was collected by missionaries, travellers and others who were not trained to carry out fieldwork systematically. The purpose for which information was collected was mostly for facilitation of governance by administrators rather than understanding culture and society in totality. It was Bronislaw Malinowski who is credited with carrying out fieldwork using ethnographic method for the first time. Let us discuss the social and academic background of Malinowski from Box 9.2.

Box 9.2 Social and Academic Background of Bronislaw Malinowski

A Polish man, born in 1884 if aristocratic parents, who studied maths, physics and philosophy at the Jagiellonian University, in Cracow, he was inspired to take up anthropology after reading Fraser’s at the London School of Economics and Political Science. He gained the financial support to undertake field research in New Guinea, but war broke out while he was in Australia. However, though legally an ‘enemy’ in Australia, he was able to move freely about the Pacific islands for the duration of the war as long as he reported his movements to the Australian government. His most famous research was carried out in the Trobriand Islands in Melanesia, off the north-east coast of Australia. (O’ Reilly 2005:7).

When Malinowski was stranded in Trobriand Islands with no connection with the outside world, the only option before him was to learn the language and interact closely with the natives. He developed an understanding of their behavior, beliefs, values, customs and everyday life. Here, certain important points emerge in the use of ethnographic method. The first is learning the language. In order

that the researcher interacts with the people under studies, she/he should acquire at least working knowledge of the language. This facilitates communication and equips the researcher to make sense of documents (such as letters, diaries, personal records) written in people's own language.

Second, living with people for a prolonged period of time is quite difficult. Traditionally, researchers were advised to be in the field for at least one period in the course of which they would live with the people and participate in their activities. While participating intensively in the activities of the people, the researcher has to remember to maintain a mental distance, i.e., not get so involved that become oblivious to their research work. They are required to adopt the 'marginal native' status in fieldwork. One has to learn the language, live with the people, participate in their activities, and yet maintain a mental distance. The mind of the researcher is aware of the affairs of the people and the same time not totally involved so as to lose sight of his/her own identity.

Third, being non-judgmental and holding respect for the culture of people under study. Researchers are expected to shed ethnocentrism i.e. their own biases and sense of superiority of themselves and their own culture over people they study. Researcher, should control their surprise, sometimes even shock on seeing certain behavior and practices of the people that contradict those of her/his own. Explicit exhibition of disgust, anger, shock have to be shielded from the people. At any cost, one has to adopt an attitude of cultural reactivism i.e., acceptance of and respect for local culture and way of life. In fact, such practices enable a researcher to establish rapport with the people more easily and quickly. On many occasion, it has been found that the researcher is treated by the people under study as a part of a particular family. Some of them have also adopted clean names acquired local skills (e.g. that of making baskets.)

Fourth, collecting data in context. Malinowski insists that data should be collected from people in their natural setting. It is only by being in context that a researcher can gain an insight into how the people deal with and experience things in their everyday life. Also, what people say they do or the norms they adhere to are not always followed in the same way. It is only when a researcher collects data in the natural setting of the people is it possible to see if there is any discrepancy in the way people say they behave and the way they actually do.

Fifth, the context and the way in which ethnographic method is used should be described clearly and in detail. Malinowski states that the researcher should describe his/her method in the same way as a scientist would explain the conditions of an experiment. Description of how the ethnographic method is used in a particular research enables subsequent researchers to understand what were the difficulties and challenges that were encountered and how these were overcome, reasons for collecting certain data and how it was done.

Malinowski devoted himself to studying the "imponderabilia of everyday life" so well explained in his book, *Argonauts of the Western Pacific*. This book is now a classic of anthropological literature. Many anthropologists have adopted the

ethnographic method of Malinowski and presented their ethnographies on the lines of *Argonauts of the Western Pacific*. In essence the fieldwork tradition established by Malinowski and which is essential to produce the rounded of description of a culture that we call an ethnography, requires a long term immersion of the researcher in the local culture. It is different from mere observation, for a casual observer can be mentally isolated from the situation under observation.

9.4 EMERGING TRENDS IN ETHNOGRAPHIC METHOD

The beginnings of ethnographic method are situated in functionalism. The basis of functionalism was the organic analogy which presented a synchronic picture of society. The focus was on what as happening in the present rather than what happened in the past. The other point of focus was on harmony as a normal state of being. Each culture was treated as unique and functional in its own way and amenable to study in its entirety. It is worthwhile to note that functionalism is a theoretical and methodological perspective which views society as a system of inter-related parts. These inter-related parts contribute to the stability of society. Functionalism seeks to understand and explain a custom or a culture feature in learns of its functions or the role it performs in society. In the book, *Argonauts of the Western Pacific*, Malinowski explained the role of magical and religious rites in stability and maintenance of social structure. Through use of ethnographic method, he found that magic arises and function in situations of uncertainty and emotional anxiety. It performs the function of relieving anxiety in situations of crisis and instils hope and faith among the people. Malinowski was inspired to analyze the ethnographic data in the perspective of functionalism.

In more recent times with wider philosophical transformations regarding the reality of the world, the ethnographic method too has undergone change from being an objective method to far more subjective method is which the researcher interprets data in light of her/his own subjectivity. It is however important for researchers to reveal their subjective position and also to concentrate more on ‘narratives’ (i.e., personal account of people) to allow the voices of the informants to come through.

Thus one does not presume any more that people within a culture would have the same experiences and similar life. In contemporary times ethnographic method incorporates the view that different people have their own version of reality. Each person’s account is treated as authentic. Care is taken to make the data representative of all segments of society and not of a few privileged, ‘key informants’. The choice of key informants should in itself take care of different segments of the social reality that is being observed.

Earlier, community was loosely referred to as a group of people living in a common territory and practices the same culture. In contemporary time, however, people of a community migrate to different parts of a country and even different

parts of the world. Hence, when a researcher sets out to make an ethnographic study of, for example, Santhals. The researcher might have to visit different parts of India and even abroad to present an ethnography. This enterprise of carrying out ethnographic study at many locales is referred as 'multi-sited ethnography'. Now, sometimes, it may not be possible to visit all the places where Santhals have migrated. In such a situation the researcher might have to use digital means (e.g., computer, smart phones etc.) to collect data. The use of digital means in ethnographic method results in what is commonly understood is 'digital ethnography'.

Data for digital ethnography is chiefly collected through online questionnaires, digital videos, social networking websites, and blogs. Let us examine each one of these in detail.

(i) Online Questionnaire: For long, researchers have used the technique of questionnaire for data collection. Questionnaires were often printed and sent to respondents through post. With the coming of computers and acquired efficiency of communication through email, questionnaires came to be sent by email. In recent times, however, questionnaire hosting services (such as survey monkey) are readily available. They make it easy to design, deploy, and analyze web-based questionnaires. Online questionnaires can be easily stored and retrieved quickly (even when the number is very large). Further, they can be conveniently subjected to quantitative analysis as it is possible to export data to SPSS or some other package for analysis. Additionally, online questionnaires have a global reach. Consider the case of an online questionnaire on the use of mobile phone by youth. Now, depending on the scope of the study, the online questionnaire could be sent to youth in different parts of the country as also different parts of the world. It has been found that respondents frequently mail the questionnaire to their friends and acquaintances with a personal request to respond. This widens the data-base of the study. Of course, it is possible to apply filters and make use of restricted data from the sample. Let us go back to the example of study on use of mobile phone by the youth. If a researcher wants to restrict analysis to women respondents from south Asia, she/he could apply filters and, from the large data set, selects data pertaining to women in south Asia.

(ii) Digital Videos: Many times, people upload their videos on different sites. These videos could present a social situation or an event as it occurs in a natural setting or it could be performed for the camera. In both cases, digital videos provide a rich source of data for researchers. If they depict a social situation in natural setting, researcher develops an understanding of how it happens; if the situation is performed for the camera, researcher develops an understanding of what is important for the people, what is it that they want to show. The researcher has to, however, be watchful and place what is shown in the digital video in its proper context or else, the data would be misleading.

(iii) Social Networking Websites: Research can be enriched with use of social networking websites. The sites generate a large data-base of people who closely or remotely share common interests. Social networking websites such as Facebook show how people put up their personal details that are available to a large number of people to discuss, comment, ridicule or even troll. These details are used by researchers. Let us find out how social networking websites can be useful to ethnographers and what difficulties they hold in Box 9.3.

Box 9.3 Advantages and Disadvantages of Social Networking Websites for Researchers

Specifically, social networking sites can be useful to ethnographers in the following ways:

- 1) they are virtual 'gatekeepers' with chains of 'friends' who are potential research respondents;
- 2) they contain vast stores of multimedia material regarding even the most marginal social movements of groups;
- 3) ethnographers can invisibly observe the social interactions of page members, gleaning a previously unavailable type of ethnographic data;
- 4) pages can be created by social researchers with the explicit purpose of conducting research online (e.g. focus groups watch an embedded video and comment on it);
- 5) the structure of relationships on the sites is a useful research method itself with, as Garton et al. (1999:78) argue, the content, direction, and strength of the relationship 'strands' fruitful approach;
- 6) pages can be created by social researchers to disseminate useful information to the public, an approach taken by the creators of the 'Cure Diabetes' MySpace page (Barsky and Purdon, 2006).

Though very alluring, the drawback of these research options is that membership of these communities is inherently restricted to the digital 'have nots', and ethnic/gender digital divides strongly persist, an issue I discuss here. While allowing for these gaps or omissions, research done through social networking sites is suited to projects whose respondents are either familiar with or can be trained on the requisite technologies. That being said, the use of social networking sites for focus groups, for example, can result in increased inclusion for those with disabilities (mobility and otherwise) as well as groups that are vulnerable or otherwise difficult to access (Murthy 2008:844-845).

(iv) Blogs: Many people write blogs on a regular basis. These are read and commented upon by others. Blogs present a person's perspective and other people's response to it. Use of blogs in research could serve as one of the means of data collection and interpretation. It would be worthwhile, to use it along with other sources rather than singularly.

Check Your Progress 2

- 1) What is multi-sited ethnography?

.....
.....
.....
.....

- 2) List the main sources of data collection used in for digital ethnography.

.....
.....
.....
.....

9.5 MERITS AND DEMERITS OF ETHNOGRAPHIC METHOD

As we have understood, the ethnographic method involves close interaction with people under study over a prolonged period of time. This has both advantages and disadvantages. Let us examine the merits of ethnographic method first followed by its demerits.

Merits of Ethnographic Method: The value of ethnographic method has gained rather than lost its importance over a period of time. The depth and authenticity of information that can be gained by being with the people and participating in their daily life activities as much as possible cannot be replaced by any other method.

Also, use of ethnographic method enables the researcher to understand the culture and behaviour of a people in its context. The researcher develops empathy for comes to respect other cultures. This empathy leads not only to precise understanding and appreciation of the people and culture under study but also to a more humane way of dealing with other issues in life.

Further, as pointed out by the famous anthropologist, Lev'i-Strauss use of ethnographic method enables us to look below the surface, at deeper levels of interconnections of social phenomenon. Let us consider the example of a group dance by people under study to attract rain. Now, for anyone who watches it, the main issue will be to find out whether the dance led to rainfall or not.

For the researcher using the ethnographic method, this issue is not of much concern. More important than finding out whether rain was caused by the dance or not, is what function the group dance performed and how the people related with it. The group dance could have fostered solidarity and cohesiveness in society, it could have brought enemies together as performers or audience; it could reflect the social organization of the people. The more time one spends with a people, the more insights one gains on their way of doing things and how they relate to the world. It is by this process that ethnographers have gained insights into different cosmologies, and philosophies. They have discovered that

gender can be constructed in many different ways, that man and woman can be constructed differently in different cultures, that death can be understood in a variety of forms.

Demerits of Ethnographic Method: At one time it was believed that societies were like as organisms constituted of different parts and that the main objective of an ethnographer was to show the interconnections between different aspects of society. If this was true, then any one ethnographer's account would be same as another ethnographer's account. There would essentially have to be only one explanation of social functioning, and that would be essentially the 'true' one. But soon such a proposition was demonstrated to be imprecise. The famous example is that of Margaret Mead-Derek Freeman controversy. The findings of Margaret Mead based on her ethnographic fieldwork in Samoa were contradicted by Derek Freeman who conducted ethnographic study in the same area. He argued that what Mead had reported was not an unbiased picture of the Samoans but coloured by her own position as a young female student. There was soon an overall agreement that ethnographic research could never be unbiased and could never yield complete knowledge that was free from the subject position of the observer.

Additionally, ethnographic method is time-consuming. There is also a likelihood that the ethnographer is not accepted by the people or not allowed to carry out in-depth fieldwork in the area. There could be many reasons for this e.g., ideological difference between researcher and people under study, overall sense of dislike for the researcher.

9.6 LET US SUM UP

In this Unit, we have discussed the ethnographic method. We began with understanding what ethnography is. We discussed how ethnographic study is carried out. Thereafter, we focused attention on the ethnographic method. We discussed key consideration in use of ethnographic method. Key considerations in use of ethnographic method were discussed. Finally, we dwelt upon the emerging trends in ethnographic research.

9.7 REFERENCES

- Berreman, Gerald D. 2004. 'Ethnography: Method and Product', *Methodology and Fieldwork*, edited by Vinay Kumar Srivastava. New Delhi: Oxford University Press.
- Hammersley, M. and P. Atkinson. 1995. *Ethnography: Principles in Practice*. London: Routledge.
- Malinowski, Bronislaw. 1922. *Argonauts of the Western Pacific*. London: Routledge and Kegan Paul.
- Murthy, Dhiraj. 2008. Digital Ethnography: An Examination of the Use of New Technologies for Social Research. *Sociology* 42(5): 837–855.
- O' Reilly, Karen. 2005. *Ethnographic Methods*. London: Routledge.

9.8 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

- 1) Ethnography as a method of research refers to a composite range of techniques of data connection that are used in combinations to develop a holistic understanding of the way of life of a group. Ethnography as a product of research refers to the descriptive account presented by the researcher as an outcome of research.
- 2) Two distinguishing features of ethnographic research are: first, it is based on data collected by researcher over a prolonged period of time in the course of which they live with the people and participate in their daily activities; and second, it presents people's point of view and interpretation of data is situated in the local context of the people under study.

Check Your Progress 2

- 1) The enterprise of carrying out ethnographic study at many locales is referred as 'multi-sited ethnography'.
- 2) Data for digital ethnography is collected through online questionnaire; digital video; social networking websites; and blogs.