**Unit 2**

**Sampling**

2.1. Meaning, importance, purpose, characteristics. 2.2. Sample size- Rationale for fixing sample size. 2.3. Methods of sampling – random, systematic, stratified and cluster sampling. 2.4. Sampling techniques – probability and non-probability sampling – types and advantages. 2.5. Errors in sampling and its control.

A population is a group of individuals with at least one common characteristics which distinguishes that group from other individuals and that are of interest to the researcher.

A sample is a subgroup of individuals selected from that population. Because of the population is large, it will be impossible for the researcher to study every individual within the potential study population, so the researcher  study a subgroup or sample. As researchers choose a sample for study, they need to make sure that the sample is representative of the larger population.

Sampling is selecting a group (subgroup) from a much larger population that is similar in its trait (i.e. gender, ethnicity, age, income, etc.) distribution of the larger population. Findings made from studying the group can then be generalized to the larger population. This means that statements about the population can be made based on a study that has been conducted. Researchers must be sure to include a large enough sample before drawing any generalizations. This is known as population generalization.

Sampling strategies, or selecting the sampling groups, involves the researcher choosing who will participate in the study. This involves careful thought, taking into consideration relevancy between the issue or phenomena under study and the group chosen to be observed. To do this a researcher develops a set of criteria that defines and sets boundaries between who should and should not be selected. Should only boys or only girls be included in the sample? or Should only  urban boys be included in the sample? etc.

**Purpose of Sampling**

The primary purpose of the sampling is to obtain accurate and reliable information about the characteristics of population with minimum of cost , time and energy and to set out the limits of accuracy of such estimates.

**Importance of Sampling**

1. When the population is very large, it can be satisfactorily covered through sampling.
2. It saves a lot of time, energy and money.
3. Especially when the units of an area are homogeneous, sampling technique is really useful.
4. When the data are unlimited, the use of this method is really useful.
5. When cent percent accuracy is not required, the use of this technique becomes inevitable.
6. When the number of individuals to be studied is manageable, intensive study becomes possible.

**Characteristics of Sampling**

1. A good sample is one which, within restrictions imposed by its size, will reproduce the characteristics of the population with the greatest possible accuracy.
2. It should be free from error due to bias or due to deliberate selection of the unit of the sample.
3. It should be free from random sampling error. It should not be selected by a procedure where there is a connection between the method of selection and the characteristics under consideration.
4. There should not be any substitution of originally selected unit by some other more convenient in any way.
5. It should not suffer from incomplete coverage of the units selected for study, i.e., it should not ignore the failures in the sample in responding to the study.
6. Relatively small samples properly selected may be much more reliable than large samples poorly selected. But at the same time it is very essential that the sample is adequate in size so that it can become really reliable.
7. In the samples only such units should be included, which as far as possible, are independent.
8. While constructing a sample, it is important that measurable or known probability sample techniques are used. This will substantially reduce the likely discrepancies.

**Advantages**

1. Reduced cost: It is economical of the cost when the data are collected from a sample which is only a fraction of the population
2. Greater speed: Sampling is less time consuming and take much less time in the case of a sample than in the case of a population
3. Greater Accuracy: Sampling ensures completeness and a high degree of accuracy due to a limited area of operation.
4. Greater scope: sample simplifies data collection with a little training.
5. Organization of convenience: Sampling involves very few organizational problems.
6. Intensive and exhaustive data: As the number is limited, therefore it is possible to collect intensive and exhaustive data.
7. Better rapport: If the sample is manageable size, it is possible for the researcher to establish meaningful rapport with the respondents.
8. In a small sample, it becomes possible to scrutinize the data collected.

**Disadvantages**

1. Chances of bias: There is a chance for biased selection of samples and thereby lead us to draw erroneous conclusions.
2. Difficulties in selecting a truly representative sample.
3. Requires specialized knowledge in sampling technique.

**Types of Sampling Techniques**

Various sampling techniques can be used depending on the type of research to be conducted. The two major types of techniques are:

1. Probability sampling and
2. Non probability sampling.
3. **Probability Sampling**

Probability sampling provides a scientific technique of drawing samples from population according to some laws of chance in which each unit has some definite pre-assigned probability of being chosen in the sample. Probability sampling procedure specifies the probability that each member of a population has of being selected. Probability sampling techniques include different methods of sampling such as:

1. **Random Sampling**

A group drawn from the population, with every member of the population having an equal chance of being selected. This is the most common and highly recommended technique.

There are four methods are generally used for drawing out a sample on random basis. They are lottery method, use of tables of random numbers, and selecting from sequential list.

*Merits*

1. It is a more scientific method which eliminates personal bias.
2. No advance knowledge of the characteristics of the population is necessary under this method.
3. It is free from errors in classification.
4. This method is simple to use.
5. It is easy to assess the sampling error in this method.
6. The sample drawn under this method is true representative of the universe.
7. This method provides us most reliable and maximum information at the least cost which saves time , money and labor.

*Demerits*

1. The sampling method requires complete list of the universe. But such up to data list is not available in many enquiries which restricts the use of this method.
2. In the case of survey, if the area of coverage is large the sampling are expected to be scattered widely geographically and thus data collection may be time consuming and costly.
3. The selected sample may not be true representative of the universe if its size is too small.
4. For given degree of accuracy it requires larger sample as compared to stratified sampling.
5. **Stratified Sampling**

When the population is heterogeneous with respect to the variable or characteristics under study then the technique of stratified sampling is used to obtain more efficient and accurate results. Stratification means division of the universe into groups according to geographical, sociological or economic characteristics.

*Process of stratifying*

1. The universe is first divided into sub-groups and required units are selected at random from each sub group.
2. The stratification should be conducted in such a way that the items in one stratum should be similar to each other but they should differ significantly from units of other strata.
3. Each and every unit in the population must belong to one and only one stratum. In other words various strata must be non overlapping.
4. The size of each stratum in the universe must be large enough to provide selection of items or random basis.
5. Size of the sample from each strata can either be proportional or disproportional to the size of each stratum.

*Merits*

1. If a correct stratification has been made even a small number of units will form a representative sample.
2. Under this sampling no significant group is left unrepresented.
3. It is more precise and to get a great extent avoids bias.
4. It saves time and cost of data collection since the sample size can be less in this method.
5. It is the only sampling plan which enables us to achieve different degrees of accuracy for different segments of the population.
6. Equal chance of being selected as sample unit.

*Demerits*

1. It is great difficult to divide the universe into homogeneous data.
2. If the strata are overlapping, unsuitable, disproportionate the selection of samples may not be representative.
3. If stratification is faulty, the result obtained may be biased. Such errors cannot be compensated even by taking large samples.
4. **Systematic Sampling**

Under this method a sample is taken from a list prepared on a systematic arrangement either on the basis of alphabetic order or on house number or any other method. In this method only the first sample unit is selected at random and the remaining units are automatically selected in a definite sequence at equal spacing from one another.

*Steps involved in systematic sampling*

1. First the population is arranged in serial numbers and size of the sample is determined.
2. The sampling interval is determined by dividing the population by the size of the sample.
3. Any number is selected at random from the first sampling interval. The sub sequent samples are selected at equal or regular intervals.

*Merits*

1. It is very easy to operate and checking can be done quickly.
2. Randomness and probability features are present in this method which makes sample representative.

*Demerits*

1. It works well only if the complete and up to date frame is available and if the units are randomly arranged.
2. Any hidden periodicity in the list will adversely affect the representativeness of the sample.
3. **Cluster Sampling**

Under this method the total population is divided into some recognizable sub-divisions which are termed as clusters and a simple random sample of these clusters is drawn and then the survey of each and every unit in the selected cluster is made.

*Ex:* to study the problems of middle class working couples in a state, the first stage will be to pick up a few districts in the state. The next stage will be to pick up at random a few rural and urban areas for the study. The third stage will come when from each area a few families belonging to middle class will be picked up. The last stage will be that of selecting working couples out of these families. Thus the stages will be:

State Districts urban rural areas Middle class families working couples in these families

*Principles of cluster sampling*

1. Cluster should be as small as possible with the cost and limitations of the survey.
2. The number of sampling units in each cluster should be approximately same.

*Merits*

1. It is easier and more practical method which facilitates the field work.
2. It is economical, because it is easier and less expensive.
3. It is useful because it may combine the advantages of both random and stratifies sampling.

*Demerits*

1. Probability and the representativeness of the sample is sometimes affected if the number of clusters is very large.
2. The results obtained under this method are likely to be less accurate if the number of sampling units in each cluster are not approximately same.
3. It may not be possible to apply its findings to another area.
4. **Multi-stage sampling**

The method is generally used in selecting a sample from a very large area. As the name suggests multi-stage sampling refers to a sampling technique which is carried out in various stages. Here the population is regarded as made of a number of primary units, each of which is further composed of a number of secondary stage units which is further composed of third stage units and so on till we ultimately reach the desired sampling unit which we are interested. At each stage there is a random selection and size of the sample may be proportional or disproportional depending on the size and character of variations is event to the purpose of inquiry. Thus the area of investigation is scientifically restricted to small number of ultimate units which are representative of whole.

*Merits*

1. It is more flexible when comparing to other methods of sampling.
2. It is simple to carry out.
3. Helps to cover large area
4. It is reliable and satisfactory method.
5. This technique is of greater significance in surveys of under developed areas where no up to date and accurate frame is generally available for subdivision.

*Demerits*

1. Errors are likely to be large in this method when compared to any other method
2. Less efficient than a suitable single stage sampling of the same.
3. **Non-probability Sampling**

Non-probability sampling is based on the personal judgment. Under this method a desired number of sample units are selected deliberately or purposely depending up on the object of the enquiry so that the important items representing the true characteristics of the population are included in the sample. Non probability sampling is used when probability sampling is not feasible. Non probability sampling techniques include different methods of sampling, they are:

1. **Purposive sampling**

Purposive sampling is also known as deliberate sampling or judgment sampling. A researcher uses his or her judgment to select a population that reflects an important aspect of the research. Here investigator purposefully selects and purposefully leaves some members. Purposive sampling is generally more appropriate for qualitative research than quantitative research.

*Merits*

1. The sampling is within the complete control of the investigator.
2. The purpose of the study can be fulfilled even with a small sample which is picked up purposely and carefully.

*Demerits*

1. Chance for biases in sampling because the investigator selects the sample
2. It may not be possible to find out mistakes and inaccuracies till the end
3. Investigator requires well knowledge about the universe to purposive sampling
4. **Quota sampling**

It is another form of purposive sampling and also a form of stratified sampling, except that it refers to a non-probability design in which the investigator, after having stratified his population, uses his judgment rather than randomness in selecting the cases within each of the strata. The number to be selected from each stratum is known as quota.

*Merits*

1. Quota sampling is advantageous over probability sampling with respect to convenience.
2. It less expensive and more quick.
3. It is the only practicable method in the case of population for which no suitable frame is possible.

*Demerits*

1. Chance for bias in the selection.
2. Standard errors cannot be calculated as can be done in the case of a random sample.
3. **Convenience Sampling**

It is known as unsystematic, accidental or opportunistic sampling. Under this method a sample is selected according to the convenience of the investigator. This convenience may be in respect of availability of data, accessibility of the units etc. In educational research, convenient sampling is used frequently by teachers who use their own classes for their research. Findings from such research generally are limited to the population studied and not extended to larger populations.

**Application of sampling technique in various types of research**

**Historical research**

In historical research the problem of sampling is not so important because historical research is based upon past records, events and facts but in Case Study-Method judgement sample and purposive sample is used because the purpose of case study method is to improve the case and not to conclude therefore non “probability sampling is applied.

**Normative survey method**

In this method random sampling is frequently used because in normative survey large sample is selectedsystematic, multi stage and multiple sampling can also be used in normative survey method.

**Experimental method**

In experimental method most precise and comprehensive method of sampling is preferred. Therefore, stratified sampling technique is used but in educational situation it is often difficult to use this stratified sampling than cluster sampling technique is preferred. In the field of education research cluster sampling techniques is most frequently used and it has

some limitations but it has usability in teaching learning situations and educational research.

**Sample size**

The size of the sample means the number of sampling units selected from the population for investigation. The size of the sample is directly related to standard of accuracy, time, cost and administration of the investigation. The larger the size of the sample, greater will be the representation of the items of the universe in it, but it brings difficulties of managing it. In large sized samples intensive study of the units becomes difficult. If the size of the sample is kept small, it will not represent the universe and degree of accuracy in the results will be limited. Hence the size of the sample should be neither too big nor too small. It should be optimum.

According to Parten, “An optimum sample survey is one which fulfills the requirements of efficiency, representativeness, reliability and flexibility.”

The size of the sample should ensure minimum of cost and minimum of standard error. A small sample will have a large standard error compared with a large sample of the same characteristics, while the cost will be less in the first as compared to the latter.

A small random or stratified sample is apt to be much superior to a large but badly selected sample.

**Rationale for fixing sample size**

The size of the sample which is required in order to make a valid inference about a population is dependent up on several factors such as:

1. Homogeneity or heterogeneity of universe

In a universe consisting of homogeneous units a small sample is suitable, while in a universe consisting of heterogeneous units, a large sized sample is inevitable, for yielding good results.

1. Number of classes proposed

If it is necessary to classify data in a large number of classes, a large sized sample should be taken to facilitate analysis of data. If the size of the sample is small and classes are many, then there will be same classes which may not get due representation in the sample. If the proposed number of classes is more, the size of the sample should be big.

1. Nature of study

Nature of study also affects the size of a sample. For an intensive and continuous study small sample will be suitable, because such study in a large sized sample will require more resources. For general survey the size of the sample should be large, but for technical studies the size of the sample may kept be small.

1. Practical considerations

The availability of finance, time and trained personnel are other practical considerations which affect the size of the sample.

1. Standard of accuracy

Though it is believed that larger is the size of the sample, great will be the degree of accuracy, yet this is not always so. If a sample is selected by experts through scientific method, a small sample cam produce more accurate results.

1. Type of sampling

In random sampling, greater accuracy in results will be achieved only in a large sample. In a properly drawn stratified sampling, a small sample can give better results.

1. Nature of units

Where it is expected that a large number of units will not respond, then a large sample should be taken.

1. Size of questionnaire

If the size of the questionnaire is large and it contains difficult questions, the size of the sample should be kept small.

**Sampling error**

Sampling errors occurs when estimates are derived a sample rather than a census of the population. The difference between sample estimate(statistics) and the population value(parameter) is called error. The measures estimated from the samples are called statistics, which will vary from sample to sample drawn from the same population due to fluctuation. Descriptive measures of population are called parameters.

Errors in sampling are of two types.

1. Sampling error (random error)
2. Non-sampling error ( systematic error)
3. **Sampling error**

The sampling error is quantified using SE or standard error. The errors which are due to sampling and of which average magnitude can be determined are called sampling errors.

With probability samples, the SE can be estimated using sample design and sample data. As the sample size increases sampling error decreases. If the population is relatively homogeneous, the SE will small. If the population is heterogeneous, SE will be high.

So we can say that SE is avoidable errors that occur whenever sampling is done. The sampling error occurs by chance.

1. **Non-sampling error**

The non-sampling error is also known as systematic error and this arises due to

1. Selection bias
2. Non-response bias
3. Response bias

Selection bias is a systematic tendency to exclude one kind of unit or another from the sample.

Non-response bias means the difficulty in the extrapolation from respondents to non respondents.

Response error means difference in the responses by different participants.

There is no simple and direct way of estimating size of non-sampling error.

**Reasons for non sampling error**

1. Sampling operations- There may be errors in sample selection, or part of the population may be omitted from the sampling frame.
2. Non interviews: Information is generally obtained for only part of the sample, Therefore there are differences between the interview population and those interviewed.
3. Adequacy of respondent: Sometimes respondents cannot be interviewed and information is obtained about them from others, but the poor respondent is not always as knowledgeable about the facts.
4. Understanding the concepts: Some respondents may not understand what investigator wanted.
5. Lack of knowledge
6. Concealment of the truth: Respondents may conceal the truth. Here the respondent may answer in a socially acceptable way when it is not actually so.
7. Loaded question: The questions may be worded to influence the respondent to answer in a specific way.
8. Processing errors: Include coding error, data keying, computer programming etc.
9. Conceptual problems: There may be difference between what is desired and what actually over.
10. Interviewer error: Interviewer may miss the question or twist the answers in their own words.

Mouly has classified sampling error in a four-way classification table.

|  |  |  |
| --- | --- | --- |
|  | Random | Systematic |
| Sampling | A | B |
| Measurement | C | D |

The cell A and B refers to sampling errors and C and D refers to non-sampling error.

The cell A refers to unavoidable errors that occur whenever sampling is done. Some shift in the sample statistics is caused by selecting at random any individual who may be high, low or average in the trait in question. Cell B refers to constant error. These are also called errors of bias in sampling, that is errors which do not cancel out each other but changes systematically in one or other direction of population.

Cell C refers to random error of measurement and are due to unreliability of the testing instrument.

Cell D refers to systematic errors of measurement. This bias arises mainly because of all the discrepancies between our observations. And the quantities we aim to measure of the systematic non cancelling type.

**Controlling of sampling errors**

Sampling error must be reduced to the minimum so that the results and conclusions are sufficiently representative of whole universe. This can be done by:

1. Keeping survey focused on measurable objectives.
2. Removing errors of bias from sample.
3. Enlarging the size of the sample. Sampling error usually decrease with increase in sampling size.
4. Survey frames should be large.
5. Using standardized tools for data collection.
6. Making sure of response rates.
7. Using proper editing and interpretation procedure in the processing system.
8. Using standard procedures and process.

**References**

1. Donna M.Mertens .(2005) Research and Evaluation in Education and Psychology -Integrating Diversity with Quantitative ,Qualitative and Mixed Method ,2nded .Sage Publication p . 140.
2. Koul,Lokesh, (2006). Methodology of Educational Research .Vikas Publishing House Pvt Ltd,Noida .
3. Appa Iyer Sivakumar ,Krishnasawamy.K.N;M .Mathirajan.(2006).Management Research Methodology : Integeration of Principles ,Methods and Techniques , Pearson Education , Delhi.
4. Chauhan , Sandeep (2002) .Research Methodology in Education , Pearson Education , Delhi.
5. Singh ,Y.K and Bajpai,R.B.(2008) Research Methodology-Data Presentation ,APH Publishing Corporation,Delhi.
6. Guptha, S.(2001). *Research methodology and statistical techniques*. Deep & Deep publications: New Delhi.
7. Singh, S.K. (1985). *Methodology of Research Education. Sterlings Publications PVT LTD.: New Delhi.*