SAMPLING TECHNIQUES AND SAMPLING ERROR

The first step in sampling is the identification of the target population, the large group to which the researcher wishes to generalize the results of the study. Once we have identified the population, the next step is to select the sample.

SAMPLING TECHNIQUES

Two major types of sampling procedures are available to researchers: probability and non-probability sampling. Probability sampling involves sample selection in which the elements are drawn by chance procedures. The main characteristic of probability sampling is that every member or element of the population has a known probability of being chosen in the sample. Non-probability sampling includes methods of selection in which elements are not chosen by chance procedures. Its success depends on the knowledge, expertise, and judgment of the researcher. Non-probability sampling is used when the application of probability sampling is not feasible. Its advantages are convenience and economy.

PROBABILITY SAMPLING

Probability sampling is defined as the kind of sampling in which every element in the population has an equal chance of being selected. The possible inclusion of each population element in this kind of sampling takes place by chance and is attained through random selection. When probability sampling is used, inferential statistics enable researchers to estimate the extent to which the findings based on the sample are likely to differ from what they would have found by studying the whole population. The four types of probability sampling most frequently used in educational research are simple random sampling, stratified sampling, cluster sampling, and systematic sampling.

Simple Random Sampling

The best known of the probability sampling procedures is simple random sampling. The basic characteristic of simple random sampling is that all members of the population have an equal and independent chance of being included in the random sample. The steps in simple random sampling comprise the following:

a. Define the population.
b. List all members of the population.
c. Select the sample by employing a procedure where sheer chance determines which members on the list are drawn for the sample.

Random sampling is the best way to obtain a representative sample. It is required for many statistical analyses. These analyses permit the researcher to make inferences about a population based on the behavior of a sample. If samples are not randomly selected, then a major assumption of many statistical analyses is violated, and inferences made from the research are suspect.

Steps in Simple Random Sampling

In general, random sampling involves defining the population, identifying each member of the population, and selecting individuals for the sample on a completely chance basis. One way to do this is to write each individual's name on a separate slip of paper, place all the slips in a hat or other container, shake the container, and select slips from the container until the desired number of participants is selected. This procedure is not exactly satisfactory if a population has 1,000 or more members. One would need a very large hat—and a strong writing hand! A much more satisfactory approach is to use a table of random numbers, also called a table of random digits.

Stratified Sampling

When the population consists of a number of subgroups, or strata that may differ in the characteristics being studied, it is often desirable to use a form of probability sampling called stratified sampling. In stratified sampling, researcher first identifies the strata of interest and then randomly draw a specified number of subjects from each stratum. The basis for stratification may be geographic or may involve characteristics of the population such as income, occupation, gender, age, year in college, or teaching level. In studying adolescents, for example, researcher might be interested not merely in surveying the attitudes of adolescents toward certain phenomena but also in comparing the attitudes of adolescents who reside in small towns with those who live in medium-size and large cities. In such a case, investigator would divide the adolescent population into three groups based on the size of the towns or cities in which they reside and then randomly select independent samples from each stratum.

Steps in stratified sampling

a. Identify and define the population.
b. Determine desired sample size.
c. Identify the variable and subgroups (i.e., strata) for which you want to guarantee a specific representation.
d. Classify all members of the population as members of one of the identified subgroups.
e. Randomly select (using a table of random numbers) an equal number of individuals from each subgroup.

An advantage of stratified sampling is that it enables the researcher to also study the differences that might exist between various subgroups of a population. In this kind of sampling, you may either take equal numbers from each stratum or select in proportion to the size of the stratum in the population. The latter procedure is known as proportional stratified sampling, which is applied when the characteristics of the entire population are the main concern in the study. Each stratum is represented in the sample in exact proportion to its frequency in the total population. The major advantage of stratified sampling is that it guarantees representation of defined groups in the population.

Systematic Sampling

Another form of probability sampling is called systematic sampling. First, you decide how many subjects you want in the sample (n). Because you know the total number of members in the population (N), you simply divide N by n and determine the sampling interval (K) to apply to the list. Select the first member randomly from the first K members of the list and then select every

Kth member of the population for the sample. For example, let us assume a total population of 500 subjects and a desired sample size of 50: $K = N/n = 500/50 = 10$.

Steps in Systematic Sampling

Systematic sampling involves the following steps:

a. Identify and define the population.
b. Determine the desired sample size.
c. Obtain a list of the population.
d. Determine *K* by dividing the size of the population by the desired sample size.
e. Start at some random place in the population list. Close your eyes and stick your finger on a name.
f. Starting at that point, take every *K* th name on the list until the desired sample size is reached.
g. If the end of the list is reached before the desired sample is reached, go back to the top of the list.

Systematic sampling differs from simple random sampling in that the various choices are not independent. Once the first case is chosen, all subsequent cases to be included in the sample are automatically determined. If the original population list is in random order, systematic sampling would yield a sample that could be statistically considered a reasonable substitute for a random sample.

Cluster Sampling

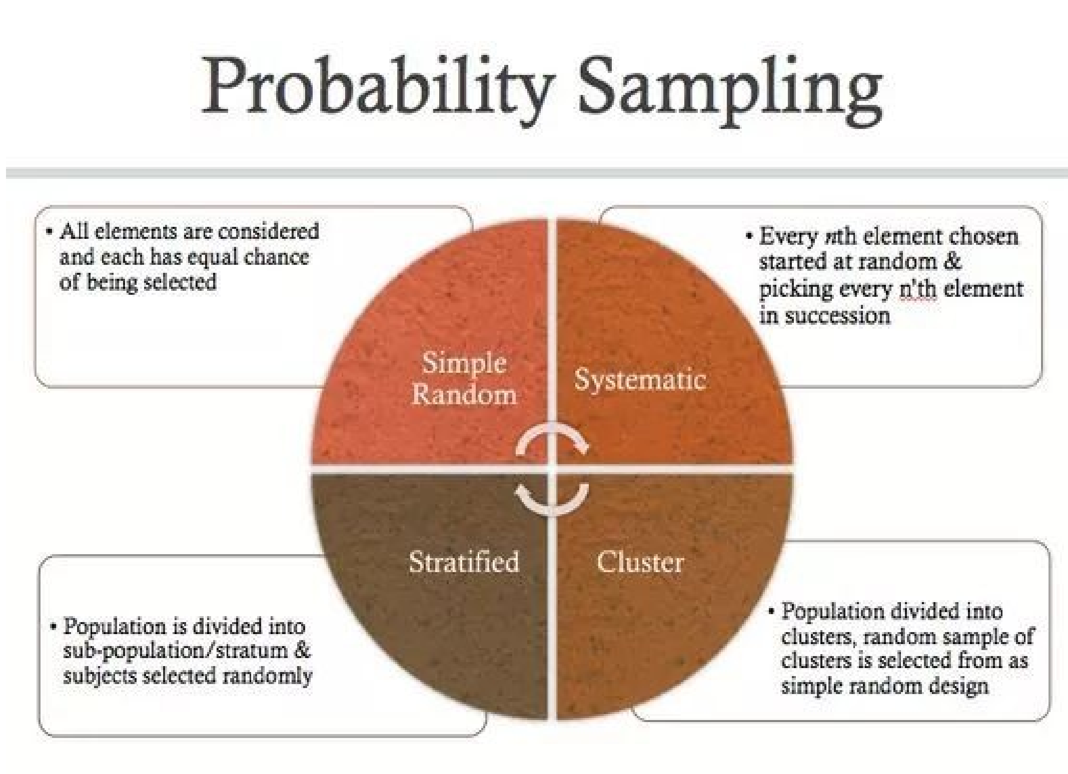
In cluster sampling, intact groups, not individuals are randomly selected. Any location within which we find an intact group of population members with similar characteristics is a cluster. Examples of clusters are classrooms, schools, city blocks, hospitals and department stores. Cluster sampling may be the only feasible method of selecting a sample when the researcher is unable to obtain a list of all members of the population. It is also convenient when the population is very large or spread over a wide geographic area. Cluster sampling usually involves less time and expense and is generally more convenient than either simple random sampling or stratified sampling.

Steps in Cluster Sampling

a. Identify and define the population.
b. Determine the desired sample size.
c. Identify and define a logical cluster (e.g., neighborhood, school, city block).
d. List all clusters (or obtain a list) that make up the population of clusters.
e. Estimate the average number of population members per cluster.
f. Determine the number of clusters needed by dividing the sample size by the estimated size of a cluster.
g. Randomly select the needed number of clusters, using a table of random numbers.
h. Include in your study all population members in each selected cluster.

Multistage Sampling

The sample to be studied is selected at random at different stages. For example, we need to select a sample of middle class working couples in Odisha state. The first stage will be randomly selecting a specific number of districts in a state. The second stage involves selecting randomly a specific number of rural and urban areas for the study. At the third stage, from each area, a specific number of middle class families will be selected and at the last stage, working couples will be selected from these families.



NONPROBABILITY SAMPLING

Non-probability sampling, also called nonrandom sampling, is the process of selecting a sample using a technique that does not permit the researcher to specify the probability, or chance, that each member of a population has of being selected for the sample. Nonrandom sampling methods do not have random sampling at any stage of sample selection and can introduce sampling bias.

When nonrandom samples are used, it is usually difficult, if not impossible, to describe the population from which a sample was drawn and to whom results can be generalized. To compensate for this problem, the researcher may obtain information from non respondents. Often, follow-up contact with non respondents provides the researcher with insights about potential bias provided by the respondents.

Nonrandom sampling approaches include convenience sampling, purposive sampling, and quota sampling. Of these methods, convenience sampling is the most used in educational research and is therefore the major source of sampling bias in educational research studies.

Convenience Sampling

Convenience sampling also referred to as accidental sampling or haphazard sampling is the process of including whoever happens to be available at the time. It is regarded as the weakest of all sampling procedures, involves using available cases for a study. Interviewing the first individuals you encounter on campus, using a large undergraduate class, using the students in your own classroom as a sample, or taking volunteers to be interviewed in survey research are various examples of convenience sampling. There is no way (except by repeating the study using probability sampling) of estimating the error introduced by the convenience sampling procedures. Probability sampling is the ideal, but in practice, convenience sampling may be all that is available to a researcher. In this case, a convenience sample is perhaps better than nothing at all. If you do use convenience sampling, be extremely cautious in interpreting the findings and know that you cannot generalize the findings.

Purposive Sampling

In purposive sampling—also referred to as judgment sampling—sample elements judged to be typical, or representative, are chosen from the population. The assumption is that errors of judgment in the selection will counterbalance one another. Researchers often use purposive sampling for forecasting national elections. In each state, they choose a number of small districts whose returns in previous elections have been typical of the entire state. They interview all the eligible voters in these districts and use the results to predict the voting patterns of the state. Using similar procedures in all states, the pollsters forecast the national results. The critical question in purposive sampling is the extent to which judgment can be relied on to arrive at a typical sample. There is no reason to assume that the units judged to be typical of the population will continue to be typical over a period of time. Consequently, the results of a study using purposive sampling may be misleading. Because of its low cost and convenience, purposive sampling has been useful in attitude and opinion surveys. Be aware of the limitations, however, and use the method with extreme caution.

Quota Sampling

Quota sampling involves selecting typical cases from diverse strata of a population. The quotas are based on known characteristics of the population to which you wish to generalize. Elements are drawn so that the resulting sample is a miniature approximation of the population with respect to the selected characteristics. For example, if census results show that 25 percent of the population of an urban area lives in the suburbs, then 25 percent of the sample should come from the suburbs.

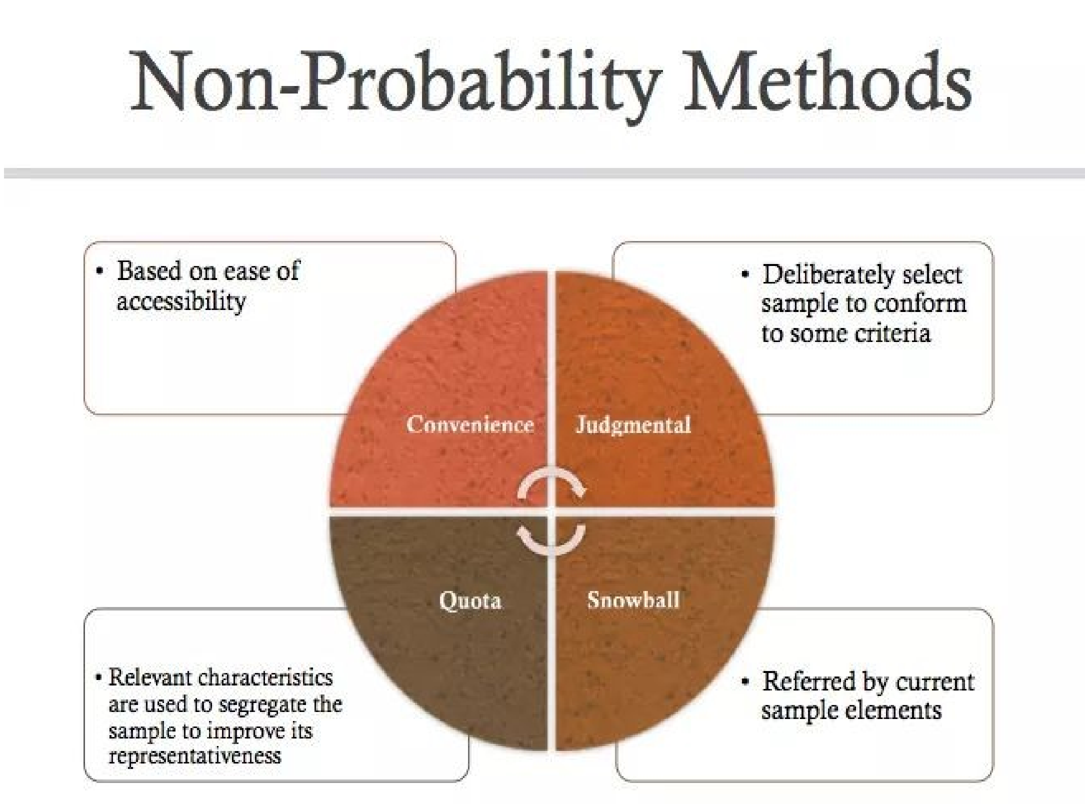
Steps in quota sampling

a. Determine a number of variables, strongly related to the question under investigation, to be used as bases for stratification. Variables such as gender, age, education, and social class are frequently used.

b. Using census or other available data, determine the size of each segment of the population.
c. Compute quotas for each segment of the population that are proportional to the size of each segment.
d. Select typical cases from each segment, or stratum, of the population to fill the quotas.

Snowball Sampling

In snowball sampling, the researcher identifying and selecting available respondents who meet the criteria for inclusion in his/her study. After the data have been collected from the subject, the researcher asks for a referral of other individuals, who would also meet the criteria and represent the population of concern.



THE SIZE OF THE SAMPLE

A larger sample is more likely to be a good representative of the population than a smaller sample. However, the most important characteristic of a sample is its representativeness, not its size. A random sample of 200 is better than a random sample of 100, but a random sample of 100 is better than a biased sample of 2.5 million. Size alone will not guarantee accuracy. A sample may be large and still contain a bias.

SAMPLING ERROR

When an inference is made from a sample to a population, a certain amount of error is involved because even random samples can be expected to vary from one to another. The mean intelligence score of one random sample of fourth-graders will probably differ from the mean intelligence score of another random sample of fourth-graders from the same population. Such differences, called sampling errors, result from the fact that the researcher has observed only a sample and not the entire population.

Sampling error is "the difference between a population parameter and a sample statistic." For example, if you know the mean of the entire population (symbolized μ) and also the mean of a random sample (symbolized-$\bar{X}$) from that population, the difference between these two ($\bar{X} - \mu$) represents sampling error (symbolized e). Thus, $e = \bar{X} - \mu$. For example, if you know that the mean intelligence score for a population of 10,000 fourth-graders is $\mu = 100$ and a particular random sample of 200 has a mean of $\bar{X} = 99$, then the sampling error is $\bar{X} - \mu = 99 - 100 = -1$. Because we usually depend on sample statistics to estimate population parameters, the notion of how samples are expected to vary from populations is a basic element in inferential statistics. However, instead of trying to determine the discrepancy between a sample statistic and the population parameter (which is not often known), the approach in inferential statistics is to estimate the variability that could be expected in the statistics from a number of different random samples drawn from the same population. Because each of the sample statistics is considered to be an estimate of the same population parameter, any variation among sample statistics must be attributed to sampling error.

Suggested Reading

Ary, D., Jacobs, L. C., Sorensen, C., and Razavieh, A. (2010), Introduction to Research in Education, Wadsworth Cengage Learning, Canada.

Best, J. W. and Kahn, J. V. (1995), Research in Education, Prentice Hall, New Delhi.

Cohen, L., Manison, L., and Morrison, K. (2018), Research Methods in Education, Routledge, London and New York.

Creswell, J. W. (2012), Educational Research Planning, Conducting, and Evaluating Quantitative and Qualitative Research, Pearson, New York.

Gay, L. R., Mills, G.E., and Airasian, P. W. (2012), Educational Research Competencies for Analysis and Application, Pearson, New York.

Kerlinger, Fred. N. (1978), Foundations of Behavioral Research, New York University.

Kothari, R.C. (2004), Research Methodology, New Delhi, New Age International (P) Limited, Publishers.

Langenbach, M., Vaughn, C., and Aagaard, L. (1993), An Introduction to Educational Research, Allyn and Bacon, Boston.