Dr. Betka Rezig Nadia

Level : Master One (Literature and Civilization)

Subject : Research Methodology

Research Design: Sampling in Social Sciences

Sampling is the statistical process of selecting a subset (called a "sample") of a population of interest for purposes of making observations and statistical inferences about that population. Social science research is generally about inferring patterns of behaviors within specific populations. We cannot study entire populations because of feasibility and cost constraints, and hence, we must select a representative sample from the population of interest for observation and analysis. It is extremely important to choose a sample that is truly representative of the population so that the inferences derived from the sample can be generalized back to the population of interest.

The Sampling Process

The sampling process comprises of several stage.

The first stage is defining the target population. A population can be defined as all people or items (unit of analysis) with the characteristics that one wishes to study. The unit of analysis may be a person, group, organization, country, object, or any other entity that you wish to draw scientific inferences about.

The second step in the sampling process is to choose a sampling frame . This is an accessible section of the target population (usually a list with contact information) from where a sample can be drawn. If your target population is professional employees at work, because you cannot access all professional employees around the world, a more realistic sampling frame will be employee lists of one or two local companies that are willing to participate in your study

The last step in sampling is choosing a sample from the sampling frame using a well-defined sampling technique. Sampling techniques can be grouped into two broad categories: probability (random) sampling and non-probability sampling. Probability sampling is ideal if generalizability of results is important for your study, but there may be unique circumstances where non-probability sampling can also be justified. These techniques are discussed in the next two sections.

A- Probability Sampling

Probability sampling is a technique in which every unit in the population has a chance (non-zero probability) of being selected in the sample, and this chance can be accurately determined. Sample statistics thus produced, such as sample mean or standard deviation, are unbiased estimates of population parameters, as long as the sampled units are weighted according to their probability of selection. All probability sampling have two attributes in common:

(1) every unit in the population has a known non-zero probability of being sampled, and

(2) the sampling procedure involves random selection at some point. The different types of probability sampling techniques include:

1-**Simple random sampling**. In this technique, all possible subsets of a population are given an equal probability of being selected. The probability of selecting any set of n units out of a total of N units in a sampling frame is N C n . Hence, sample statistics are unbiased estimates of population parameters, without any weighting. Simple random sampling involves randomly selecting respondents from a sampling frame, but with large sampling frames, usually a table of random numbers or a computerized random number generator is used.

2-Systematic sampling. In this technique, the sampling frame is ordered according to some criteria and elements are selected at regular intervals through that ordered list. Systematic sampling involves a random start and then proceeds with the selection of every k th element from that point onwards, where k = N / n , where k is the ratio of sampling frame size N and the desired sample size n , and is formally called the sampling ratio . It is important that the starting point is not automatically the first in the list, but is instead randomly chosen from within the first k elements on the list.

3-Stratified sampling. In stratified sampling, the sampling frame is divided into homogeneous and non-overlapping subgroups (called "strata"), and a simple random sample is drawn within each subgroup.

4-Cluster sampling. If you have a population dispersed over a wide geographic region, it may not be feasible to conduct a simple random sampling of the entire population. In such case, it may be reasonable to divide the population into "clusters" (usually along geographic boundaries), randomly sample a few clusters, and measure all units within that cluster.

4-Matched-pairs sampling. Sometimes, researchers may want to compare two subgroups within one population based on a specific criterion. For instance, why are some firms consistently more profitable than other firms? To conduct such a study, you would have to categorize a sampling frame of firms into "high profitable" firms and "low profitable firms" based on gross margins, earnings per share, or some other measure of profitability. You would then select a simple random sample of firms in one subgroup, and match each firm in this group with a firm in the second subgroup, based on its size, industry segment, and/or

other matching criteria. Now, you have two matched samples of high-profitability and low-profitability firms that you can study in greater detail. Such matched-pairs sampling technique is often an ideal way of understanding bipolar differences between different subgroups within a given population.

5-Multi-stage sampling. The probability sampling techniques described previously are all examples of single-stage sampling techniques. Depending on your sampling needs, you may combine these single-stage techniques to conduct multi-stage sampling. For instance, you can stratify a list of businesses based on firm size, and then conduct systematic sampling within each stratum. This is a two-stage combination of stratified and systematic sampling. Likewise, you can start with a cluster of school districts in the state of New York, and within each cluster, select a simple random sample of schools; within each school, select a simple random sample of grade levels; and within each grade level, select a simple random sample of students for study. In this case, you have a four-stage sampling process consisting of cluster and simple random sampling.

B- Non-Probability Sampling

Nonprobability sampling is a sampling technique in which some units of the population have zero chance of selection or where the probability of selection cannot be accurately determined. Typically, units are selected based on certain non-random criteria, such as quota or convenience. Because selection is non-random, nonprobability sampling does not allow the estimation of sampling errors, and may be subjected to a sampling bias. Therefore, information from a sample cannot be generalized back to the population. Types of non-probability sampling techniques include:

1- Convenience sampling. Also called accidental or opportunity sampling, this is a technique in which a sample is drawn from that part of the population that is close to hand, readily available, or convenient

This is a non-probability sample because you are systematically excluding all people who shop at other shopping centers.

Other examples of convenience sampling are sampling students registered in a certain class or sampling patients arriving at a certain medical clinic. This type of sampling is most useful for pilot testing, where the goal is instrument testing or measurement validation rather than obtaining generalizable inferences.

3-Expert sampling. This is a technique where respondents are chosen in a non-random manner based on their expertise on the phenomenon being studied. For instance, in order to understand the impacts of a new governmental policy such as the Sarbanes-Oxley Act, you can sample an group of corporate accountants who are familiar with this act. The advantage of this approach is that since experts tend to be more familiar with the subject matter than non-experts, opinions from a sample of experts are more credible than a sample

that includes both experts and non-experts, although the findings are still not generalizable to the overall population at large.

4-Snowball sampling. In snowball sampling, you start by identifying a few respondents that match the criteria for inclusion in your study, and then ask them to recommend others they know who also meet your selection criteria. For instance, if you wish to survey computer network administrators and you know of only one or two such people, you can start with them and ask them to recommend others who also do network administration. Although this method hardly leads to representative samples, it may sometimes be the only way to reach hard-to-reach populations or when no sampling frame is available.