10.2478/seeur-2022-0023

SAMPLING TECHNIQUES (PROBABILITY) FOR QUANTITATIVE SOCIAL SCIENCE RESEARCHERS: A CONCEPTUAL GUIDELINES WITH EXAMPLES

Md. Mizanur Rahman
BRAC Business School
BRAC University
mizanur.rahman@bracu.ac.bd

Mosab I. Tabash
College of Business, Al Ain University, United Arab Emirates.
mosab.tahash@aau.ac.ae

Aidin Salamzadeh
Faculty of Management, University of Tehran, Tehran, Iran
salamzadeh@ut.ac.ir

Selajdin Abduli
South East European University, Tetovo, North Macedonia
s.abduli@seeu.edu.mk

Md. Saidur Rahaman
Metropolitan University Sylhet
saidurmgt@gmail.com

ABSTRACT

Collecting data using an appropriate sampling technique is a challenging task for a researcher to do. The researchers will be unable to collect data from all possible situations, which will preclude them from answering the study's research questions in their current form. In light of the enormous number and variety of sampling techniques/methods available, the researcher must be knowledgeable about the differences to select the most appropriate sampling technique/method for the specific study under consideration. In this context, this study also looks into the basic concepts in probability sampling, kinds of probability sampling techniques with their advantages and disadvantages. Social science researchers will benefit from this study since it will assist them in choosing the most suitable probability sampling technique(s) for completing their research smoothly and successfully.

Key words: Sample, sampling, probability sampling, quantitative research, social science

BACKGROUND

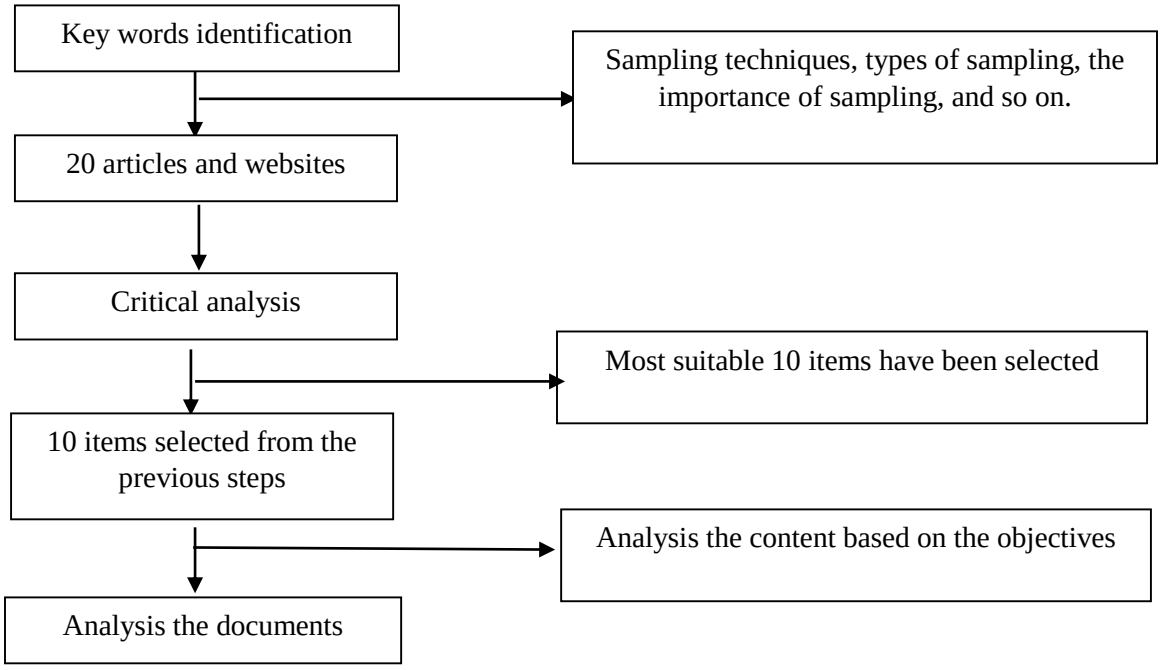
Research is the process of determining how to solve an issue, obtaining additional knowledge about it, or devising a better solution to address the problem. To accomplish its objective, research must be organized and carried out systematically. Sampling procedures are a critical component of quantitative research. The primary aim of sampling is to obtain a representative sample composed of a small number of units or instances drawn from a much larger group or population. Thus, the researcher can research the smaller group and make valid generalizations about the more extensive group based on the findings of the smaller group. The researchers concentrate on the approaches that would provide them with highly comparable samples (i.e., very much like the population). They employ probability sampling, a sort of sampling based on mathematical ideas about the likelihood of certain events occurring. Based on the sample data, a determination can be made about the entire population. Samples assist in reducing the time and cost associated with conducting a survey. On the other hand, the probability sampling must be sufficiently representative to enable the study population to conclude. Cost and time constraints would preclude a comprehensive study of the entire population in this type of study. Thus, sampling is the optimal method for doing research in the social sciences. However,

the goal of this research is to investigate a few of these concerns in relation to the following questions:

- Is there any confusion between sample and sampling?
- Why are probability sampling techniques essential for social science researchers?
- What are the stages of the probability sampling technique?
- What are the different techniques of probability sampling techniques?

METHODS/DESIGN

This is an informative general literature review article that will provide clear guidelines with necessary examples for social science researchers interested in using the probability sampling approaches in their research works. A slight modified version of a five-step literature review technique developed by Mascarenhas et al. (2018) has been used in this study. However, instead of five steps, we have used the following four steps:



Source: Authors' creation

After selecting the study's goal, the researchers began looking for relevant papers in WoS, SCOPUS, Google Scholar, DOAJ, JSTOR, and other relevant websites using keywords linked to sampling techniques, types of sampling, the importance of sampling, and so on. Finally, based on 10 papers and websites, the researchers used literature to support their conceptual definitions of probability sampling strategies and their benefits and downsides. However, all examples are based on the researchers' personal experiences.

SAMPLE AND SAMPLING: DO NOT BE CONFUSED

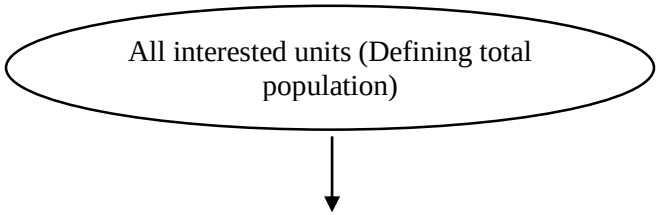
In quantitative data analysis, a sample is a subset of the total amount of data that has been gathered by surveys or lengthy observations. It can be thought of as a smaller unit of measurement that represents the actual data. Besides, to select the respondents within a population is also difficult for quantitative researchers (Dana and Dana 2005; Hanlon and Larger, 2011)

Example: 500 postgraduate students from Bangladesh have completed a survey relating to psychological status during COVID-19.

In relation to the above example, the method that will applied for collecting those samples is called sampling. Sampling is a critical component of most studies because it is the foundation of virtually all research. Sampling is described as the act, process, or technique of selecting a representative sample of a population to observe and analyze the characteristics of the entire population. To put it mildly, sampling is defined as the process of selecting a random sample from a population using specialized sampling procedures. While collecting samples, you can take a variety of approaches. We will discuss the approaches later.

Stages of probability sampling technique

Before selecting the appropriate probability sampling techniques, it is an important responsibility to know the stages of probability sampling techniques. Usually, the researchers need to focus on two important stages before selecting the techniques, which are 1. Defining total population 2. Defining the sampling frame. Thus, it is clear that there are three stages of probability sampling techniques shown in the following figure.



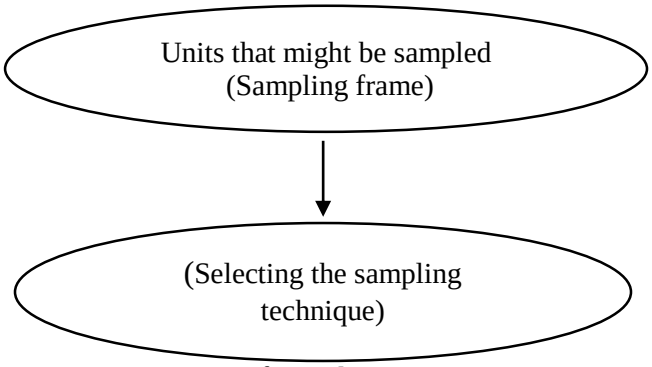


Figure 1: Stages of sampling process

Defining the target population (all interested units):

The "target population" is a group of people or items about which researchers wish to make broad generalizations. In some cases, a sample can be obtained from a group of people, often known as a "target population."

Example: A researcher wants to see the condition of psychological status of postgraduate students of Bangladesh during COVID-19. In that case, all of the postgraduate students will be treated “target population”.

Selecting the sampling frame (units might be sampled):

A complete list of sample units from a population is usually called a sampling frame. One thing that needs to remember; a population differs from a sample frame because a population is more general, whereas a sample frame is more specific.

Example: Now, the researcher has decided to collect the data from the postgraduate students of 10 private universities in Bangladesh. Thus, all of the postgraduate students of these 10 private universities are the sampling frame.

Selecting the sampling techniques:

Sampling is the process of selecting a sample from a population. Sampling procedures are critical in social science and other experimental research (Suresh et al., 2011). The two main types of sampling techniques are probability sampling and non-probability sampling (Elfil and Negida, 2017; Shorten, Moorley, 2014]. However, in this article, the researchers have only focused on the probability sampling techniques.

PROBABILITY SAMPLING:

When everyone in the population has an equal chance of being chosen is usually called the probability sampling technique. It is thought to be the best way to ensure that all sampling units are equally representative of their populations (Curtin et al., 2005; Fowler, 2009).

Why the researchers will use this technique?

The primary purpose of probability sampling is to obtain a sample that is representative of the population from which it was taken. The use of random sampling does not imply that each sample is 100% representative of the population. Instead, it indicates that most random samples will be near to the population most of the time and that it is possible to determine how probable a sample is correct.

Example: In the 10 private universities, there are 5000 postgraduate students. So, under probability sampling, when each student gets an equal chance to participate in the mentioned research, it will be considered probability sampling.

COMMON TYPES OF PROBABILITY SAMPLING:

Simple random sampling: when, how and why not

Simple random sampling is a sampling method that ensures that each member of a population has an equal chance of being chosen as a respondent (Thomas, 2020). For specific findings in social science research investigations, this strategy is used. Simple random sampling works well when the study's goal is to find a generalized result that can be applied to the entire population. In social science research, simple random sampling, on the other hand, has a variety of limitations and obstacles. One of the most notable drawbacks of simple random sampling, for example, is that it cannot be used in situations where the population units are heterogeneous.

Example: Among the 5000 postgraduate students, when 500 students will be selected either using random number table or by lottery method will be treated as simple random sampling.

Systematic random sampling: when, how and why not

The initial subject in systematic random sampling is picked at random, while the subsequent subjects are chosen in a systematic manner. It is one of the procedures in which every Kth item is chosen, where K is the number of items in the sampling frame divided by the number

of items required for the sample size (Acharya et al., 2013). A random approach is utilized to choose a beginning point, and then every Kth number on the list is selected. Researchers will be able to perform, analyze, and compare the samples in a rapid manner because only a few random components are required to select the first person. In other words, this method saves both time and money. Furthermore, the costs of implementation are cheaper. Besides, if the population's size in a particular demographic is not defined, then this method cannot run properly. And for this, different areas such as research on animals are in the field area; this method is not that useful.

Example: Using the same example with N=5000 and n=500, the sampling interval is k = N/n, or 10. Now, the researcher can pick 1 and 10. In this situation (he) can select a random number between 1 and 10. Suppose s(he) has chosen 5 as his or her first respondent. After that, add "10" to that number. Then 15, 25, 35, 40, and so on until the sample size of 5000 was attained.

Stratified Sampling: when, how and why not

The data is classified into multiple subgroups (strata) based on common characteristics such as age, gender, race, income, education, and ethnic origin. Each stratum is randomly sampled. The advantages are that it ensures representation of all parts of the population due to this. Stratified random sampling provides better population coverage since the researchers have more control over the subgroups and ensure that they are included. Stratified sampling techniques can be classified into two types.

Proportionate Stratified Sampling

In proportionate stratified sampling, variables are selected for the sample based on their original distribution in the population of interest. This shows that the relative size of a stratum in the population of interest determines the likelihood of selecting a variable from that stratum for the sample.

Example: In the 10 private universities, there are 5000 postgraduate students. Among them, 3000 students are male, and 2000 students are female. Now the researcher wants 1/10=500 respondents. It must be 500/1.66=300 male and 500/2.5=200 female students as per proportionate stratified sampling.

Disproportionate Stratified Sampling

It is common for social science researchers to employ this method when they do not select a sample representing the entire population they wish to research. The idea is that people from various groups will not have the same possibilities of being included in the research sample.

Example: The researcher needs 500 postgraduate students. In this method, the researcher does not need the male or female ratio. Just s(he) needs 500 respondents.

Cluster sampling: when, how and why not

It is a two-step procedure in which the overall population is divided into groups or clusters, typically locales such as villages, schools, wards, blocks, and other similar locations (Acharya et al., 2013). Epidemiologic studies are more prone than clinical studies to have this type of incident occur to their participants. It is particularly beneficial for surveys that are conducted throughout a large geographic area. Clusters are picked at random from a pool of respondents. Every single respondent in the cluster has been included in the sample population. The majority of the time, larger sample size is required. When a population is dispersed, and it is impossible to obtain a representative sample of all elements, cluster sampling is a useful method of gathering information about them. As with stratified sampling, a cluster sample is more successful when its elements are not all the same kind of thing, unlike stratified sampling. There are certain advantages to cluster sampling, such as its low cost and speed, but there are also many disadvantages, such as the high potential for inaccuracy (Etikan and Babtope 2019).

Example: In this situation, if the researcher needs 500 postgraduate students from different districts, it will be treated as cluster sampling. However, in order to get these 500 respondents, the researcher can use either simple random or systematic random sampling.

Multi-stage sampling technique: when, how and why not

Multi-stage sampling is also known as multi-stage cluster sampling. It is one of the complex forms of cluster sampling that consists of two or more sample selection steps in one operation. In another way, it is the process of dividing large populations into smaller clusters to make primary data collection more effective and efficient. The researchers' usually employ this technique to avoid the problems that come with randomly sampling from a large group of people.

Researchers with limited resources and time can take the sample from groups like this because they use multi-stage sampling. In a nutshell, this procedure is a way to cut down on the size of the population by breaking it up into smaller groups, which can then be randomly chosen from. People with minimal between-group variance can use this type of sampling to make the population easier to understand. On the other hand, multi-stage sampling does not cover all survey participants; study outcomes cannot be 100% accurate. While multi-stage sampling aims to reduce variance within and across groups (which should be reduced), it is impossible to tell if the demographics eliminated from the study were useful. Inadequate data collecting leads to lost information.

Example: In this situation, if the researcher needs 500 postgraduate students from five districts, it will be treated as cluster sampling. In the second step, from each district, the researchers need 50% male and 50% female (stratified sampling). In the final step, for 50% male and 50% female, they can select simple random sampling. This is all about multi-stage sampling technique.

CONCLUSION AND IMPLICATIONS

The purpose of this review paper is to provide clear guidelines with necessary examples for social science researchers interested in using the probability sampling approaches in their research works. The article starts with clear concepts about sample and sampling as there is confusion between these two terms. The importance of probability sampling is discussed later on. Finally, each probability sampling technique has been discussed benefits, drawbacks, and examples. It is worth noting that if any social science researcher reads the full article from beginning to end, s (he) will get clear guidelines about the probability sampling technique.

REFERENCES

1. Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it. Indian Journal of Medical Specialties, 4(2), 330-333.
2. Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. Public opinion quarterly, 69(1), 87-98.
3. Dana, L. P., & Dana, T. E. (2005). Expanding the scope of methodologies used in entrepreneurship research. International Journal of Entrepreneurship and Small Business, 2(1), 79-88.
4. Elfil, M., & Negida, A. (2017). Sampling methods in clinical research; an educational review. Emergency, 5(1), 1-5.
5. Etikan, I., & Babtope, O. (2019). A basic approach in sampling methodology and sample size calculation. Med Life Clin, 1(2), 50-54.
6. Fowler, F.J (2009). Sampling. Survey Research Methods. 4th ed. Thousand Oaks: Sage Publications; 19-47.
7. Hanlon, B., & Larget, B. (2011). Samples and populations. Department of Statistics University of Wisconsin—Madison, 14(2), 10-22.
8. Shorten, A., & Moorley, C. (2014). Selecting the sample. Evidence-based nursing, 17(2), 32-33.
9. Suresh, K., Thomas, S. V., & Suresh, G. (2011). Design, data analysis and sampling techniques for clinical research. Annals of Indian Academy of Neurology, 14(4), 287-290
10. Thomas, L. (2020). An introduction to simple random sampling. Retrieved 19 January 2022, from https://www.scribbr.com/methodology/simple-random-sampling/