

A Review of Sampling Techniques in Social Science Research

Dr. Anviti Rawat ^{1*}

¹*Associate Professor, University School Of Education, Guru Gobind Singh Indraprastha University.*

Abstract

The sampling procedure has a significant impact on a study's accuracy. Because they have the power to determine the validity of results, choosing the right sample size and selection procedure for a given research problem is essential. Examine the many studies on sampling methods in social science research that have been published in the literature. This review highlights the critical role of sampling techniques in social science research, emphasizing the trade-offs between probability and non-probability methods. While probability sampling ensures representativeness and statistical reliability, non-probability sampling offers practical advantages in terms of cost, speed, and feasibility, especially in digital research contexts like Twitter. However, limitations such as lack of generalizability and inability to compute confidence intervals must be acknowledged. Ultimately, the choice of sampling technique should align with the research objectives, available resources, and population characteristics to ensure valid, reliable, and meaningful findings in social science inquiry.

Keywords; Sampling techniques, Cluster sampling, Probability and non-probability methods, Sample size, Population size, Convenience, and snowball sampling, etc.

1 Introduction

Finding solutions to problems, gaining more knowledge about them, or devising a better approach to them are all considered forms of research. For research to accomplish its objective, it must always be planned and executed methodically. The use of sampling procedures is crucial in “quantitative research” [1]. A representative sample is obtained by selecting only a few of units or examples from a much broader

* ISBN No. 978-93-49028-02-9

set or population. This is the primary goal of sampling. From the findings of the smaller sample, the researcher can make reliable generalisations about the larger group [2]. The techniques that would provide highly similar samples are the main focus of the researchers. They make use of probability sampling, a kind of sampling that is based on mathematical ideas about the likelihood that certain occurrences would occur. The sample data allows for the drawing of conclusions about the whole population [3]. A sufficiently representative probability sample is necessary to enable the study population to make inferences. The lack of funds and time would make it difficult to conduct a comprehensive examination of the whole population in any kind of research. A sample, as used in statistics, is the portion of the population from which inferences about the population may be made [4]. For the researcher, gathering data with an appropriate sample strategy is a difficulty. The researchers won't be able to collect data from every case, thus the present techniques for answering the data gathering questions are problematic. To choose which of the many sampling techniques and processes available is best for the specific topic being examined, the researcher must be aware of the differences between them [5]. The sample size must be enough to allow for significant conclusions to be drawn. This means that the lowest sample size needed to estimate the true population percentage with the specified degree of confidence and margin of error is this one [6]. Consequently, a typical problem in statistical analysis is determining the appropriate sample size. Its equation may be found using the normal distribution's crucial parameter, population size. It serves as the foundation for all research designs and is also a cost-effective and time-saving approach [7], [8].

A. Sampling

Sampling is the intentional process of selecting a subset or individuals from a group with the intention of predicting the characteristics of the entire population and deriving statistical conclusions. Given the expense, effort, and often impossibility of studying the whole population, it offers a practical and practical way to examine its traits. Using a range of sampling methodologies, market researchers collect samples from a large population to provide relevant insights. A study's goal, the time and money available, and the research hypothesis are some of the factors that influence the optimal sampling approach [9], [10].

B. Sampling techniques

Study participants are selected using two different methods—probability sampling and non-probability sampling—to differentiate between the two. Everyone has pros and cons of its own. There is a known, nonzero possibility of selection for every member of the population when using probability sampling. This design's randomisation reduces selection bias and makes the sample representative of the whole population [11]. With probability sampling, the researcher may use the sample to make inferences about the population and predict the error of sampling with a certain level of confidence. This category includes techniques including basic random "sampling, stratified sampling, and cluster sampling". Probability sampling is precise but sometimes time-consuming and difficult since it relies on a thorough inventory of the population and choices must be made via a complicated process, especially when the target population is vast or scattered [12].

C. Probability sampling methods

The sampling technique in question is characterised by an equal chance of selecting each unit or part of the population for the final sample. Random sampling emphasises the element of sample selection that is random and has a non-zero probability. By ensuring a more impartial and representative sample, this sampling method permits reliable conclusions about the whole population [9].

Simple random sampling: In spite of its simplicity, the random sample ensures that each member of the population has a fair opportunity of being selected. Following the definition and creation of a comprehensive list of population members, the sample size is determined, and then the real random selection process—either by drawing lots or using a random number generator—begins. Though logistically challenging, particularly when the population is large or scattered, this is highly prized since it lessens selection bias.

Stratified sampling: Using consistent criteria like age, sex, or education, stratified sampling would separate the population into discrete groups, or strata. From each stratum, random samples would then be selected to provide a sample representative of the entire spectrum of categories. Given that each subgroup is well-represented, this will improve the estimates' accuracy and make the technique more appropriate for populations with high levels of variability in important features. But without a thorough understanding of the population, stratification is hard to establish in a meaningful way.

Cluster sampling: Cluster sampling is a method that simplifies the data collection process by performing a random selection of individuals for analysis from a sample that has been divided into clusters, which may represent institutions or geographic regions. The sample will then include each member of the selected clusters. The compilation of a comprehensive inventory of the population would be nearly impossible in large-scale research. Consequently, this will be extremely beneficial. The cost of cluster sampling is low; however, biases may be introduced if the clusters selected are not representative of the entire population.

Systematic sampling: The systematic sampling method involves the continuous selection of each Kth person from a population list, beginning at a random point. The method is both straightforward and effective when applied over a consistent sample period. However, systematic sampling may be biased if the population exhibits a natural pattern that corresponds to the sample interval.

Multi-stage sampling: In multi-stage sampling, a population is divided into numerous stages or levels. Cluster sampling, for example, is a sophisticated form of multi-stage sampling that involves the division of populations into substantial clusters (such as areas or organisations), from which additional random samples are obtained in a succession of stages.

For example, a random sample of individuals from each area may be selected when geographical areas are selected. By concentrating on broad or dispersed populations, stepwise stratification aids in gradually reducing the sample size in large-scale studies. Even while this might save money and time, if the sample is not representative at every step of the selection process, there is a chance that sampling error will rise.

D. Non-probability sampling methods

Another sampling method is non-probability sampling, which essentially involves drawing data by choosing individuals non-randomly based on preset criteria. Although the final sample may not be a perfect representation of the total population, this provides a simple method of collecting data [9].

Convenience sampling: Among the several data collecting techniques, convenience sampling entails selecting samples from the population segment that the researcher finds most accessible. When resources or time are scarce, this is an expeditious and cost-effective method of acquiring data. Although the sample may not be representative, this approach often increases bias since it makes it impossible to extrapolate the results to the whole population.

Purposive sampling: The researcher uses purposive sampling, often referred to as judgemental or expert sampling, to choose participants based on their potential value in providing the needed data. Research that calls for specific individuals with certain traits or specialities often uses this. Although the approach may provide very valuable insights, it bears a significant risk of researcher bias and may not be representative of the whole community.

Snowball sampling: When examining populations that are hidden or difficult to reach, snowball sampling is used. In this instance, participants must create a network of referrals by recommending others who meet the requirements for research participation. In exploratory research, this method is helpful. But since social networks drive samples to become homogenous over time, it has been argued that this results in biased samples.

Quota sampling: In order to satisfy specific quotas within each segment, quota sampling is a method of selecting participants from a group that has been divided into mutually exclusive subgroups. The sample is biased and less representative due to the fact that selection within the subgroup is not random, despite the fact that quotas ensure that the sample has specific demographic features.

2 Literature Review

(Ahmed, 2024) [13] Thoroughly revised the rules governing sample size computation and sampling techniques, providing sufficient proof to help researchers boost the validity and statistical strength of their work. In exploratory settings, non-probability sampling is helpful, but only probability can guarantee generalisability. A further critical procedure is the determination of the optimal sample size, which necessitates taking into account of several factors: the confidence level, the margin of error, the effect size, the size of the entire population, and the statistical power. The study provides researchers with the theoretical direction and useful tools they need to choose suitable sampling procedures and validate methodologically sound sample size estimations.

(Ali & Hatef, 2024) [14] It is the objective of this methodological review to underscore the importance of sampling and sample size estimation methodologies and to offer suggestions for the precise definition of sample sizes. The sample size for research on health is determined using methods that are elucidated with examples, depending on the type of study design. Researchers need to determine acceptable

accuracy levels, study power, the confidence level to be utilised, and the statistical analysis that will be used in order to estimate sample size. This method review provides a comprehensive and straightforward explanation of "the sample size formula and sampling technique", which will assist in the demystification of the complexities associated with statistical equations for sample size estimates in social health care and science research.

(Adeoye, 2023) [15] Provides a comprehensive outline of the steps involved in producing articles for publication in a variety of academic subjects, including data collection methods and the use of sample methodologies. Probability and non-probability sampling are two types of sampling procedures used in statistics. One advantage of probability sampling is that it guarantees the sample's representativeness of the body. Among the several types of probability sampling techniques are "stratified, cluster, systematic, and basic random sampling". In non-probability sampling, the sample is selected using personal judgement. Since non-probability sampling incorporates the representation of every topic in the population, it is seen to be the optimum method. Numerous non-probability sampling techniques exist, such as convenience sampling, quota sampling, judgemental sampling, snowball sampling, and sequential sampling.

(Makwana et al., 2023) [16] A study's sampling procedure has a big impact on its accuracy. The many sampling techniques used in research are summarised in the article. Probability sampling techniques and non-probability sampling techniques are the two basic categories into which these methods may be divided. Simple, systematic, and stratified random sampling are examples of probability sampling procedures. Non-probability sampling methods, on the other hand, include purposive, convenience, snowball, self-selection, and quota sampling.

(Mishra & Lavater, 2023) [7] Sampling is a vital part of every research project. Choosing the right sample size and selection method for your specific research problem is essential since they may determine the reliability of your results. In order to get reliable and accurate data, this article will go over the optimal sample size, some of the most popular sampling methods, and how they could be used in practice. This study looks at the fundamentals of probability sampling as well as the benefits and drawbacks of non-probability sampling and various probability sampling methods. Social science scholars will find this study useful in choosing the most suitable sampling method.

(Vicente, 2023) [17] Since Twitter has a vast user base and vast amounts of data, sampling is usually necessary when utilising it for study. Despite its importance for legitimate and trustworthy research results, there is currently a lack of expertise about how to choose representative samples of Twitterverse users. This report presents a thorough quantitative literature assessment of sample strategies developed and used in social science research on Twitter. The review covers the following topics: (1) target population definition; (2) sampling frames used to support sample selection; (3) sampling techniques used to get Twitter user samples; (4) gathering of information from Twitter users; (5) "sample size; and (6) research validity". This review may be used as methodological advice for scholars and professionals who want to do social science research with Twitter users and the Twitterverse.

(Zickar & Keith, 2023) [18] Examine various sampling strategies, such as snowball, probability-based, convenience, and purposeful sampling. To assist organisational researchers in selecting the best sample strategies for their research issues, we point out the advantages and disadvantages of each strategy. In order to enhance the quality of online sampling, we analyse screening methods and highlight recommended practices that researchers may apply to improve the quality of their samples. In order to draw conclusions about the lack of methodological and sample diversity in organisational research, the over-reliance on a small number of sampling techniques, the necessity of disclosing important aspects of sampling, and participant quality concerns, we lastly reviewed the sampling practices of all empirical research articles published in the Journal of Applied Psychology during the previous five years.

(Raifman et al., 2022) [19] Gave instructions on how to use one of the most often used strategies, respondent-driven sampling (RDS), and a summary of sampling techniques for groups that are difficult to reach. Estimates derived from convenience or non-probability-based sample data may be skewed or not applicable to the intended audience. In an attempt to produce representative samples, inclusion-influencing variables may be evaluated and taken into consideration in RDS and time-location sampling (TLS). RDS is especially well-suited to connect with the most obscure members of communities that are difficult to reach. In addition to more generalisable estimates of population characteristics, TLS, RDS, or a combination of the two may provide a rigorous approach to find and enrol samples from communities that are difficult to reach. These techniques should be added to the toolkits of researchers that are interested in sampling populations that are difficult to reach.

(Rahman et al., 2022) [20] It might be difficult for a researcher to get data using a suitable sampling strategy. The study's research questions cannot be addressed in their present form as the researchers will not be able to gather data from every scenario. To choose the best sampling technique or method for the particular study being considered, the researcher has to be aware of the distinctions between the vast array of possible sampling techniques and procedures. In this regard, the research also examines the fundamental ideas of probability sampling, including the many types of probability sampling methods and their benefits and drawbacks. This study will help social science researchers choose the best probability sampling approach or techniques to use in order to conduct their research efficiently and effectively.

(Cornesse et al., 2020) [21] There is ongoing debate in the survey research literature on the reliability of estimates of a larger population using "probability and nonprobability sample surveys". As a consequence of survey design, statistical theory supports confidence in probability sampling, but findings from nonprobability sampling are entirely dependent on models for validity. This article reviews the controversy surrounding "probability and nonprobability sample surveys". Theoretically, nonprobability sample surveys may provide reliable results, and we discuss empirical evidence on which sample types offer the most accuracy in reality. We derive best-practice suggestions and suggest directions for future study from these theoretical and empirical concerns.

3 Conclusion

In conclusion, this review underscores the critical role of sampling techniques in shaping the validity and reliability of social science research. Even though statistical rigour and representativeness are provided by probability sampling techniques such as "simple random, stratified, cluster, and systematic sampling", they often need substantial resources and population access. In contrast, non-probability methods, including convenience and snowball sampling, provide pragmatic alternatives, especially in exploratory studies or when dealing with platforms like Twitter. However, these methods come with limitations such as the inability to calculate confidence intervals or margins of error. Despite these drawbacks, the rising cost and complexity of data collection have led many researchers to adopt non-probability approaches. Importantly, researchers must carefully align their sampling strategy with study objectives, population characteristics, and resource constraints. By acknowledging the strengths and limitations of each method, social science researchers can make informed methodological choices, enhancing the credibility and interpretability of their findings.

References

- [1] P. Cash, O. Isaksson, A. Maier, and J. Summers, "Sampling in design research: Eight key considerations," *Des. Stud.*, vol. 78, p. 101077, 2022, doi: 10.1016/j.destud.2021.101077.
- [2] A. E. Berndt, "Sampling Methods," *J. Hum. Lact*, vol. 36, no. 2, pp. 224–226, 2020, doi: 10.1177/0890334420906850.
- [3] M. Elfil and A. Negida, "Sampling methods in clinical research; an educational review," *Arch. Acad. Emerg. Med.*, vol. 7, no. 1, pp. 3–5, 2017.
- [4] M. R. W. Hiebl, "Sample Selection in Systematic Literature Reviews of Management Research," *Organ. Res. Methods*, vol. 26, no. 2, pp. 229–261, 2023, doi: 10.1177/1094428120986851.
- [5] S. Noor and J. Golzar, "Simple random sampling," *Handb. Stat.*, vol. 6, no. November, pp. 97–109, 2022, doi: 10.1016/S0169-7161(88)06006-7.
- [6] M. Memon, H. Ting, J.-H. Cheah, R. Thurasamy, F. Chuah, and T. Huei Cham, "SAMPLE SIZE FOR SURVEY RESEARCH: REVIEW AND RECOMMENDATIONS," *J. Appl. Struct. Equ. Model*, vol. 4, no. 2, pp. 2590–4221, 2020.
- [7] S. K. Mishra and M. Lavater, "Sampling Methods in Social Science Research," no. July, 2023, [Online]. Available: <https://www.researchgate.net/publication/371987239>
- [8] N. Saxena, "A systematic Survey of Natural Language Processing (NLP) Models & its Application," *Int. J. Innov. Sci. Eng. Manag.*, pp. 25–31, 2022.
- [9] S. J. Stratton, "Population Sampling: Probability and Non-Probability Techniques," *Prehosp. Disaster Med.*, vol. 38, no. 2, pp. 147–148, 2023, doi: 10.1017/S1049023X23000304.
- [10] S. Dehariya and D. Sharma, "Study On Highway Roadside Safety Prabhat Chauraha to Bajrang Chauraha Raisen Road Bhopal," *Int. J. Innov. Sci. Eng. Manag.*, pp. 44–50, 2024.

- [11] Hamed Taherdoost, "Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research," *Int. J. Acad. Res. Manag.*, vol. 5, no. 2, pp. 18–27, 2016.
- [12] M. Haque, "Sampling Methods in Social Research," *Glob. Res. Methodol.*, pp. 1–6, [Online]. Available: http://www.pansoi.com/grmgrlaranya.org/Journals/SAMPLING_METHODS_IN_SOCIAL_RESEARCH.pdf
- [13] S. K. Ahmed, "How to choose a sampling technique and determine sample size for research: A simplified guide for researchers," *Oral Oncol. Reports*, vol. 12, no. September, p. 100662, 2024, doi: 10.1016/j.oor.2024.100662.
- [14] M. D. Ali and E. A. J. Al Hatef, "Types of Sampling and Sample Size Determination in Health and Social Science Research," *J. Young Pharm.*, vol. 16, no. 2, pp. 204–215, 2024, doi: 10.5530/jyp.2024.16.27.
- [15] M. A. Adeoye, "Review of sampling techniques for education," *ASEAN J. Sci. Educ.*, vol. 2, no. 2, pp. 87–94, 2023, [Online]. Available: <https://ejournal.bumipublikasinusantara.id/index.php/ajsed>
- [16] D. Makwana, P. Engineer, A. Dabhi, and H. Chudasama, "Sampling methods in research: A review," *Int. J. Trend Sci. Res. Dev.*, vol. 7, no. 3, pp. 762–768, 2023, [Online]. Available: <https://www.researchgate.net/publication/371985656>
- [17] P. Vicente, *Sampling Twitter users for social science research: evidence from a systematic review of the literature*, vol. 57, no. 6. Springer Netherlands, 2023. doi: 10.1007/s11135-023-01615-w.
- [18] M. J. Zickar and M. G. Keith, "Innovations in Sampling: Improving the Appropriateness and Quality of Samples in Organizational Research," *Annu. Rev. Organ. Psychol. Organ. Behav.*, vol. 10, pp. 315–337, 2023, doi: 10.1146/annurev-orgpsych-120920-052946.
- [19] S. Raifman, M. A. DeVost, J. C. Digitale, Y.-H. Chen, and M. D. Morris, "Respondent-Driven Sampling: a Sampling Method for Hard-to-Reach Populations and Beyond," *Curr. Epidemiol. Reports*, vol. 9, no. 1, pp. 38–47, 2022, doi: 10.1007/s40471-022-00287-8.
- [20] M. M. Rahman, M. I. Tabash, A. Salamzadeh, S. Abduli, and M. S. Rahaman, "Sampling Techniques (Probability) for Quantitative Social Science Researchers: A Conceptual Guidelines with Examples," *SEEU Rev.*, vol. 17, no. 1, pp. 42–51, 2022, doi: 10.2478/seeur-2022-0023.
- [21] C. Cornesse et al., "A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research," *J. Surv. Stat. Methodol.*, vol. 8, no. 1, pp. 4–36, 2020, doi: 10.1093/jssam/smz041.
- [22] M. Medhi, "a Study on the Green Marketing Practices Adopted By Various Companies in India," vol. 6, no. 3, pp. 83–88, 2015, [Online]. Available: <http://www.iaeme.com/ijmhrm.asp83http://http://www.iaeme.com/currentissue.asp?JType=IJMHRM&VType=6&IType=3http://www.iaeme.com/currentissue.asp?JType=IJMHRM&VType=6&IType=3>