Poonam Shodh Rachna (ISSN 2456-5563)
(A multidisciplinary, peer reviewed and refereed Research Journal)
Vol. 1, Issue.V, July 2022, pc – PSR-2207013

The Art of Choosing a Research Sample

Authors : Dr. Sumer Singh, Singhania University, Rajasthan
Dr.Uma Shankar Yadav, Singhania University, Rajasthan
Dr. Neeraj Kumar Sharma, Singhania University, Rajasthan
Dr. Shatrughan Bhadwaj, MGNCRE, Ministry of Education, Hyderabad

Abstract
The percentage of a target population that is included in a study is referred to by researchers as the "sample size," and they use this term to indicate the proportion of the population that is studied. The total number of people who took part in a study is referred to as the "sample size." After the study is completed, the total number of participants is frequently segmented into subgroups samples are stratified by demographic variables like age, sex, and location to ensure that they are representative of the population under study. Choosing the appropriate size of the sample to analyze Indeed, is a crucial part of every statistical study.

Key words : Research Sample, Population, Sample Size, Demographic variable

Introduction

The percentage of a target population that is included in a study is referred to by researchers as the "sample size," and they use this term to indicate the proportion of the population that is studied. The total number of people who took part in a study is referred to as the "sample size." After the study is completed, the total number of participants is frequently segmented into subgroups samples are stratified by demographic variables like age, sex, and location to ensure that they are representative of the population under study. Choosing the appropriate size of the sample to analyze Indeed, is a crucial part of every statistical study. If the sample is too small, its findings cannot be considered reliable. In the event that there are insufficient individuals included in the sample, the results will not be valid or accurate. However, despite the fact that a larger sample size leads to lower error margins and a sample that is more representative of the population as a whole, Additionally, it may cause a dramatic rise in both cost and time requirements to carry out the research if the sample size is too large.

When determining the appropriate size of your sample, there are a few aspects you should keep in mind as well as a few different approaches you might use.

The Extent of Our Confidence and the Interval of Our Confidence

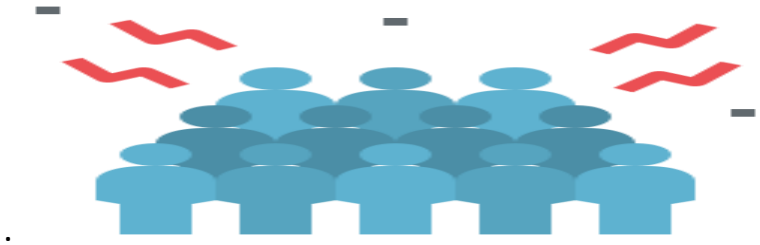
When selecting a sample, it is important to keep in mind that sampling and non-sampling errors are not the only possible causes of bias in research; other factors, such as those listed

above, must also be taken into consideration. In statistics, sample sizes are typically used to determine error metrics like the standard error and the confidence level. The confidence level is the more frequent of the two.

Confidence Interval, or CI (Margin of Error)

Confidence intervals are used to quantify the range of possible values for a particular statistic, as well as the degree to which one may be confident in the results of a particular sampling strategy. This is accomplished by dividing the range of possible values by a number known as the standard deviation. The level of certainty with which one can conclude that The study results are representative of the kinds of things one might find in the field if a representative sample of the population under investigation were surveyed is what is meant to be conveyed by the term "confidence interval," which translates literally to "interval of certainty." Many times, confidence intervals are written with a plus or minus sign after them. If 60% of the people in your sample select an answer and your confidence interval is 6, then you can be certain that between 54% and 66% of the population would have selected that answer if you had questioned the entire population. This range is calculated by subtracting 60 from 60 and adding 60 to 60.

How self-assured are you, Really? The confidence level is the number of times (or degree of certainty) that a sample taken at random would contain the actual population parameter within the confidence interval. This is measured in terms of the probability that the parameter would be included. Confidence intervals are numerical ranges that represent the chance that a specific proportion of the population will pick a certain response. These ranges are often expressed as percentages. As an example, if you have a confidence level of 99% in your findings, you may do the same experiment or survey several times with a good chance that the outcomes will provide results that are comparable to those acquired from the population as a whole



Population

The dependability of the findings is improved as the sample size is increased since this enhances the possibility that the answers obtained correctly reflect the population as a whole. To restate this idea in a different way, a larger sample size correlates to a smaller confidence interval for any given degree of certainty.

The number of standard deviations from the mean.

When determining the appropriate size of a sample, it is essential to take into account the standard deviation, which is a measurement of the dispersion of a data set in relation to its mean. When determining the appropriate size of a sample from a population, the standard deviation is a useful tool for providing an estimate of the standard deviation of the population. It is feasible to utilise it to make predictions about the variety of potential replies and the dispersion of those answers around the mean.

The mean and the standard deviation will both be high when there is a lot of variability or dispersion in the data. After you have sent your survey, for example, do you expect a significant amount of variety in the responses, if any at all? The standard deviation is a statistical metric that determines how much different solutions might differ from one another.

When determining an appropriate sample size, The size of the intended audience is an essential factor to think about. The whole group from which generalisations are drawn is referred to as the "population," and the word "population" is used to refer to this group. There are two types of samples that may be obtained from a population: a probability sample and a non-probability sample. Even when working with relatively small or easily measurable samples of a population, it is necessary to have an accurate estimate of the population size even if the size of population is known.

In research, sample sizes of 400 and 500 are often used since these numbers enable assumptions to be derived about any size population with 95% confidence and a 5% error rate, as shown in the computation that follows below. You will, however, require a larger sample size if you wish to compare subgroups within a larger group, such as provinces within a country. This is because comparing subgroups inside a larger group is more difficult. For the majority of research projects, GeoPoll recommends a sample size of at least 400 respondents per country as a minimum viable sample, 800 respondents per country for an investigation that includes 1200 responses or more per nation for third-level breakdowns, such as men 18-24 in Nairobi.

Methods for Figuring Out the Right Amount of Subjects to Sample

Since we are all on the same page with regard to the fundamental ideas, we can move on to a concise discussion of the Andrew Fisher Formula for computing the sizes of the samples.

- Make a tally of the individuals (if known).
- Determine the degree of confidence that corresponds to 95%.
- The third step is to determine the level of certainty.
- Standard deviation value (a SD of 0.5 is a safe choice).
- Determine a Z-score for the data depending on the amount of confidence.

The following degrees of confidence are represented by their respective z-scores in the table of data that follows:

Confidence Level	80%	85%	90%	95%	99%
Z- Score	1.28	1.44	1.65	1.96	2.58

Applying the sample size computation to the data you have just acquired will allow you to calculate the size of your sample.

$$\text{Sample Size} = \text{Z-Score}^2 \times \text{StdDev X (1-StdDev)} / (\text{Confidence Interval})^2$$

The following is a representation of an example computation:

To get a 95% level of certainty in your study, you'll need a standard deviation of 0.5 and an error margin of 5%, all you need to do is put those figures into the following formula:

$$\text{Sample Size} = \text{Z-Score}^2 \times \text{StdDev X (1-StdDev)} / (\text{Confidence Interval})^2$$

$$((1.96)^2 \times .5(.5)) / (.05)^2 = (3.8416 \times .25) / .0025 = .9604 / .0025 = 384.16$$

385 is the best number for the size of your sample.

There are a number of useful calculators that can be located on the internet and utilized for this specific purpose. The Easy Calculation demo calculator is a cost-free tool that you may start using right now. After you tell the software your target confidence level, population size, and 95 % confidence level, it will compute the appropriate sample size for you.

References :

1. Cochran, W. G. 1963. Sampling Techniques, 2nd Ed., New York: John Wiley and Sons, Inc.
2. Sharma, N. K. (2022, May 15). *How to Write an Article/Research Paper of Social Science for Publication in an Indexed Journal*. How to Write an Article/Research Paper of Social Science for Publication in an Indexed Journal. http://dx.doi.org/10.13140/RG.2.2.27844.71049
3. Israel, Glenn D. 1992. Sampling The Evidence Of Extension Program Impact. Program Evaluation and Organizational Development, IFAS, University of Florida. PEOD-5. October.
4. Kish, Leslie. 1965. Survey Sampling. New York: John Wiley and Sons, Inc.
5. Miaoulis, George, and R. D. Michener. 1976. An Introduction to Sampling. Dubuque, Iowa: Kendall/Hunt Publishing Company.
6. Sharma, N. K. (2018, February 28). *Corporate Governance and Its Relation to Business* . Corporate Governance and Its Relation to Business . http://dx.doi.org/10.13140/RG.2.2.16541.74729
7. Smith, M. F. 1983. Sampling Considerations In Evaluating Cooperative Extension Programs. Florida Cooperative Extension Service Bulletin PE-1. Institute of Food and Agricultural Sciences. University of Florida.
8. Sharma, N. K. (2015, November 4). *Industry Initiatives for Green Marketing in India*. Industry Initiatives for Green Marketing in India. http://dx.doi.org/10.4172/2151-6219.1000192
9. Sudman, Seymour. 1976. Applied Sampling. New York: Academic Press.
10. Sharma, N. K. (2016, February 28). *Corporate Social Responsibility Is Not a Charity but a Responsibility in India*. Corporate Social Responsibility Is Not a Charity but a Responsibility in India. http://dx.doi.org/10.13140/RG.2.2.22472.75520
11. Yamane, Taro. 1967. Statistics, An Introductory Analysis, 2nd Ed., New York: Harper and Row
12. Morse, J. M. (1991). Strategies for sampling. In J. M. Morse (Ed.), Qualitative nursing research: A contemporary dialogue (pp. 127-145). Newbury Park, CA: Sage.
13. Morse, J. M. (1994). Designing qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative inquiry (pp. 220-235). Thousand Oaks, CA: Sage.
14. Sharma, N. K. (2022, May 15). *How to Write an Article/Research Paper of Social Science for Publication in an Indexed Journal*. How to Write an Article/Research Paper of Social Science for Publication in an Indexed Journal. http://dx.doi.org/10.13140/RG.2.2.27844.71049
15. Ramesh, R., Shukla, A. K., & Sharma, N. K. (2017, May 31). *Corporate Social Responsibility in Our Changing Business World*. Corporate Social Responsibility in Our Changing Business World. http://dx.doi.org/10.13140/RG.2.2.30674.58562
16. Pandey, R. N., & Sharma, N. K. (2018, February 28). *Management of Stress Life* . Management of Stress Life . http://dx.doi.org/10.13140/RG.2.2.20795.03361
17. Shukla, A. K., Ramesh, R., & Sharma, N. K. (2018, February 18). *An Overview of Corporate Social Responsibility in India*. An Overview of Corporate Social Responsibility in India. http://dx.doi.org/10.13140/RG.2.2.21633.89446
18. Sharma, N. K. (2022, May 31). *Instruments Used in the Collection of Data in Research*. Instruments Used in the Collection of Data in Research. http://dx.doi.org/10.2139/ssrn.4138751
19. Rachna, S. R., & Sharma, N. K. (2022, July 31). *How Garbage Dumps affect Urban Environment : A Case Study of Prayagraj District*. How Garbage Dumps Affect Urban Environment : A Case Study of Prayagraj District. http://dx.doi.org/10.13140/RG.2.2.23364.09603
20. Kumar , P., & Sharma, N. K. (2022, April 30). *NGO Impact On India's Development Process*. NGO Impact On India's Development Process. http://dx.doi.org/10.13140/RG.2.2.31972.24963
21. Yadav, G. P., & Sharma, N. K. (2022, March 31). *Marketing in India is adapting to shifting consumer attitudes and behaviours*. Marketing in India Is Adapting to Shifting Consumer Attitudes and Behaviours. http://dx.doi.org/10.13140/RG.2.2.24422.50241
22. Sharma, N. K. (2021, December 31). *Easy Way to Determine the Sample Size*. Easy Way to Determine the Sample Size. http://dx.doi.org/10.13140/RG.2.2.35758.84808
23. Bellhouse, David. 2000. ''Survey Sampling Theory over the Twentieth Century and Its Relation to Computing Technology.'' Survey Methodology 26:11–20.
24. Blumberg, Stephen, and Julian Luke. 2010. Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, July–December 2009. Hyattsville, MD: U.S. Centers for Disease Control and Prevention, National Center for Health Statistics. http://www.cdc.gov/nchs/ data/nhis/earlyrelease/wireless201005.pdf.
25. Sharma, N. K. (2016, February 28). *Penetration Of E-Commerce And Its Acceptance : An Exploratory Study Of Sme's In India*. Penetration Of E-Commerce And Its Acceptance : An Exploratory Study Of Sme's In India. http://dx.doi.org/10.13140/RG.2.2.24150.47689
26. Bowley, Arthur. 1926. ''Measurement of the Precision Attained in Sampling.'' Bulletin of the International Statistical Institute 22: Supplement to Liv. 1:1–62.

27. Brick, J. Michael. 2010. ''Allocation in Dual Frame Telephone Surveys with Nonsampling Errors.'' Paper presented at the Statistical Society of Canada Annual Meeting, Quebec City, Canada. Couper, Mick. 2000. ''Web Surveys: A Review of Issues and Approaches.'' Public Opinion Quarterly 64:464–94.
28. Deville, Jean-Claude, and Carl-Erik Sa¨rndal. 1992. ''Calibration Estimators in Survey Sampling.'' Journal of the American Statistical Association 87:376–82.
29. Sharma, N. K. (2022, March 31). *Post-Pandemic Human Resource Management: Challenges and Opportunities*. Post-Pandemic Human Resource Management: Challenges and Opportunities. http://dx.doi.org/10.13140/RG.2.2.31311.56484
30. Dillman, Don. 1978. Mail and Telephone Surveys: The Total Design Method. New York: Wiley & Sons. Duncan, Joseph, and William Shelton. 1978. Revolution in United States Government Statistics, 1926–1976. Washington, DC: U.S. Department of Commerce, Office of Federal Statistical Policy and Standards. Frankel, Martin, and Lester Frankel. 1987. ''Fifty Years of Survey Sampling in the United States.'' Public Opinion Quarterly 51(Part 2):S127–38.
31. Godambe, Vidyadhar. 1966. ''A New Approach to Sampling from Finite Populations. II. DistributionFree Sufficiency.'' Journal of the Royal Statistical Society Series B (Methodological) 28:320–28. Green, Donald, and Alan Gerber. 2006. ''Can Registration-Based Sampling Improve the Accuracy of Midterm Election Forecasts?'' Public Opinion Quarterly 70:197–223.
32. Sharma, N. K. (2019, March 31). *CSR Expenditure of BSE Listed Companies in India: An Analytical Study* . CSR Expenditure of BSE Listed Companies in India: An Analytical Study . http://dx.doi.org/10.13140/RG.2.2.23626.18882
33. Groves,Robert,andMickCouper.1998.NonresponseinHouseholdInterviewSurveys.NewYork:Wiley.
34. Groves, Robert, and Steven Heeringa. 2006. ''Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs.'' Journal of the Royal Statistical Society Series A 169:439–57.
35. Groves, Robert, Stanley Presser, and Sarah Dipko. 2004. ''The Role of Topic Interest in Survey Participation Decisions.'' Public Opinion Quarterly 68:2–31.
36. Sharma, N. K. (2020, August 21). *An Analysis of Corporate Social Responsibility in India*. An Analysis of Corporate Social Responsibility in India. http://dx.doi.org/10.2139/ssrn.3676827
37. Hansen, Morris, William Madow, and Benjamin Tepping. 1983. ''An Evaluation of Model Dependent and Probability-Sampling Inferences in Sample Surveys.'' Journal of the American Statistical Association 78:776–93.
38. Heckathorn, Douglas. 1997. ''Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations.'' Social Problems 44:174–99.
39. Holt, D. Tim, and Smith. 1979. ''Post-Stratification.'' Journal of the Royal Statistical Society Series A 142:33–46.
40. Jabine, Thomas, and Fritz Scheuren. 1985. ''Goals for Statistical Uses of Administrative Records: The Next 10 Years.'' Journal of Business & Economic Statistics 3:380–91.
41. Sharma, N. K. (2015, October 31). *Emergence of SNS as Marketing Communication Tool* . Emergence of SNS as Marketing Communication Tool . http://dx.doi.org/10.13140/RG.2.2.32958.51526
42. Kaier, Anders. 1895–96.''Observations et experiences concernant des de´nombrements representatives.'' Bulletin of the International Statistical Institute 9:176–83.
43. Kalton, Graham, and J. Michael Brick. 1995. ''Weighting Schemes for Household Surveys.'' Survey Methodology 21(1):33–44.
44. Keeter, Scott. 2006. ''The Impact of Cell Phone Noncoverage Bias on Polling in the 2004 Presidential Election.'' Public Opinion Quarterly 70:88–98.
45. Kruskal, William, and Frederick Mosteller. 1980. ''Representative Sampling IV: The History of the Concept in Statistics, 1895–1939.'' International Statistical Review/Revue Internationale de Statistique 48:169–95.
46. Lavallec´e, Pierre. 2007. Indirect Sampling. New York: Springer. Lee, Sunghee, and Richard Valliant. 2009. ''Estimation for Volunteer Panel Web Surveys Using Propensity Score Adjustment and Calibration Adjustment.'' Sociological Methods & Research 37:319–43.