

REVIEW ARTICLE

Sample Size Determination in Applied Research: A Critical Review of Cochran's, Yamane's and Power-Analysis Approaches

Sarika B. Patil*

Research Scholar, Punyashlok Ahilyadevi Holkar Solapur University, Solapur, Maharashtra, India

* Corresponding author

OPEN ACCESS

This article is licensed under CC BY-NC 4.0.

**ARTICLE HISTORY**

Received: 21 Sep 2026

Accepted: 06 Oct 2026

Published: 14 Oct 2026

CORRESPONDING AUTHORMs. Sarika B. Patil, Punyashlok
Ahilyadevi Holkar Solapur University,
Solapur, Maharashtra, India.**HOW TO CITE**Patil, S. B. (2026). Sample Size
Determination in Applied Research: A
Critical Review of Cochran's, Yamane's
and Power-Analysis Approaches. *Global
Journal of Advanced Science and
Technology*, 1(1), 22–26.

© 2026 The Author(s).

ABSTRACT

Few questions are asked more often by research students than how large a sample should be, and few are answered more mechanically. Many theses and articles in India justify their sample size with a single formula, most often Yamane's or the Krejcie and Morgan table, without explaining whether the formula suits the study's purpose. This review examines the three approaches most commonly used in applied research: Cochran's formula for estimating a proportion, Yamane's simplified formula, and statistical power analysis for testing hypotheses. It sets out the assumptions behind each, shows with worked calculations how they compare across population sizes and effect sizes, and discusses common errors. The comparison shows that Cochran's formula and the Krejcie and Morgan table give practically identical results, that Yamane's formula is equivalent to Cochran's at a confidence level of about 95.5 per cent and therefore gives slightly larger samples, and that none of these formulas is appropriate when the aim is to compare groups or test a relationship, where power analysis is needed. A study comparing two groups for a small effect needs nearly 400 participants per group at 80 per cent power, far more than survey formulas suggest. The review recommends that researchers state the purpose of the sample, the assumptions used and adjustments for non-response and design effects, and that they justify sample size in the way the analysis will actually be carried out.

Keywords: Sample size; Cochran's formula; Yamane's formula; Power analysis; Research methodology; Effect size

1. Introduction

At some point in every research project, usually when the questionnaire is ready, the researcher must decide how many people to survey or how many units to observe. Too small a sample gives estimates so uncertain that they mean little, or tests that cannot detect real effects. Too large a sample wastes time, money and the goodwill of respondents. Choosing the size well is therefore an ethical as well as a statistical matter.

In practice, the choice is often made by habit. A formula is copied from an earlier thesis, a population size is entered, and the result is reported as the sample size, sometimes rounded up for safety. Whether the formula was designed for the kind of question the study asks is rarely discussed. Button et al. (2013) estimated that the median statistical power of studies in neuroscience was only about 21 per cent, a reminder that inadequate samples are a problem across disciplines, not only in student research.

This review examines the three approaches that applied researchers in India use most: Cochran's formula, Yamane's formula and power analysis. It explains what each is designed to do, compares their results with worked examples, and offers practical recommendations.

2. Two different purposes

The most important distinction is between two purposes of a study. Some studies aim to estimate a quantity, such as the proportion of farmers who use crop insurance or the average household income in a ward. Others aim to test a hypothesis, such as whether a training programme improves test scores or whether income differs between two groups. The first purpose calls for a sample large enough to give an estimate within a chosen margin of error. The second calls for a sample large enough to detect an effect of a given size with a chosen probability. Cochran's and Yamane's formulas serve the first purpose; power analysis serves the second (Lenth, 2001; Lakens, 2022).

3. Cochran's formula

For estimating a proportion p with margin of error e at a confidence level corresponding to the standard normal value z , Cochran (1977) gives the sample size for a large population as $n_0 = z^2 p(1 - p) / e^2$. When the proportion is unknown, $p = 0.5$ is used because it gives the largest sample. At 95 per cent confidence ($z = 1.96$) and a 5 per cent margin of error this gives $n_0 = 384.16$, which is rounded up to 385. For a finite population of size N , the finite population correction gives $n = n_0 / [1 + (n_0 - 1) / N]$.

The table published by Krejcie and Morgan (1970), widely reproduced in Indian methodology textbooks, is based on essentially the same calculation, using the chi-square value 3.841 (which equals 1.96^2) with $p = 0.5$ and $e = 0.05$. The two give the same sample sizes, as Table 1 shows. Bartlett et al. (2001) extended the approach to continuous variables, for which the formula uses the expected standard deviation of the variable in place of $p(1 - p)$.

4. Yamane's formula

Yamane (1967) gave a simpler expression, $n = N / (1 + Ne^2)$, which needs only the population size and the margin of error. Its popularity comes from this simplicity. It is not a separate method, however. Substituting $p = 0.5$ and $z = 2$ into Cochran's formula with the finite population correction gives almost exactly Yamane's result, so Yamane's formula assumes maximum variability and a confidence level of about 95.5 per cent (Israel, 1992). This is why it gives slightly larger samples than Cochran's formula at 95 per cent confidence, and why its results approach 400 rather than 385 for large populations.

Table 1. Required sample size by population size (95% confidence for Cochran and Krejcie–Morgan; $e = 0.05$; $p = 0.5$).

Population (N)	Cochran (with correction)	Yamane	Krejcie and Morgan
100	80	80	80
500	218	223	218
1,000	278	286	278
5,000	357	371	357
10,000	370	385	370
1,00,000	383	399	383
10,00,000	385	400	384

Source: calculated by the author using the formulas in Sections 3 and 4; values rounded up to the next whole number.

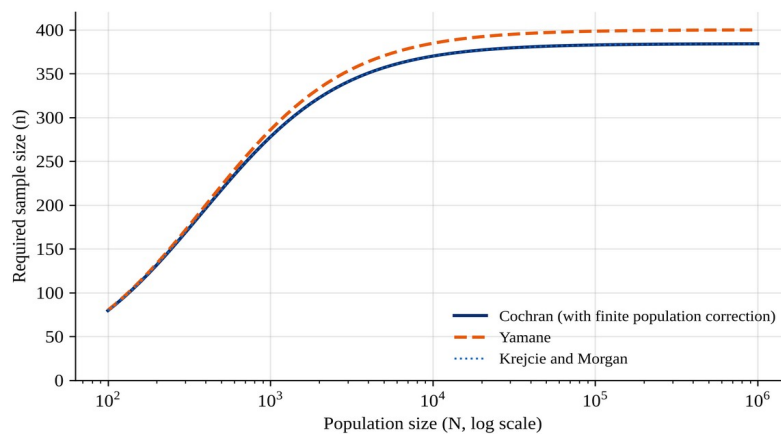
**Figure 1.** Required sample size as population size increases, for the three formulas.

Figure 1 brings out a point that is often misunderstood. Beyond a population of about 20,000, the required sample barely changes. A survey of a district with ten lakh households needs hardly more respondents than one of a town with twenty thousand. The precision of an estimate depends mainly on the size of the sample, not on the fraction of the population it represents.

5. Power analysis

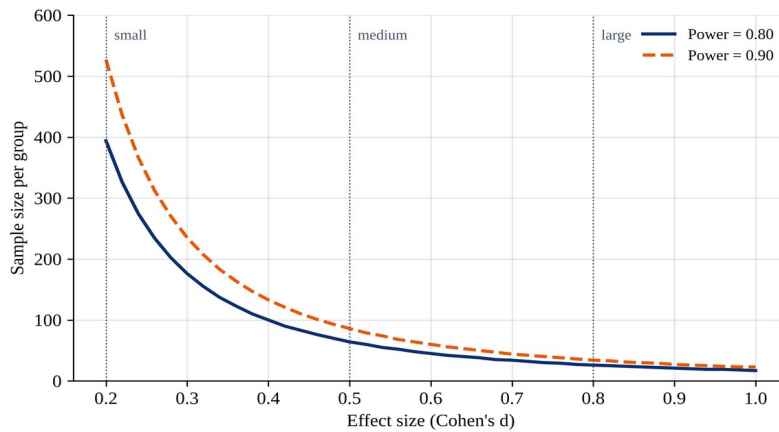
When a study aims to test a hypothesis, the relevant question is different: how large must the sample be to have a good chance of detecting an effect, if a real effect of a given size exists? Power analysis answers this by combining four quantities: the significance level (usually 0.05), the desired power (usually 0.80 or 0.90), the expected effect size and the sample size. Fixing any three determines the fourth (Cohen, 1988).

Cohen (1992) proposed conventional values for effect sizes, such as $d = 0.2$, 0.5 and 0.8 for small, medium and large differences between two means. Table 2 and Figure 2 show the resulting sample sizes for comparing two independent groups with a t-test. Free software such as G*Power performs these calculations for most common tests (Faul et al., 2007).

Table 2. Sample size per group for a two-sample t-test (two-sided, $\alpha = 0.05$).

Effect size (Cohen's d)	Description	Power = 0.80	Power = 0.90
0.2	Small	394	527
0.5	Medium	64	86
0.8	Large	26	34

Source: calculated by the author using the non-central *t* distribution; values agree with those given by G*Power.

**Figure 2.** Sample size per group needed to detect a given effect size in a two-sample t-test ($\alpha = 0.05$).

The contrast with Table 1 is striking. A survey formula might suggest that 385 respondents are enough, yet if the aim is to compare men and women and the true difference is small, nearly 400 respondents are needed in each group. Conversely, for a large effect, 26 per group may be sufficient. Lenth (2001) warns against relying on the conventional small, medium and large values: the expected effect should come from previous studies, pilot data or the smallest difference that would matter in practice.

6. Common errors

- Using Yamane's or Cochran's formula to justify the sample for a study whose main analysis is a test of difference or association.
- Applying the formula to the whole population when results will be reported separately for subgroups; each subgroup then needs an adequate sample.
- Ignoring non-response. If 70 per cent of those approached are expected to respond, the number contacted should be the required sample divided by 0.70 (Bartlett et al., 2001).
- Ignoring the design effect in cluster or multistage sampling, which usually requires a larger sample than simple random sampling.
- Reporting the formula but not the assumptions used: the confidence level, margin of error, expected proportion or effect size.

7. Recommendations

Lakens (2022) describes several legitimate ways to justify a sample size, including measuring the whole population, accepting the limits set by available resources, aiming for a given precision of estimates, and conducting an a priori power analysis. Whatever the approach, the justification should be transparent. Drawing on the literature reviewed here, we recommend that researchers:

1. state whether the main purpose of the study is estimation or hypothesis testing, and choose the method accordingly;

2. report every assumption: confidence level, margin of error, expected proportion or standard deviation, significance level, power and effect size;
3. base the expected effect size on earlier studies or pilot data rather than on conventions alone;
4. adjust the sample for expected non-response and for the design effect where cluster sampling is used;
5. ensure that every subgroup to be analysed separately has an adequate sample.

8. Conclusion

Cochran's formula, the Krejcie and Morgan table and Yamane's formula are closely related tools for one task, estimating a proportion with a chosen precision, and they give very similar answers. They are not suitable for justifying the sample of a study designed to compare groups or test relationships, for which power analysis is the appropriate method and often calls for considerably larger samples. A sample size is not a number to be produced by a formula and forgotten; it is a design decision that should follow from what the study is trying to find out.

Conflict of interest: The author declares no conflict of interest.

Funding: This research received no external funding.

Declaration of generative AI use: During the preparation of this work the author used an AI writing assistant (Claude, Anthropic) to help structure and draft the text. The author reviewed, edited and verified the content and takes full responsibility for this article.

References

- Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, 19(1), 43–50.
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Israel, G. D. (1992). *Determining sample size* (Fact Sheet PEOD-6). Institute of Food and Agricultural Sciences, University of Florida.
- Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, 30(3), 607–610. <https://doi.org/10.1177/001316447003000308>
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8(1), Article 33267. <https://doi.org/10.1525/collabra.33267>
- Lenth, R. V. (2001). Some practical guidelines for effective sample size determination. *The American Statistician*, 55(3), 187–193. <https://doi.org/10.1198/000313001317098149>
- Yamane, T. (1967). *Statistics: An introductory analysis* (2nd ed.). Harper and Row.