Calculating the Sample Size in Quantitative Studies

Ganesh Prasad Adhikari

Central Department of Education

Tribhuvan University, Nepal

Email: gpadhikarin@gmail.com

Abstract

The determination of the required sample size is always an important step during the planning of the quality of research for a quantitative study. The researchers need to have sound prerequisite knowledge of the inferential statistics concept for determining a minimum sample size requirement and its estimation because an insufficient sample size may not be able to answer the research questions and a large sample is wastage of resources, may add to the complexity, and take more time for the study. Both situations are ethically unacceptable for the researchers. Therefore, this study provides researchers with appropriate sample size determination methods for quantitative study. The Google Scholar, PubMed, and Library Genesis were used to find out the required information for the review of this paper. This study presents the sample size calculation formula in a simplified manner with relevant examples so that researchers may effectively use them in their research.

Keywords: Effect size, power of a test, sample size, type I error, type II error

Introduction

A sampling strategy is more than often necessary because it is not always possible to collect data from every unit of the population. Therefore, determining an appropriate sample size is the act used to select the number of observations to include in a statistical sample, the sample size is an important characteristic of any empirical study in which the aim is to make an inference about a population from a sample (Bujang & Adnan, 2016). The sample size is the number of participants or other study units that will be included and required to answer the study's research questions (Gupta, 2011). In particular, too-large a sample is only a wastage of resources, cost, and time and, on the other hand, too small a sample size fails to produce conclusive and reliable results (Gowda et al., 2019). Therefore, it is essential

for the researcher to estimate appropriate sample size to produce reliable results using the statistical procedure. However, the sample size estimate is unclear and indicates less power due to suboptimal sample size. The estimation of the sample size is also essential to know the feasibility of the study in terms of cost and time required (Bujang, 2021). There is no magic solution for sample size estimation (Gowda et al., 2019); however, different statistical formulas are available for calculating sample size when researchers are deal with different types of parameters and variables in different study designs in quantitative studies (Hazra & Gogtay, 2016). But many researchers still don't know which one to use to determine the appropriate formula for determining sample size in their studies. Even if, researchers have an understanding of basic statistical concepts for calculating sample size, they can also use available software for appropriate calculation of sample size in their study. Accordingly, the researcher briefly explains the basic statistical principles and formulas for calculating sample size in quantitative studies.
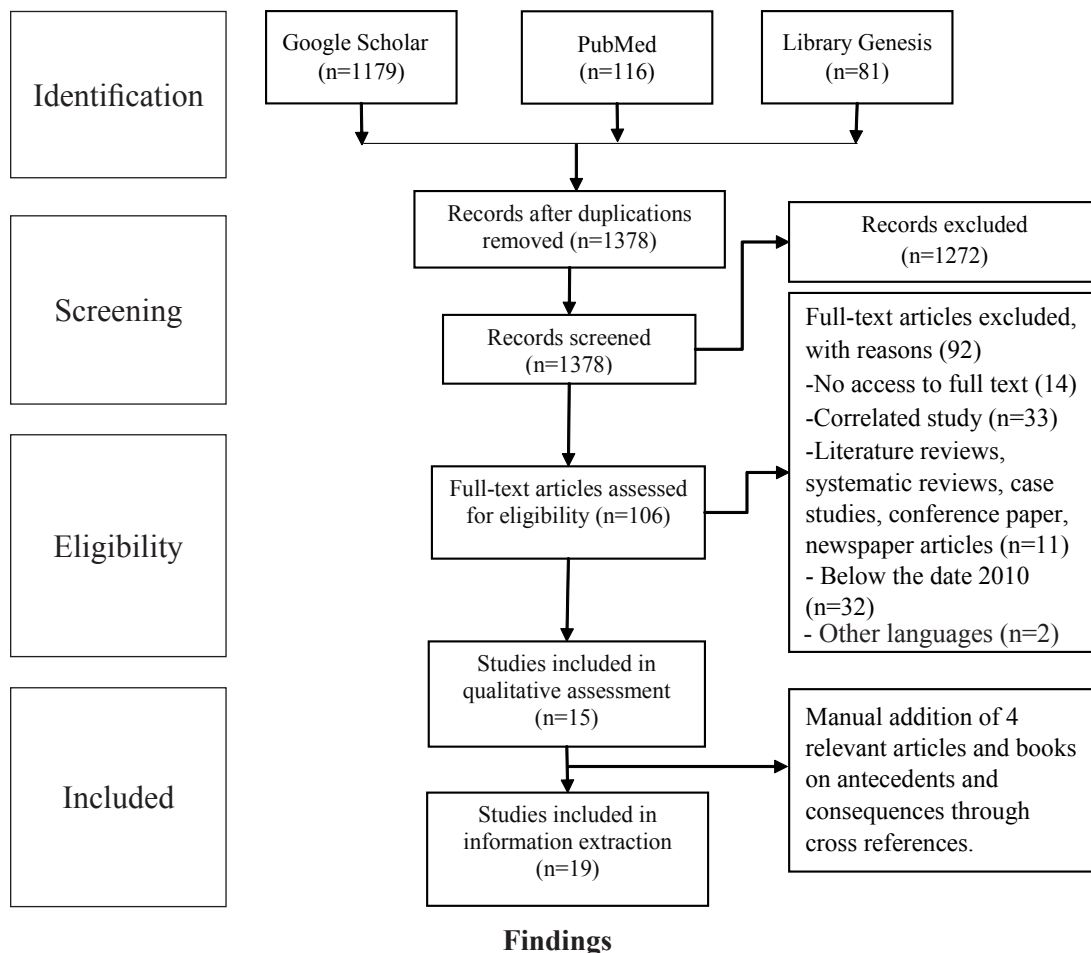
Methods and Procedures

This article is based on original research articles, review articles, and books. These were selected from various sources, as shown in the flowchart (Figure 1). The Google Scholar, PubMed, and Library Genesis databases were used to search the literature. The search terms were used "sample size", "calculation" and "research". To maintain the quality of the exam, all duplicates have been carefully checked. The abstracts of the articles were extensively reviewed for analysis and article purification to ensure the quality and relevance of the academic literature included in the review process. A careful evaluation of each research paper was performed at a later stage. A total of 1378 non-duplicate citations were reviewed. As this is a narrative review, articles with relevant information and illustrations meeting the objective of this review have been included and all other articles have been excluded. The research period was 2010-2021. All articles prior to 2010 were excluded from the search and the inclusion criteria was to limit articles published only in English. There were two articles in a language other than English and they were excluded from the study. In addition, after filtering out duplicate records, by purposive sampling, four additional articles were submitted for the study because they were most relevant to this study. The researcher selected the total of 19books and articles after evaluating each book and article on the inclusion and exclusion criteria. Figure 1 shows the inclusion and exclusion of literature at each stage.

Figure 1
Literature Review Process on Databases

| Identification | Google Scholar (n=1179) | PubMed (n=116) | Library Genesis (n=81) |

Records after duplications removed (n=1378)

Records excluded (n=1272)

Screening

Records screened (n=1378)

Full-text articles excluded, with reasons (92)
-No access to full text (14)
-Correlated study (n=33)
-Literature reviews, systematic reviews, case studies, conference paper, newspaper articles (n=11)
- Below the date 2010 (n=32)
- Other languages (n=2)

Full-text articles assessed for eligibility (n=106)

Eligibility

Studies included in qualitative assessment (n=15)

Manual addition of 4 relevant articles and books on antecedents and consequences through cross references.

Included

Studies included in information extraction (n=19)

Findings

Basic Statistical Conceptions of Sample Size Calculation

In this study, researcher gives the appropriate guidelines to determine the representative sample size for quantitative research works. For this purpose usually we will need to know the basic statistical conception to determine the appropriate sample size for the quantitative study. Each of these is reviewed below:

Null Hypothesis and Alternative Hypothesis

An assumption or a guess or a statement of quantity about the population parameter is called a hypothesis (Gupta, 2011). In quantitative studies, the null hypothesis (H_0) refers to a hypothesis of no significant difference between the statistic and the parameter, i.e. there is no significant difference between a sample statistic and a population parameter (Charan et al., 2021; Rao, 2012). In addition, an alternative hypothesis (H_1) is the oppositee of the null hypothesis, which indicates a significant difference (Das et al., 2016). This is why it is also called the difference hypothesis. It should be noted that the alternative hypothesis

is a mutually exclusive and complementary statement of the null hypothesis. If the null hypothesis is rejected, the alternative hypothesis is accepted. If we test the hypothesis that the population mean (μ) has a specified value μ_0, we define a null hypothesis and an alternative hypothesis as follows:

$H_0 : \mu = \mu_0$. That is, the population mean is equal to μ_0. And the alternative hypothesis will be any of the following three cases

(i) $H_1 : \mu \neq \mu_0$. That is, the population mean is not equal to μ_0. (Two-tailed test).
(ii) $H_1 : \mu > \mu_0$. That is, the population mean is greater than μ_0. (One-sided test on the right).
(iii) $H_1 : \mu < \mu_0$ That is, the population mean is less than μ_0. (One-sided test on the left).

Thus, the researcher should have clearly define null and alternative hypotheses in quantitative studies (Gupta, 2011).

Types of Error

There is always a chance of making a possible error in deciding to accept or reject a null hypothesis based on the sample information. Using only partial information based on a sample, our decision about the population may not be 100% correct. Hypothesis testing involves an inevitable but quantifiable risk of making incorrect decisions and leading to error. When testing the hypothesis, there are two types of errors. (i) Type I error: A type I error occurs when the null hypothesis (H_0) is true but it is rejected, and (ii) Type II error: A type II error occurs when the null hypothesis (H_0) is false ,but it is accepted (Das et al., 2016).The states of nature, the decisions are taken and their consequences are shown in the following table (Gupta, 2011).

Table 1
Decision Making on Null Hypothesis

		The true about H_0	
		True (There is no difference)	False (There is difference)
Decision of the research on the basis of results	Accept H_0 (There is on difference)	Correct decision (no error) probability=$1-\alpha$	Type II error (β)
	Reject H_0 (There is difference)	Incorrect decision (Type I error) probability=α	Correct rejection

Thus, the probability of type I, and Type II errors are denote by α, and β respectively. Therefore, $\alpha = $ P (Type I error)

$\qquad = $ P (Rejecting H_0, when it is true), and

$\beta = $ P (Type II error)

$\qquad = $ P (Accepting H_0, when it is false)

Figure 2

Type I and Type II Errors under Null Hypothesis (H_0) and Alternative (H_1) Hypothesis Testing, c = Critical Value
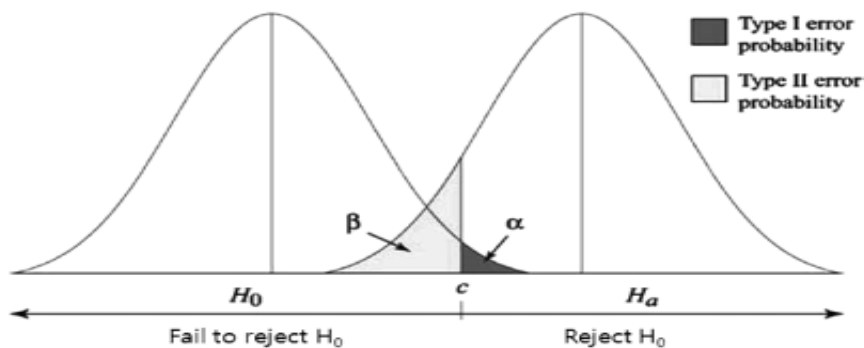
(Pak & Oh, 2012)

Table 2

Relation between Type I Error and Sample Size

Type I Error	Type II Error	Power of Test	Sample Size
1%	13%	87%	55
2%	13%	87%	48
3%	13%	87%	43
5%	13%	87%	38
10%	13%	87%	31

The information mentioned in Figure 2 and Table 2 shows that if the value of α = P (type I error) decreases, then the value of β= P (type II error) increases and vice versa. In any quantitative research, if the type I error increases, then the sample size should be decreased. If the type I error decreases, then the sample size should be increased. Therefore, the researcher should know these concepts in the related research field to find the appropriate sample size.
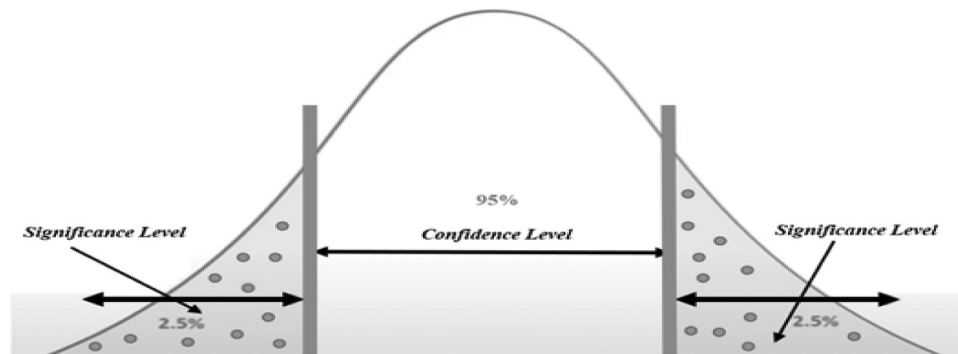
Significance Level

The maximum probability or permissible risk of rejecting the null hypothesis when it is true is called the level of significance. In other words, the maximum size of the committing type I error which we are prepared to risk in the decision process is called the level of significance. It is usually expressed as % and is denoted by α. Hence, α = P (Type I error)

= P (RejectingH_0, when $H_0$is true)

Generally, α is taken as 1% or 5%. If the level of significance (α) is set at 5% i.e. α=5%, we are ready to take a chance of 5% rejecting a true null hypothesis when the experiment is repeated under identical consideration. Therefore, α measures the risk of committing Type I error, (1-α) is the degree of confidence or confidence level or the probability of accepting H_0 when $H_0$is true. Thus, if α=5%, then (1-α) %=95%. This means that we are 95% confident of the correct decision. The higher the significance level we use for testing a hypothesis, the higher the probability of rejecting the null hypothesis when it is true (Stallings & Singhal, 1969).

Figure 3

Significance Level and Confidence Level

(Sayyed, 2020)

Figure 3 shows that when the significance level increases, the confidence level also decreases and vice versa. When a researcher employs a higher level of significance to test a null hypothesis, the likelihood of rejecting the null hypothesis when it is true increases. Thus, the researcher should have to demonstrate a required level of confidence in the correct decision.

One Tail and Two Tail Inferential Statistical Tests

The choice of one-sided or two-sided test depends on the objective of the research. The research hypothesized that a new ICT is a more effective educational tool for reducing students' poor performance in mathematics; then the one-sided test might be sufficient to test the hypothesis, but if you are not sure whether new ICT tools can be more or less effective in student performance compared to existing ICT tools, it is always better to use the two-tail test. The inputs for one-sided testing and two-sided testing are the same except for the critical Z-value; which is different in the one-tailed test ($Z_{1-\alpha}$), and the two-tailed test ($Z_{1-\frac{\alpha}{2}}$) as shown in Table 2.

Table 3
Critical Values of Test-statistic Z

Type of test	Critical values (Z_{α}) at the level of significance (α)		
	1%	5%	10%
Two-tailed test	$\lvert Z_{\alpha}\rvert = 2.58$	$\lvert Z_{\alpha}\rvert = 1.96$	$\lvert Z_{\alpha}\rvert = 1.645$
Right-tailed test	$Z_{\alpha} = 2.33$	$Z_{\alpha} = 1.64$	$Z_{\alpha} = 1.28$
Left-tailed test	$Z_{\alpha} = -2.33$	$Z_{\alpha} = -1.645$	$Z_{\alpha} = -1.28$

(Gupta, 2011)

The information mentioned in Table 2 indicates that the calculated statistical value lies within the intervals of -2.58 to 2.58, -1.96 to 1.96, and -1.645 to 1.645 at the level of

significance (α) = 1%, 5%, and 10% respectively, then the test null hypothesis is accepted, otherwise rejected. Therefore, the researcher should be careful about the decision making in testing the null hypothesis.

Study Power

The study power means the probability of excluding a significant difference when it exists. In other words, it is a probability of generalization of the results of the study to the whole population. An increase in statistical power will reduce the possibility of the occurrence of a type II error (β); this means that it reduces the risk of false negative results. The power is denoted P=1-β. In most statistical trials, the power of 0.8 (80%) or greater is considered more appropriate to find a statistically significant difference. A power of 80% means that there is a 20% chance of accepting the null hypothesis in error i.e. beta is 20% or 0.20 (Charan et al., 2021). For the determination of the sample size, one should know the power of the study. The power reflects the ability to pick up an effect that is present in a population using a test based on a sample from that population (Noordzij et al., 2010).

Figure 4

Power Change when Mean (μ_A) of the Alternative Hypothesis Closes to the Mean (μ_0) of the Null Hypothesis.
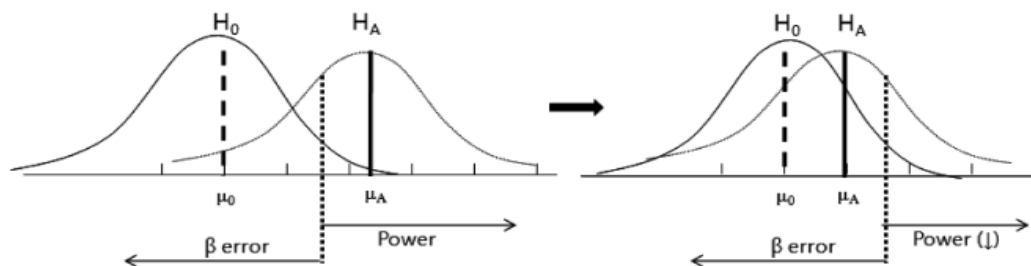
(Pak & Oh, 2012)

Table 4
Relation between the Power, Type II Error, and Sample Size

Type I Error	Type II Error	Power	Sample Size
5%	20%	80%	20
5%	18%	82%	21
5%	16%	84%	22
5%	15%	85%	23
5%	13%	87%	24
5%	12%	88%	25
5%	11%	89%	26
5%	10%	90%	27

Figure 4 and Table 4 show that as the power of the test increases, the type II error decreases and the sample size increases at a 5% level of significance. As a result, the

researcher should understand these concepts and increase the power of the test and sample size to reduce the type II error, which results in the correct decision for testing the null hypothesis in his research field.

Effect Size

The effect size is the smallest difference or the effect that the researcher considers clinically relevant. Determining the size of the effect can be a difficult task. In some cases, it may be based on data from previous studies (Schmidt et al., 2018). A pilot study may be necessary for this purpose or an expert Clinical judgment could be sought. For circumstances where neither of these options apply, Cohen determined standardized effect sizes described as "small", "medium" and "large" (see Table 5). These vary for different study models. For smaller effect sizes, a larger sample size would be required (Bujang, 2021; Cohen, 1988).

Table 5
Small, Medium, and Large Effect Sizes as Defined by Cohen

Test	Effect size	Small	Medium	Large
Difference between two means	d	0.20	0.50	0.80
Difference between many means	f	0.10	0.25	0.40
Chi-squared test	w	0.10	0.30	0.50
Pearson's correlation coefficient	ρ	0.10	0.30	0.50

Table 6
Relation between Effect Size and Sample Size

Effect Size	Sample Size
0.20	196
0.50	32
0.80	13

Table 5 shows that different statistical tests have different effect sizes, and Table 6 shows that when the effect size is increased, the sample size decreases, so that the researcher should increase the sample size to reduce the effect size in his study. However, since the researcher has little time and money, at that time, a large effect size will be taken.

Margin of Error

The margin of error is sometimes referred to as the sampling error or level of precision, which is a probability of variation in the results of sample for the population. It is also defined as the range of values above and below a confidence interval. The margin of error can be calculated in two ways, depending on whether we have parameter or statistics. Therefore, (i) the margin of error = critical value × standard deviation of the population, and (ii) the margin of error = critical value × standard deviation of the sample. In calculating sample size, the margin of error decreases, when the sample size is increased, which show in Table 7. This variation (range) is often expressed in percentage points (i.e. $\pm$ 5%). A

researcher finds that 75% of the students in the sample adopted a recommended practice with a ± 5% margin of error, and then the researcher concludes that 65 % and 80% of the students of the population have adopted the practices (Sharma et al., 2020).

Table 7

Relation Between Sample Size and Marginal Error

*	100	150	200	250	300	350	400	450	500	700	1000	1200	1500	1700	2000	2500	3000
*¥	3.9	3.2	2.8	2.5	2.3	2.1	1.9	1.8	1.6	1.5	1.2	1.1	1.0	0.9	0.9	0.8	0.7

Note. * Sample Size
*¥ Margin of Error

Figure 5
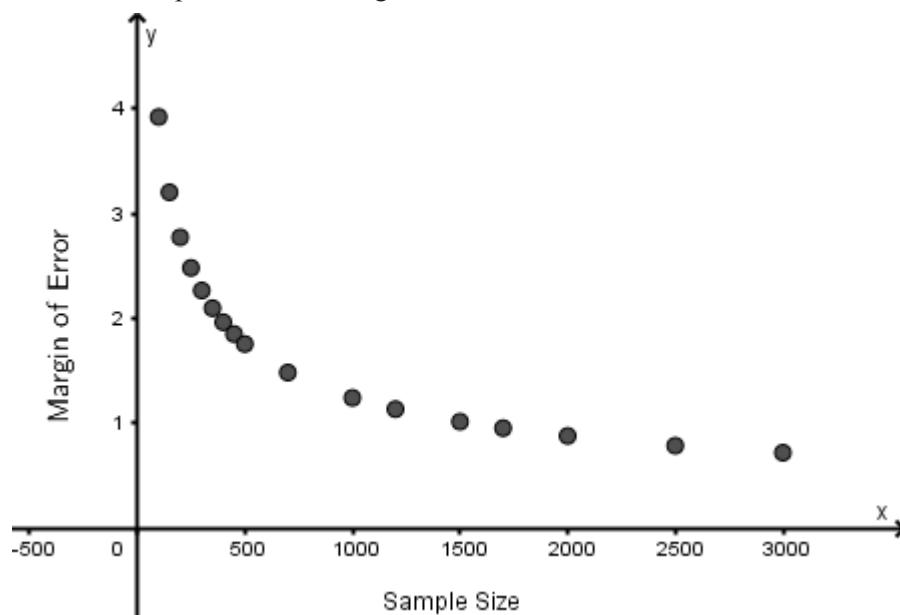Relation between Sample Size and Marginal Error

Figure 5 shows that when the margin of error increases, the sample size decreases. Also, the margin of error degreases. In this case, the sample size has increased. As a result, the researcher should be aware that increasing the sample size in his study reduces the margin of error.

Example 1: Suppose that, a group of researchers want to test the effect of new ICT tools taken by secondary school teachers in Nepal on the performance of mathematics students. As evidence, they want to estimate the IQ scores of thirteen-years-old students how teaching by ICT tools in classroom was. Previous studies suggest that the S.D. of IQ scores of thirteen-years-old students is 25.5. How many such students should the researchers' sample to obtain a 95% confidence interval with a margin of error is 6 points?

Solution: Here, M.E.= 6, Confidence interval = 95%, therefore, the Z-score is 1.96, and σ = 25.5. Now, we have M.E. = $Z \times \frac{\sigma}{\sqrt{n}} \Rightarrow 6 = 1.96 \times \frac{25.5}{\sqrt{n}} \Rightarrow n = (\frac{1.96 \times 25.5}{6})^2 \approx 69$.

Thus, we need at least 69 such students in the sample to a maximum of error of points 6.

The Variability

Finally, for calculating the sample size, the researcher must anticipate the variance of the population of a given result a variable that is estimated using the standard deviation (SD). Researchers often use an evaluation obtained from information in previous studies because the variance is usually an unknown quantity. For a homogeneous population, we need a smaller sample size as the variance or SD will be less in this population. Supposed to study the effect of ICT tools of the mathematics classroom on the performances, we include a population whose performances vary from 30 to 85. Now, it is easy to understand that the SD of this group will be more and we would need a larger sample size to detect a difference between the interventions, if not the difference between the study, the groups would be masked by the inherent difference between them due to the SD. If, on the other hand, we take a sample from population performances between 70 and 85, we would naturally have obtained a more homogeneous group, thus reducing the SD and the sample size (Gupta et al., 2016). Thus, the sample size calculation for continuous outcome measures require an estimate of variability (or SD). Large variability requires a large sample size and small variability requires small sample (Schmidt et al., 2018), which is shown in Figure 6.

Figure 6

Relationship between Sample size and Standard Deviation
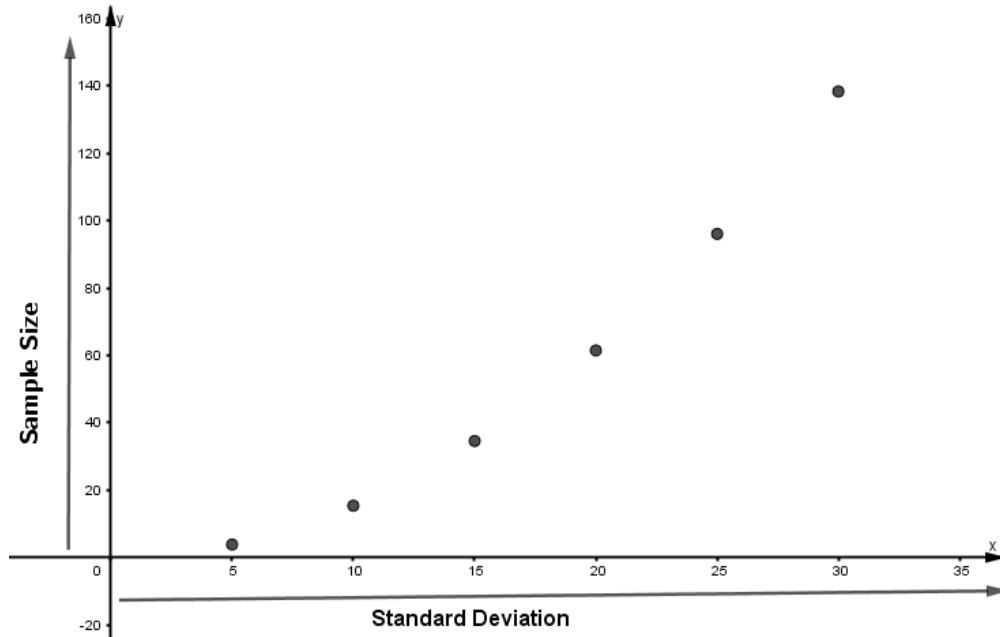
Table 8
Relationship between Sample Size and Standard Deviation

Standard Deviation	5	10	15	20	25	30
Sample Size	4	15	35	61	96	138

Examples for Sample Size Estimation for Quantitative Studies

Several methods used to calculate the sample size, depending on the type of data or study design. In this paper, the following formulas are used to calculate the sample size using the above basic concept for the quantitative study.

Estimation of the Sample Size for Finite Populations

To determine the minimum sample size for finite populations, and the proportion of the population is estimated. The formula for determine of sample size can be used by (Cochran, 1963).

Sample size (n) $= \dfrac{\frac{Z^2 \cdot p(1-p)}{e^2}}{1+\frac{Z^2 \cdot p\,(1-p)}{N e^2}}$, where

n = Sample size
N = Population size
Z = Critical value of desired level of confidence
e = Margin of error/ desired level of precision
p = Maximum probability of variation in the distribution.
For example: Suppose that, we want to evaluate to extension programme using ICT tools in the schools level in which students were encouraged to adopt a new ICT tools in practices at classroom activities. Assume that 2500 students were affected to adoption of new ICT tools in classroom practices, but that we do not know the variability in the proportion that will adopt the practice; therefore, assume p=.5 (maximum variability). Furthermore, suppose we desire a 95% confidence level and ±5% precision. Then, the required sample size is demonstrated as below:

Now, Sample size (n) $= \dfrac{\frac{Z^2 \cdot p(1-p)}{e^2}}{1+\frac{Z^2 \cdot p\,(1-p)}{N e^2}}$, where

n = Sample size, N = 2500, Z = 1.95, e = 0.05, and p = 0.5

Thus, n $= \dfrac{\frac{(1.96)^2 \cdot 0.5 \cdot 0.5}{(0.05)^2}}{1+\frac{(1.96)^2 \cdot (0.5) \cdot (0.5)}{2500(0.05)^2}} \approx 182$

Hence, for evaluation the students' performance in using ICT tools in mathematics classrooms of extension using ICT program, minimum of 182 students will be required in the sample.
Similarly, the sample size determine by (Almeda et al., 2010), when the population is known as below:
Sample size (n) $= \dfrac{N(Z_{\frac{\alpha}{2}})^2}{(Z_{\frac{\alpha}{2}})^2 + 4 N e^2}$, where N= population, e= margin of error, and Z=

critical value of the desired level of confidence.

For example: A researcher is want to test the effect of blended learning during the time of the pandemic to the constitute campus students of Tribhuvan University. The total population of campuses is 350000. Taking all the students in campuses is difficult and expensive, so we decided to take a sample of this population at 5% margin of error. How many students must be considered in the sample?

Here, N= 350000, e = 0.05, and $Z_{\frac{\alpha}{2}}$ = 1.96

Thus, sample size (n) = $\dfrac{N(Z_{\frac{\alpha}{2}})^2}{(Z_{\frac{\alpha}{2}})^2+4Ne^2}$ = $\dfrac{350000 \times (1.96)^2}{(1.96)^2+4\times350000\times(0.05)^2} \approx 384$

Hence, the required sample for the study is 384.

This formula may be used to estimate the proper sample size in circumstances when the population is finite yet heterogeneous, and a margin of error is provided. But when we do not know the variability of proportion, then we use the formula for the sample size (n) = $\dfrac{\frac{Z^2 \cdot p(1-p)}{e^2}}{1+\frac{Z^2 \cdot p(1-p)}{Ne^2}}$.

Estimation of the Sample Size for Infinite Population

To determine the minimum sample size for infinite populations, and the proportion of the population is estimated. The formula for determining sample size can be used:

Sample size (n) = $\dfrac{Z^2 \cdot p(1-p)}{e^2}$, where

n = Sample size

Z = Critical value of the desired level of confidence

e = Margin of error/ desired level of precision

p = Maximum probability of variation in the distribution.

Also, when the population is unknown, then generally the formula for estimating

the sample size can be used by (Almeda et al., 2010) is n = $\dfrac{(Z_{\frac{\alpha}{2}})^2}{4 \cdot (e)^2}$, where n= sample size, $Z_{\frac{\alpha}{2}}$ = 1.96, and e = margin of error. For example, how many possible number of students (n) do the researcher need when conducting research on the efficiency and usability of new ICT tools in mathematics classroom at the university level of the margin of error is 5%.

Here, sample size (n) = $\dfrac{(Z_{\frac{\alpha}{2}})^2}{4 \cdot (e)^2}$ = $\dfrac{(1.9)^2}{4 \cdot (0.05)^2} \approx 384$

Hence, the required sample size is 384. This formula is used to estimate the proper sample size in circumstances when the population is infinite and the margin of error is considered. But the formula n = $\dfrac{Z^2 \cdot p(1-p)}{e^2}$ is used when the population parameter is proportion and margin of error are given.

Estimation of Sample Size for Cross-sectional or Descriptive Studies

In the studies for cross-sectional or descriptive studies in case the data are on nominal/ordinal scale and proportion is a parameters, then generally the formula to determine the sample size can be used (Charan et al., 2021):

Sample size (n) = $\dfrac{(z_{1-\frac{\alpha}{2}})(p)(q)}{(e)^2}$, where

n = desired sample size, $Z_{(1-\frac{\alpha}{2})}$= critical value and a standard value for the corresponding level of confidence (at 95% confidence level or 5% level of significance (type-I error) it is 1.96 and 99% confidence level it is 2.58), p = Expected prevalence or based on previous research, q = 1-p, and e = Margin of error or precision (Sharma et al., 2020). For example: A researcher wants to carry out a descriptive study to understand the proportion of performance of students' use of ICT tools in the mathematics classroom in Kathmandu city. A previous study stated that students' performance in using ICT tools was 45%. At a 95% confidence level and 5%margin of error, calculate the sample size required to conduct the other new research. Here, $Z_{(1-\frac{\alpha}{2})}$= 1.96, p =0.45, q = 1-p = 0.55, and e = 0.05.

Thus, the required sample size (n) $= \frac{\left(z_{1-\frac{\alpha}{2}}\right)(p)(q)}{(e)^2} = \frac{1.96 \times (0.45)(0.55)}{(0.05)^2} \approx 194.$

Now, n = 194 + 19= 213 (let us suppose that 10% will be the dropout of students in the study).

Hence, for conducting a new cross-sectional study to identify the proportion of students' performance using ICT tools in mathematics classrooms, minimum of 213 students will be required in the sample. This formula for finding the appropriate sample size is similar to that used by (Pourhoseingholi et al., 2013).

Also, if the sample size in the case the data is on interval or ratio scale, then the formula of simple size determine is n $= \frac{\left(z_{1-\frac{\alpha}{2}}\right)^2 \cdot (\sigma)^2}{(e)^2}$, where n = sample size, $Z_{1-\frac{\alpha}{2}}$ = standardized value for the corresponding level of confidence at 95% confidence level, it is 1.96, at 99% confidence level or 1% type I error it is 2.58), e = margin of error or rate of precision, and σ = S.D. which is based on a pilot study or previous study (Sharma et al., 2020). For example, a teacher wants to know the average performance of mathematics among bachelor level student in the Tribhuvan University at 95% confidence level, and the margin of error is 0.05. From the pilot study, the S.D. of the performance scores of bachelor-level students in mathematics was found to be 5.6. How many students will be required to conduct a new study?

Here, $Z_{1-\frac{\alpha}{2}}$ = 1.96, e = 0.05, and σ = 5.6,

Thus, sample size (n) $= = \frac{\left(z_{1-\frac{\alpha}{2}}\right)^2 \cdot (\sigma)^2}{(e)^2} = \frac{(1.96)^2 \cdot (5.6)^2}{(0.05)^2} \approx 48.$

Therefore, sample size (n) = 48+ 5 (considering 10% drop out of study students)= 53. Hence, for conducting a new survey study to obtain the average performance of bachelor-level students in Tribhuvan University, minimum 53 students will be required. This result is similar to that used by (Vishwakarma, 2014). This formula will be used when the outcome variable is continuous and the population parameter is mean.

Slovin's Formula in Determining the Sample Size

Let N be the population size and the margin of error (e) denotes the allowed probability of committing an error in selecting a small representative of the population, the sample size (n) can be obtained from the formula n $= \frac{N}{1+Ne^2}$ (Slovin, 1960). To use this formula, first, the researchers figure out the margin of error. For example, a confidence level of 95% gives a margin of error 0.05, 98% confidence level gives a margin of error 0.02, and so on

may be required. If the researcher to fit the population size and required margin of error into the formula, then the result equal the number of sample size required to evaluate the population. For example, suppose that a researcher wants to conduct a research in his/her school/university, which new ICT tools are best for mathematics classrooms to grab good performance in the total number of students of 5567. For this survey, a margin of error of 0.05 is considered sufficient accurate. By using the Slovin's formula, the required sample survey size equals

$$n = \frac{N}{1+Ne^2} = \frac{5567}{1+5567 \cdot (0.05)^2} \approx 373$$

Hence, the required sample size for the survey is 373. In situations when the population is finite and the researchers do not have enough knowledge about the population's behaviour (or distribution of behaviour) to determine the optimal sample size, this formula can be used to estimate it.

Standard Sample Size Determination Formula

According to Cochran (1963). The standard sample size formula is $n = \frac{Z^2 (p)(q)}{(e)^2}$, where n= sample size, Z= standard error associated with the chosen level of confidence, p= variability/ standard deviation (it can be taken from previous studies or pilot study), q=1-p, and e= acceptable sample error. It is normally used 0.05. For example, in conducting the survey research, how many students are fully satisfied with the teacher when using the ICT tools in mathematics classroom? Assume that, variability (p)= 0.5, confidence level (1-α)= 95%, and sampling error (e) = 5%, then $n = \frac{Z^2 (p)(q)}{(e)^2} = \frac{(1.96)^2 (0.5)(0.5)}{(0.05)^2} \approx 384$.

Hence, the required sample size is 384.

Given a desired degree of precision, a desired confidence level, and the expected fraction of the attribute existing in the population, the Cochran method may be used to find an appropriate sample size. In cases involving enormous populations, Cochran's formula is very useful.

Conclusion

The primary goal of the study, the type of outcome variable, and the statistical analysis plan all influence on the choice of a sample size formula. Sample size estimation is very important and choosing the right formula is crucial in all types of research studies to be carried out. Calculating the sample size can be a difficult task and can lead to misleading or incorrect results without prior knowledge of the basic concept of inferential statistics, especially for beginner researchers. Therefore, this study provides brief and clear guidance on determining the appropriate sample size for quantitative research depending on the study type, the null hypothesis, alternative hypothesis, types of errors, level of significance, one-sided testing, two-tailed test, power of the study, effect size, the margin of error, and variability. Thus, this paper will be very useful for new researchers as it provides a rough guide to obtaining the minimum sample size required for their studies to be conducted on quantitative research with minimum cost, manpower, availability of subjects, and time constraints in particular, and feasibility in general. Finally, the researcher suggested that the G power test and other statistical software be used to calculate precise sample sizes in future investigations.

Almeda, J., Capostrano, T., & Sarte, G. (2010). *Elementry statistics*. University of Phillippines Press.

Bujang, M. A. (2021). A step-by-step process on sample size determination for medical research. *Malaysian Journal of Medical Sciences*, *28*(2), 15–27. https://doi.org/10.21315/mjms2021.28.2.2

Bujang, M. A., & Adnan, T. H. (2016). Requirements for minimum sample size for sensitivity and specificity analysis. *Journal of Clinical and Diagnostic Research*, *10*(10), YE01–YE06. https://doi.org/10.7860/JCDR/2016/18129.8744

Charan, J., Misra, S., Kaur, R., Bhardwaj, P., Singh, K., & Ambwani, S. R. (2021). Sample size calculation in medical research: A prime. *National Academy of Medical Sciences (India)*, *57*(2), 74–80. https://doi.org/https://doi.org/ 10.1055/s-0040-1722104

Cochran, W. G. (1963). *Sampling techniques* (2nd ed.). Johnwilley and Sons Inc.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates Inc.

Das, S., Mitra, K., & Mandal, M. (2016). Sample size calculation: Basic principles. *Indian Journal of Anaesthesia*, *60*(9), 652–656. https://doi.org/10.4103/0019-5049.190621

Gowda, G. S., Komal, S., Sanjay, T. N., Mishra, S., Kumar, C. N., & Math, S. B. (2019). Sociodemographic, legal, and clinical profiles of female forensic inpatients in Karnataka: A retrospective study. *Indian Journal of Psychological Medicine*, *41*(2), 138–143. https://doi.org/10.4103/IJPSYM.IJPSYM_504_19

Gupta, K. K., Attri, J. P., Singh, A., Kaur, H., & Kaur, G. (2016). Basic concepts for sample size calculation: Critical step for any clinical trials! *Saudi Journal of Anaesthesia*, *10*(3), 328–331. https://doi.org/10.4103/1658-354X.174918

Gupta, S. C. (2011). *Fundamentals of statistics* (9th ed.). Himalaya Publishing House.

Hazra, A., & Gogtay, N. (2016). Biostatistics series module 5: Determining sample size. *Module on Biostatistics and Research Methodology for the Dermatologist*, *61*(5), 496–504. https://doi.org/10.4103/0019-5154.190119

Noordzij, M., Tripepi, G., Dekker, F. W., Zoccali, C., Tanck, M. W., & Jager, K. J. (2010). Sample size calculations: Basic principles and common pitfalls. *Nephrology Dialysis Transplantation*, *25*(10), 1388–1393. https://doi.org/10.1093/ndt/gfq563

Pak, S. I., & Oh, T. H. (2012). Sample size calculation in medical research. *Journal of Veterinary Clinics*, *29*(1), 68–77. www.ksvc.or.kr

Pourhoseingholi, M. A., Vahedi, M., & Rahimzadeh, M. (2013). Sample size calculation in medical studies. *Gastroenterology and Hepatology from Bed to Bench*, *6*(1), 14–17. https://doi.org/10.22037/ghfbb.v6i1.332

Rao, U. K. (2012). Concepts in sample size determination. *Indian Journal of Dental Research*, *23*(5), 660–664. https://doi.org/10.4103/0970-9290.107385

Sayyed, S. (2020). *What is the significance level, confidence level and P-value and how they related to hypothesis testing?* https://medium.com/@saima.sayyed/what-is-the-significance-level-confidence-level-and-p-value-and-how-are-they-related-to-692e8b2dbf4d

Schmidt, S. A. J., Lo, S., & Hollestein, L. M. (2018). Research techniques made simple: Sample size estimation and power calculation. *Journal of Investigative Dermatology*, *138*(8), 1678–1682. https://doi.org/10.1016/j.jid.2018.06.165

Sharma, S. K., Mudgal, S. K., Thakur, K., & Gaur, R. (2020). How to calculate sample size for observational and experimental nursing research stydies? *Nation Journal of Physiology, Pharmacy and Pharmacology*, *10*(1).

Slovin, E. (1960). *Siovin's formula for sampling technique*. Retrieved on February,13,2013.

Stallings, W. M., & Singhal, S. (1969). Confidence level and significance level. *Journal of Experimental Education*, *37*(4), 57–59. https://doi.org/10.1080/00220973.1969.11011150

Vishwakarma, G. (2014). Statistical basis of calculation of sample size in nursing. In *Wiley StatsRef: Statistics Reference Online* (Issue July). https://doi.org/10.1002/9781118445112.stat06352