Sample Size

The sample size of a statistical sample is the number of observations that constitute it.

The sample size is typically denoted by n and it is always a positive integer. No exact sample size can be mentioned here and it can vary in different research settings. However, all else being equal, large sized sample leads to increased precision in estimates of various properties of the population.

What Should Be the Sample Size?

Determining the sample size to be selected is an important step in any research study. For example let us suppose that some researcher wants to determine prevalence of eye problems in school children and wants to conduct a survey [1].

The important question that should be answered in all sample surveys is "How many participants should be chosen for a survey"? However, the answer cannot be given without considering the objectives and circumstances of investigations.

The choosing of sample size [2] depends on non-statistical considerations and statistical considerations. The non-statistical considerations may include availability of resources, manpower, budget, ethics and sampling frame [3]. The statistical considerations will include the desired precision of the estimate of prevalence and the expected prevalence of eye problems in school children.

Following three criteria need to be specified to determine the appropriate samples size:

1. The Level of Precision

Also called [sampling error](#) [4], the level of precision, is the range in which the true value of the population is estimated to be. This is range is expressed in percentage points. Thus, if a researcher finds that 70% of farmers in the sample have adopted a recommend technology with a precision rate of ±5%, then the researcher can conclude that between 65% and 75% of farmers in the [population](#) [5] have adopted the new technology.

2. The Confidence Level

The [confidence interval](#) [6] is the statistical measure of the number of times out of 100 that results can be expected to be within a specified range.

For example, a confidence interval of 90% means that results of an action will probably meet expectations 90% of the time.

The basic idea described in Central Limit Theorem is that when a population is repeatedly sampled, the average value of an attribute obtained is equal to the true population value. In other words, if a confidence interval is 95%, it means 95 out of 100 samples will have the true population value within range of precision.

3. Degree of Variability

Depending upon the target [population](#) [5] and attributes under consideration, the degree of [variability](#) [7] varies considerably. The more heterogeneous a population is, the larger the sample size is required to get an optimum level of precision. Note that a proportion of 55% indicates a high level of variability than either 10% or 80%. This is because 10% and 80% means that a large majority does not or does, respectively, have the attribute under consideration.

There are number of approaches to determine the sample size including: using a census for smaller populations, using published tables, imitating a sample size of similar studies, and applying formulas to calculate a sample size.

Source URL: https://verify.explorable.com/sample-size

Links
[1] https://verify.explorable.com/survey-research-design
[2] http://en.wikipedia.org/wiki/Sample_size
[3] https://verify.explorable.com/population-sampling
[4] https://verify.explorable.com/sampling-error
[5] https://verify.explorable.com/research-population
[6] https://verify.explorable.com/statistics-confidence-interval
[7] https://verify.explorable.com/statistical-variance