ORIGINAL ARTICLE ▬▬▬▬▬▬

DETERMINATION OF SAMPLE SIZE

Nyi Nyi Naing

Unit of Biostatisitcs and Research Methodology
School of Medical Sciences, Universiti Sains Malaysia, Health Campus
16150 Kubang Kerian, Kelantan, Malaysia

There is a particular importance of determining a basic minimum required 'n' size of the sample to recognize a particular measurement of a particular population. This article has highlighted the determination of an appropriate size to estimate population parameters.

Introduction

As scientific merit and ethical issue go hand-in-hand, the awareness of determination of minimum required sample size and application of appropriate sampling methods are extremely important in achieving scientifically and statistically sound results.

Statistical analysis is based on the key idea that observation on a sample of subjects is made and then draws inferences about the population from which the sample is drawn. If the study sample is not representative of the population, it will well mislead and statistical procedure cannot help. However, even a well-designed study can give only an idea of the answer sought because of random variation in the sample. Thus results from a single sample are subject to statistical uncertainty which is strongly related to the size of the sample.Quality or strength of statistical inference depends largely on the size of the sample selected (1).

In this regard, there are some concerns about determining adequate sample size for studies where particular parameters of populations are estimated. In these studies, the common parameters are either to estimate a 'mean' or 'proportion' of a parameter of a population concerned. Common examples are the estimation of mean body weight of children under five in a population, the estimation of mean blood sugar level of adults aged more than thirty in a population, the estimation of proportion of smokers in a population and the estimation of prevalence of disease 'X' in a population.

Appropriate inclusion and exclusion criteria are also important to obtain participants who have particular characteristics which are under research interest. The other area, which cannot be simply ignored, is testing validity and reliability of measurement tools before actual data collection. A pilot study or a pretest is required to test the validity and reliability of measurement tools. In addition, there are a number of potential biases which can be avoided while designing study and during planning of data collection.

The following is the brief guide to approach how to determine a sample size in this kind of studies. It is based on whether a parameter is a 'mean' or a 'proportion' (2,3,4).

If the parameter is a mean

Sample size n = $\left(z \dfrac{\sigma}{\Delta} \right)$

σ = population standard deviation

$\sigma / \sqrt{n}$

The quantity z representing the width of the confidence interval (precision) is given by Δ.

For a 95% confidence interval :

For a 95% confidence interval :

$$n = \left(\frac{1.96\sigma}{\Delta}\right)^2$$

Some information about the varience of the observations $\left(\sigma^2\right)$ is required before sample size can be calculated. It can be obtained from previous studies in literature or from a pilot study if there was no similar studies conducted before.

Example

We want to estimate the mean systolic blood pressure of Malaysian females. The standard deviation is around 20 mmHg and we wish to estimate the true mean to within 5mmHg with 95% confidence. What is the required sample size ?

Answer

We are given $\sigma = 20$, $\Delta = 5$ and $z = 1.96$.

$$n = \left(\frac{1.96 \times 20}{5}\right)^2 = 61.47 \text{ women}$$

The minimum sample size required is round up to 62 women.

If 20% drop-out is assumed, 62+12 = 74 women needed to be sampled.

If the parameter is a proportion

Sample size

$$n = \left(\frac{z}{\Delta}\right)^2 p\,(1 - p)$$

• p is expected propotion of individuals in the sample with the charecteristic of interect at the $100\,(1-\sigma)\%$ confidence interval.

• The quantity $z\sqrt{\dfrac{p(1 - p)}{n}}$ representing the width of the confidence interval is given by Δ.

To determine the sample size required to estimate the proportion with the desired level of precision, some idea is required beforehand about the possible magnitude of the proportion. It can be obtained from previous studies in literature or from a pilot study if there was no similar study conducted before.

Example

To estimate the proportion of Malaysian males who smoke, what sample size is required to achieve a 95% confidence interval of width ±5% (that is to be within 5% of the true value)? A study some years ago found that approximately 30% were smokers.

Answer

$P = 0.30$, $\Delta = 0.05$ and $z = 1.96$

$$n = \left(\frac{1.96}{0.05}\right)^2 0.3(1 - 0.3) = 322.69$$

The sample of 323 men is required at analysis stage. If 75% response rate is assumed, 323+81= 404 men need to be sampled.

Discussion

These studies have been conducted in many parts of the world. There were some studies which have been done with a very minimum number of samples. The statistical importance of required sample size, to be earmarked as a standard value for a particular population, was rarely emphasized though values have been recognized as references for ages. Health personnel need to be aware of the fact that central to the planning of any such study is the decision on how large a sample to select from the population under study. This article has tried to discuss to a certain extent about important issue of sample size in these studies. However, there are definitely some more facts to be addressed in this particular area of health sciences.

Correspondence :

Dr. Nyi Nyi Naing MBBS, DTM & H MSc(CTM) MPH MMedStats FRSS
Unit of Biostatisitcs and Research Methodology
School of Medical Sciences,
Universiti Sains Malaysia, Health Campus
16150 Kubang Kerian, Kelantan, Malaysia

References

1. Gardner MJ, Altman DG. Statistics with confidence – confidence intervals and statistical guidelines. 1989: 1-13.

2. Dobson AJ. Calculating sample size. Trans. Menzies Foundation, 1984; **7**: 75-9.

3. Lemeshow S, Hosmer DW, Klar J, Lwanga SK. Adequacy of sample size in health studies. World Health Organization. 1990: 1-4.

4. Lwanga SK, Lemeshow S. Sample size determination in health studies, A practical manual. World Health Organization. 1991: 1-3.