International Journal of Multidisciplinary Research and Development
Online ISSN: 2349-4182, Print ISSN: 2349-5979; Impact Factor: RJIF 5.72
Received: 04-12-2018; Accepted: 10-01-2019
www.allsubjectjournal.com
Volume 6; Issue 2; February 2019; Page No. 78-81

The effect of standard error and population proportion in sample size determination

Mohammed Omer A Mohammed[1*], Adil MY waniss[2], Afra Hashim[3]
[1, 3] Department of Statistics, College of Science, Sudan University of Science and Technology, Khartoum, Sudan
[2] College of Business and Econometrics, Qassim University, Kingdom of Saudi Arabia

Abstract
The aim of this study was to determine the sample size using simple random sampling techniques. Test from two tails was used at the level of significant, and value of test with Standard error ranging from (0.001 to 0.02), and value of proportions ranging from (0.12 to 0.5). Excel office and SPSS package were applied for sample size generation and analysis. The results showed that, the sample size increases as, the standard error decreases with a normal curve in all values, except in three cases where the standard error values were too small (0.001, 0.002 and 0.003). The three cases showed very large sample size compared with the other values. As a conclusion it was found that, sample size decreases as, the proportion decreases. The analysis of results showed significant correlation between sample size () and standard error (e), and there is significant difference between sample sizes obtained from one tail and two tails.

Keywords: simple random sampling, sample size, select critical value, margin of error, population proportion

1. Introduction

The Survey studies are generally related to the inferences about the characteristics of population under study. The sample size determination is the act of choosing the number of observations or replicates to include in a statistical sample. Sample size is an important feature of any empirical study in which the goal is to make inference about a population from a sample. The sample size determination depends mainly on the degree of precision (standard error) proportion, variance, confidence level, effect of size. If a sample is not true representative of the target population it may lead to unreliable conclusions. So the determination of proper sample size using appropriate technique of sampling is vital in this type of studies. It is very much necessary to have an idea about the effect of precision and proportion on sample size determination [1].

In survey studies, once data are collected, the most important objective of a statistical analysis is to draw inferences about the population using sample information's. E.g. how big a sample is required, this is one of the most frequently asked question by the investigators the aim of the calculation is to determine an adequate sample size to estimate the population with a good precision. To draw inference or to generalize about the population from sample data. The inference to be drawn is related to some parameters of population such as the mean, standard deviation or some other features like the proportion of an attribute occurring in the population. It should be noted that, a parameter is a descriptive measure of some characteristics of the population whereas if the descriptive measure is computed from the observations in the sample it is called a statistics. The Parameters is constant for a population, but the corresponding statistic may vary from sample to sample. The statistical inference generally takes one of two forms, namely, the estimation of population parameters or the testing of hypothesis [1].

Before collecting data, it is essential to determine the sample size requirements of the study. For calculating the sample size requirements of a study we must answers some questions. Do we want to learn about a mean, mean difference, proportion, proportion ratio, or odds ratio. Are we like to estimate something with a given precision, Or with given power [4].

The sample size is to be determined according to some pre assigned degree of precision. The degree of precision can be specified in terms of two criteria. The margin of permissible error between the estimated value and the population value. It is the measure of how close an estimate is to the actual characteristic in the population. The level of precision, may be termed as sampling error. It is the range to which the true value of the precision is to be estimated [1].

When designing a simple random sampling we need to decide what amount of sampling error in the estimates is tolerable and what precision of the estimates must the balanced with the cost of the survey. Although many variables may be measured, an investigator can often focus on one or two responses that are of primary interest in the survey, and use these for estimating a sample size. To estimate the sample size we may need to answer some questions such as, what is the expected of the sample, how much precision we need, what are the consequences of the sample results? And how much error is tolerable? Instead of asking about required precision, many people ask, "What percentage of the population should be include in the sample?" This is usually the wrong question to be asked. Except in cases of small populations, precision is obtained through the absolute size of the sample and not the proportion of the population covered [4].

The desired precision may be made by giving the amount of error that need to tolerable in the sample estimates. The difference between the sample statistic and the related population parameter is called the sampling error. It depends on the amount of risk a researcher is willing to accept while using the data for decisions making. It is often expressed in percentage. If the sampling error or margin of error is ±5%, and 70% unit in the sample attribute some criteria can be concluded that, as (65% to 75%) of units in the population [3].

The reliability of sample information depends mainly on four characteristics, namely the sampling method, the variability of population, the sample size, and the population size. With these four variable features, there are many situations in which different proposed sampling schemes can be employed these depending on desired features. Eventually a judgement may be required as to which of the features exerts the greater influence. Random sampling of a larger proportion of the same population, to increase the size of the sample, will certainly improve the quality of information about the population under study. However, what happens, if we take a sample which is larger in size but the proportion smaller proportion because the population it is taken from is also larger. A superficially attractive argument is to say that if we take a higher proportion of the population we may get better quality information. But, this ignores the fundamental statistical behaviour which gives rise to the laws of large numbers. In small samples individual results exert great influence on the overall picture, whereas their contribution in a larger sample is considerably less influential [2].

2. Methodology

There are different formulae for determination of appropriate sample size when different techniques of sampling are used. Here, we need a formula for representative sample size, when, simple random sampling technique has been used. Simple random sampling is the most common and the simplest method of sampling because each unit of the population has the equal chance of being drawn in the sample. Therefore it is a method of selecting units out of a population of size by giving equal probability to all units [1].

Cochran (1977) developed a formula to calculate a representative sample for proportions as Where, is the pilot sample size, is the selected critical value of desired confidence level, is the estimated proportion of an attribute that is present in the population, e is the desired level of precision [3].

Cochran pointed that if the population is finite, then the sample size can be reduced slightly. This is due to the fact that a very large population provides proportionally more information than that of a smaller population. He proposed a correction formula to calculate the final sample size as. The above formula (1) and (2) we employed in this study.

3. Results and Discussion

Z-test value was then used depending on level significant and test from (one tail, two tails) as below table (table 1). Standard error (margin of error) ranging (0.001 to 0.02), the value of the proportions ranging "between" (0.12 to 0.5) were used.

Following tables shows sample sizes for different proportions and margin of error.

Table 1: Z test values

Two tails	1.96	2.58
One tail	1.645	2.33

Table 2: When and is constant and is variety

					Ratio
1.96	0.5	0.02	2401	2345	1
1.96	0.5	0.019	2660	2591	1.11
1.96	0.5	0.018	2964	2879	1.23
1.96	0.5	0.017	3323.	3216	1.38
1.96	0.5	0.016	3752	3616	1.56
1.96	0.5	0.015	4268	4093	1.78
1.96	0.5	0.014	4900	4671	2.04
1.96	0.5	0.013	5683	5377	2.37
1.96	0.5	0.012	6669	6252	2.78
1.96	0.5	0.011	7937	7353	3.30
1.96	0.5	0.01	9604	8762	4
1.96	0.5	0.009	11857	10600	4.94
1.96	0.5	0.008	15006	13048	6.25
1.96	0.5	0.007	19600	16388	8.16
1.96	0.5	0.006	26678	21060	11.11
1.96	0.5	0.005	38416	27754	16
1.96	0.5	0.004	60025	37510	25
1.96	0.5	0.003	106711	51623	44.44
1.96	0.5	0.002	240100	70597	100
1.96	0.5	0.001	960400	90570	400

By using z-test value from two tail, with level of significance proportion and standard error ranging (0.001 to 0.02) different sizes of samples were obtained (table 2).

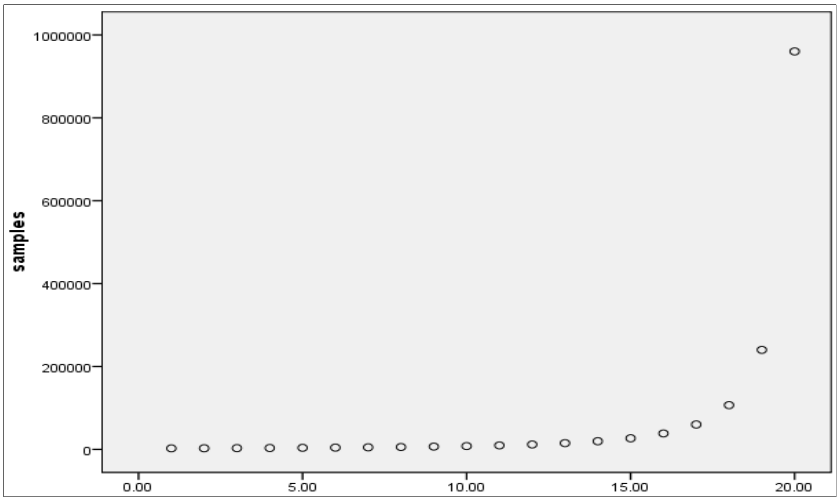

Fig 1: scatter diagram for sample sizes

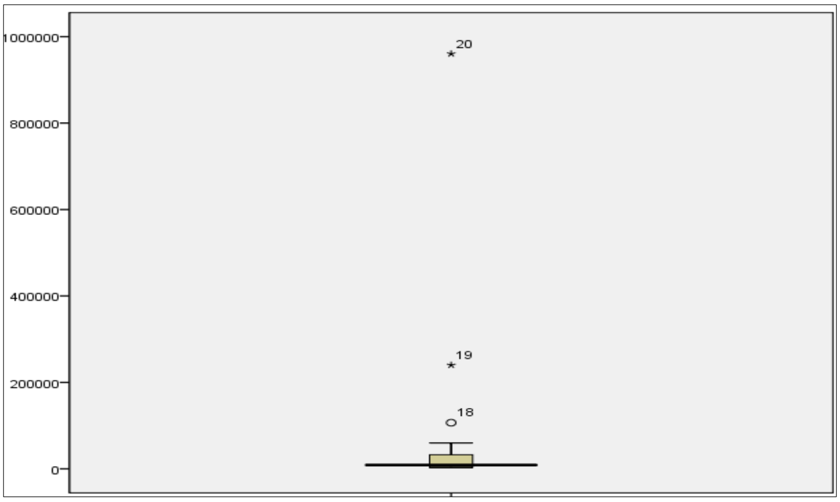

Fig 2: box plot for sample sizes

According to the obtained results in (table 2 and fig1&2) the sample size increases, as the standard error increases, and there is three values of samples out layer of box.

Table 3: When and is constant and is variety

					Ratio
1.96	0.5	0.02	2401	2345	1
1.96	0.48	0.02	2397	2341	0.9984
1.96	0.46	0.02	2386	2330	0.9936
1.96	0.44	0.02	2366	2311	0.9856
1.96	0.42	0.02	2340	2286	0.9744
1.96	0.4	0.02	2305	2253	0.96
1.96	0.38	0.02	2263	2213	0.9424
1.96	0.36	0.02	2213	2165	0.9216
1.96	0.34	0.02	2155	2110	0.8976
1.96	0.32	0.02	2090	2047	0.8704
1.96	0.3	0.02	2017	1977	0.84
1.96	0.28	0.02	1936	1899	0.8064
1.96	0.26	0.02	1848	1814	0.7696
1.96	0.24	0.02	1752	1722	0.7296
1.96	0.22	0.02	1648	1621	0.6864
1.96	0.2	0.02	1537	1514	0.64
1.96	0.18	0.02	1418	1398	0.5904
1.96	0.16	0.02	1291	1275	0.5376
1.96	0.14	0.02	1156	1143	0.4816
1.96	0.12	0.02	1014	1004	0.4224

By using z-test value from two tail, with level of significance proportion ranging (0.12 to 0.5) and standard error is 0.02 different sizes of samples were obtained (table 3).

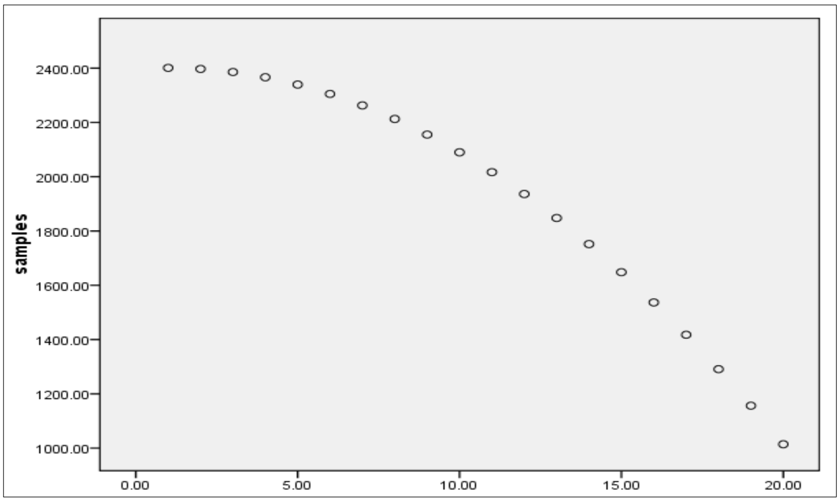

Fig 3: scatter diagram for sample sizes

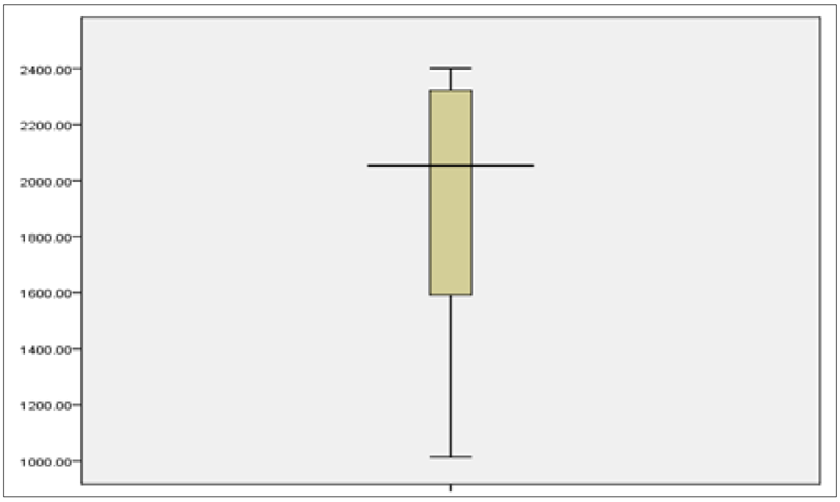

Fig 4: box plot for sample sizes

According to the obtained results in (table 2 and fig 3 & 4) the sample size increases, as the standard error increases.

5. Conclusion

From the obtained result we may conclude that, the standard error and population proportion has strong effect on sample size. The population proportion has negative effect on the sample size, i.e. the sample size there for depends mainly on the degree of precision, and population proportion. In a sample surveying if large sample size is needed we may decrease the standard error or increase the proportion (e.g. clinical research).

6. References

1. Sarmah HK, Bora Hazarika B, Choudhury G. An Investigation On Effect Of Bias On Determination Of Sample Size On The Basis Of Data Related To The Students Of Schools Of Guwahati. 2013; 2(1):33-48.
2. James Nicholson. The Relative Eestimation Sample Size and Proportion, Teaching Statistics. 1997; 19:2.
3. William G Cochran. Sampling Techniques (3rd edition). John Wiley &Sons, New York, 1977.
4. Sharon L Lohr. Sampling Design and Analysis, Second edition, Cengage Learning, v.
5. James A Hanley, Erica EM Moodie. Sample Size, Precision, and Power calculations: A Unifed approach. Journal of Biometrics and Biostatistics. 2011; 2:124.
6. Gene Shack man. Sample size and design effect, 2001.
7. Jones SR, Carley S, Harrison M. An introduction to power and sample size estimation, Emergency Medicine Journan. 2003; 20(5):453.
8. Paula Lagares Barreiro, Justo Puerto Albandoz. Populatechniquesample. Sampling techniques. University of Seville: Management Mathematics and European Schools, 2001.
9. Elizabeth Burmeister, Leanne M Atiken. Sample Size, How Many Is Enough? Offecial. Australian Critical Care. 2012; 4(25):271-274.