

## **A STUDY ON HOW TO PREFER A SAMPLING PROCEDURE AND EMPHASIZE SAMPLE SIZE FOR SCIENTIFIC INVESTIGATION**

**Vivek Kumar Chauhan<sup>1</sup>, Vikas Tyagi<sup>1</sup>, Arun Kumar Bharti<sup>1</sup>, Pradeep Chaudhary<sup>2</sup>**

**Department of Statistics, Meerut College Meerut 250004 India.**

**<sup>1</sup>Department of Statistics, Meerut College Meerut 250004 India.**

**<sup>2</sup>Department of Statistics, CCS University Meerut 250004 India.**

### **Abstract**

An Adequate sampling procedure with precise and emphasizing the sample size the combination of various dynamic selective process which is absolutely vital for any scientific investigation. It is perspicuous that these approach and give outcomes would greatly influence the inclusion and exclusion criteria and the overall generalizability of the investigation locate. This paper has provide systematically transformed the instruction on sampling procedures and determine the sample size, Hence giving the sufficient documentation that will be fruitful in help of the experimenter to furnish the authenticity and gives the statistical power of the scientific cause. Probability sampling is the only can assure the hypothesize the representative sample while non-probability sampling is fruitful in exploratory circumstances. Another significant method that emphasize the optimal sample size. While among distinctive creature has to take into the description of total population size, effective proportion methods, statistical power, confidence level, margin inaccuracy. The paper provides both theoretical instruction and applied tools that experimenter compulsion in choosing the proper techniques for sampling and authenticate sample size procedures for sampling and authenticate procedure for sample size calculation. In aggregate, such a paper sets the criterion for best tendency in research procedures that will prompt the reliability, credibility and empirical and scientific investigation across the distinctive investigation.

**Keyword :** Sample size, Effect Size, Margin Error, Qualitative investigation.

## 1. Introduction

The experimental investigation will be exact and empirical when the proper sampling procedure will be conventional preferred. Sampling approach enclose the choice of a part of the community. And are major parts are obtained through the scientific investigation since through the sampling procedures depends on the characteristics, and discovered of various judgement established on it [6]. The Circumstance of any investigation would based on the suitable of the sampling procedure. The Analytical choice for sampling procedure is the competence of such methods to derive the properties about the population. Without the inconceivable assumption is complex to analyze an comprehensive assessment an complete enumeration of the investigation. This is very significant and compulsory in cover where complete enumeration becomes impractical for a given community due to such functioning consequence as time, opportunity. A adequate preferred sample would obtain the extraneous effectiveness and conception of the decisions [2]. It is generally separated into two categories. Probability and non probability sampling methods [6,11]. Probability sampling includes random sampling , stratified sampling, and cluster sampling where techniques of choosing and established on the arbitrary methods as a vitalize methods to humiliate selection tendency .These procedures of exxagregation of sound statistical principle and are normally borrowed when observation is intentional . Non probability sampling involve the convenience and purposive sampling. Whenever the random nomination is not acceptable. These may be accessible for pragmatic intention. But they will quickly guide to tendency in the sample that than will influence of the outcomes.[10].

The nomination of the sampling procedure is inherently recognized with a few specific exploration objective. Some demographic illuminate of the pick out population and some methodological limitation of the investigation. Proper nomination of sampling procedures bears skilled applicability of to the credibility and generalization of the consideration investigation. Proper selection of sampling procedure bears the proficient to the credibility and observation on the portion of the experimenter who intends to the procedure is accurate . The article reliable to consider the different approaches to sampling in research procedure. And emphasize the most suitability the sample size calculation so that the investigator may choose the best fit for a specific scientific circumstance .This paper provides a close look at the professional and non professional of the separate methods of

sampling in promoting best possibilities of scientific investigation and enhancing empirical credibility.

## 2. Sampling technique Used in scientific investigation

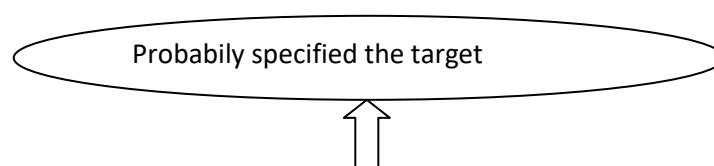
The sampling procedure is most important part used in the research technique. Through which the correctness and crediability of decision from the investigation are assured.[12]. This procedure begins with the authentication of the population. Which could be a entire classification of persons or individual identifying to the scientific query. Defining surely and appropriately what incorporate to the community is extremely fruitful to make assured that the sample illustrate that the specific group under investigation. Basically Population is a group of items unit or subjects which is under reference of study. Which is used such as attribute like demographic components, geographical classification, and every other component is identify and used in research.

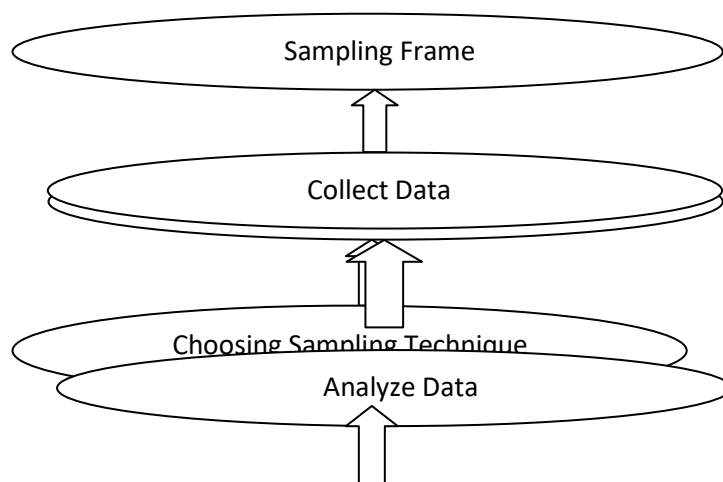
Having definitive the community , the subsequent methods is the preference of the capable of the sampling procedure. In view of the scientific investigation and opportunity and a obligation for conception experimenter have to select the probability sampling methods and non probability sampling methods. The rationale for inclination of the objective, and make the hypothesis about the community. Normally includes the simple random sampling, stratified sampling, cluster sampling and systematic sampling etc.

Once the selected sampling plan is generated, the dimension of the representative sample is to be evaluated by the investigator. And this is taken as broadly in the validity of the outcomes . In determination the sample size different factors are to be included such as size of the population, confidence level and the error such as type 1 error or type 2<sup>nd</sup> error. This outcome can be achieved by the statistical tools like spss and sas software .

Finally the experimenter must analyze the data obtained from sample taking into contemplation that no conclusion can be if the proper sampling technique is not intended. A systematic sampling plan which empower the validity and reliability of the outcomes and therefore the accuracy and purposeful of his conclusions.

The procedure are showing and the representative of the sample technique are showing in the following diagram





### 3 Sampling Approach

Discriminating among the two probability sampling and non-probability sampling are mainly two distinctive arrangement adopted in investigation for the nomination of sharer. [6]. Every procedures has its possess the opportunity and decay. Probability sampling ensures that every unit from the community carries a known, no body prospect of nomination. The Arbitrary in this sampling frame dismiss bias and make the sample delegate the population as a whole. Probability sampling provides the investigator with the possibility to make the hypothesis from sample to the population and estimate the sampling error with the initiative of composure. Example of such category include simple random sampling, stratified sampling and cluster sampling. While probability sampling is accurate it tends to be inconvenient and need much duration and possibility. Especially when the base of community being targeted is very considerable and selection must be accomplished using a very complicated Methods.

On the opposed non-probability sampling enclose no used the random technique therefore not all concern in the community carry on equal probability of choice. This is mainly used when it is not achievable. For the investigator and time are circumscribed. This latter procedure while faster, affordable, and effortless to carry out has the impairment of selection bias that is sample chosen may contain performed or not performed of groups. Outcomes from non-probability sampling therefore have the limited conclude to the community at broad level and statistical inference can not carried out as positively. It is however very beneficial in exploratory investigation qualitative investigation and where the ambition of the investigation is to investigation some of those to carry subgroups rather than to description the complete enumeration of the sample.

Where the scientific investigation gives the sound to the high amount of exactness and postulate the high degree of accuracy and preferably the outcomes to the complete enumeration the inclination is for probability sampling. When we take the exploratory investigation qualitative investigation and where the purpose is to resort to non-probability sampling it is more important than representativeness. The different technique used in position of investigation but the choice of one or other things depends upon the scientific investigation like objective of the research, resources and needs regarding statistical reliability and Validity.

### 3.1 Probability sampling procedure

#### **3.1.1 Simple Random Sampling**

In this sampling we have the unit selected in random manner. The merits of this sampling is to have the equal possibility to selected in the sample. There is a very less chance the sample is become bias. And it is very easy to performed to apply the statistical tools because the nature is the random . And the selection of bias would be less likely to happen with more reliable outcomes and discovered the recommendation. In other hands the demerits of this sampling are it requires a complete and accurate allegation of the community and the process is very complicated when we have taken the large population or mixed population. And it is very instant involving process especially if we have take the broad population. And if the sample size is very large then the process is very complicated.

#### **3.1.2 Stratified sampling**

In this sampling we have ensure the representation of key subgroups or strata leads to have the more sufficient and relevant estimates. In this sampling we ensure that the the efficiency is increased and we used the variability reduced within the strata Hence it is used the increasing the reliability. It will be the could enhance the comparability between different strata and in other hands the demerits of this sampling are the formation of the strata need for detailed knowledge of population properties which may not be convenient. It is very complex process to interpret than straightforward methods and requirement of more advance rational.

### **3.1.3 Cluster sampling**

Cluster sampling is used when the community segment is divided into small segment. this sampling is very beneficial inexpensive and effective for the geographically migrated population with the help of the sampling we can reduce the price associated with wandering. Save expedient and capture the data hence making the large investigation possible. If the complete enumeration is not possible than cluster sampling is used. In other hand the analysis is very complicated if the representative of the population is poorly defined.

### **3.1.4 Systematic Sampling**

In systematic sampling every  $k^{\text{th}}$  individual in a sample from a representative segment of the community . This method is very easy as comparison in simple random sampling. This sampling is premised on the ordered so it is very beneficial to various organisation. In this sampling may bring bias .If the population given in periodic form in the population with the interval used for sampling in result it is representation of particular properties . other hand it reduces the diversity of the sample.

The non probability are as given in the following manner

## **4.0 Convenience sampling**

In this sampling the data collection is fast as we compare to the other sampling Hence one can get the outcome in a very short duration. In preliminary investigation it is very fruitful when the time are inadequate. It is very easy to conduct the in seperable settings such as neighbourhood. In other hands it is based on the experiences that the possibility the outcome will be biased. This sampling is comes under the category of unrestricted sampling . There is a very high risk of collection of tendency in which implies that the outcome can not be formed in particularly across the complete enumeration of the sample than the chance arise the poor validity .

## **4.1 Purposive Sampling**

This sampling is related to the one can investigation matters and population particularly relevant to the scientific investigation and find the research query in depth. This is Suitable for qualitative investigation when the certain information is available based on the experience. This sampling it is based on the specific representation of the persons. In other hand there may be subject to the biases of the investigator himself because the

sample is not represent the complete enumeration of the sample which can develop the over tendency of specific moment. It is very hard to particular due to the extent the broad of population.

#### **4.2 Snowball Sampling**

In this sampling the sampling frame is beneficial hard to reach complete enumeration of the sample. It will be increasing a person willingness to share the exposure. It is very ideal sampling if we conduct the exploratory investigation when one has the little evidence under consideration the investigation. In other hand the diversity will be reduced hence the outcome is not biased. Basically the participants network are non-representative of broad population.

#### **4.3 Quota Sampling**

Quota sampling ensures that the specific subgroup are performed in the sample which is a prime requirement in scientific investigation concerned with such organisation. It is very beneficial in conducting the scientific investigation such as opinion polls by extorting different decisions simultaneously. It will be implemented relatively very fast thus it is possible to allow the timely data collection. In other hand if the quotas are not well settled than the selection procedure is might get unreasonable.

### **5.0 Sample size determination in scientific investigation**

The sample size emphasize is very broad in scientific investigation for coherent and trustworthy outcomes. [7,8]. In the estimate of population proportion apply for the accurate formula ensures that the sample taken the population and diminish the errors.

#### **5.1 Approach for estimating the population segments**

Among all the techniques used in the scientific investigation perhaps the most familiar is for estimating a population proportion [10]. A commonly used formula is:

$$n = N^2 \cdot p(1-p) / E^2$$

Where

n = imperative sample size

N= z value associated with preferred confidence level.

P= ascertained population proportion

E= preferred exactness of population.

This formula is applicable when a investigator is engaged in the proportion of a community with a specific attribute. A larger sample size degrades the margin of error and hence provides an estimate near to existence.

### 5.2 Approach for evaluating the population mean

The formula for sample size in investigation to assess the population signify such an average remuneration is given by :

$$n = N^2 \times \sigma^2 / E^2$$

Where

n= imperative sample size

N= z value is connected with preferred 1.96 for 95% confidence level .

$\sigma$ = ascertained population standard deviation.

E= disparity between the sample mean and population mean .

This is very beneficial approach when one is functioning the continuous variable. If the scatteredness of the community and the representative of the community is very substantial which is provide the correct ascertain. Similarly the inferior the disparity between the average and the population mean. then the exceptional size is needed.

### 5.3 Cochran's Sample size Approach

This formula is used when the population size is very broad and it is applicable both the cases when population segment and mean are given.

$$n = Z^2 \cdot p (1-p) / E^2$$

Where

n= Original sample size for broad population

Z = z value is associated with 1.96 for 95% confidence interval.

p= ascertained population segments (use 0.5 if unknown) .

E= margin of error



Cochran formula is very beneficial to when cross sectional investigation is applied. When the sample size is very small than adjusted the confined community correction. This is done in the following manner:

$$n_0 = z^2 / 1 + z - 1/N$$

Where

$n_0$  = adjusted sample size

$z$  = original sample size from Cochran's Approach

$N$  = total population size.

#### 5.4 Yamane's Formula

Yamane gives the formula in 1967 is one of the frequent procedure in survey or review of the application. it is beneficial to sample measure for any given community with concern for precision and feasibility. [12]. this method is appropriate when the 95% confidence level and it can be confined the other confidence levels .

The formula for Yamane's sample size calculation is:

$$n = N / 1 + N(e^2)$$

Where

$n$  = sample size

$N$  = population size

$e$  = margin of error

#### 5.4 The involvement of tables in samle size calculation

One common methods of easing the computing complexibilities in emphasizing the sample size is through the utilization of pre determined tables. These tables provide the margin of error, confidence Level, and estimated segment of the population. These Methods are provide a scientific investigation for the researcher.

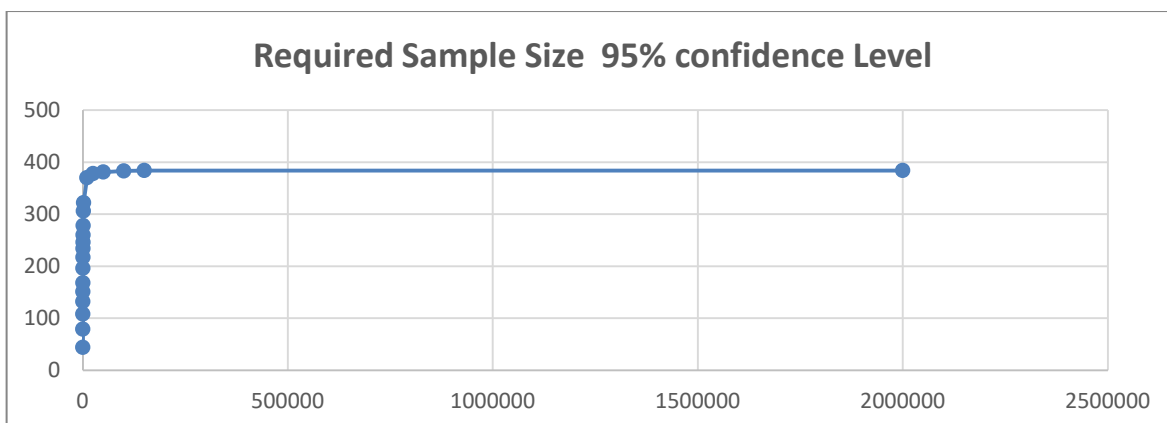
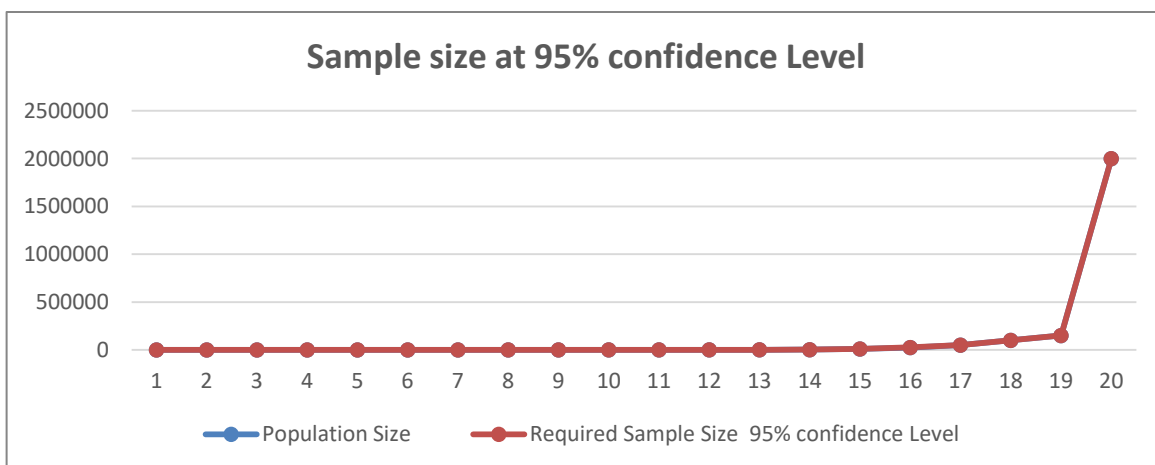
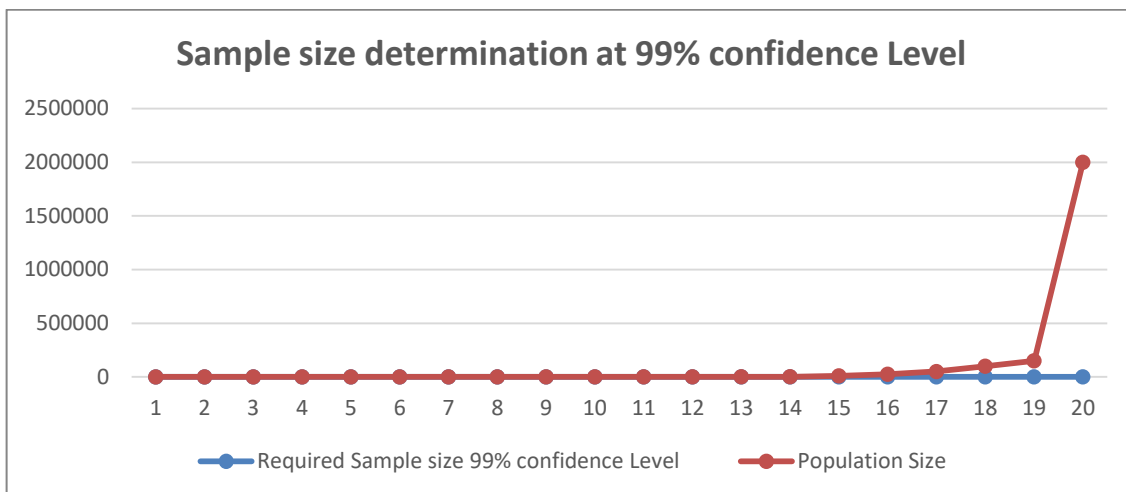
#### 5.5 Various component affected on the sample size

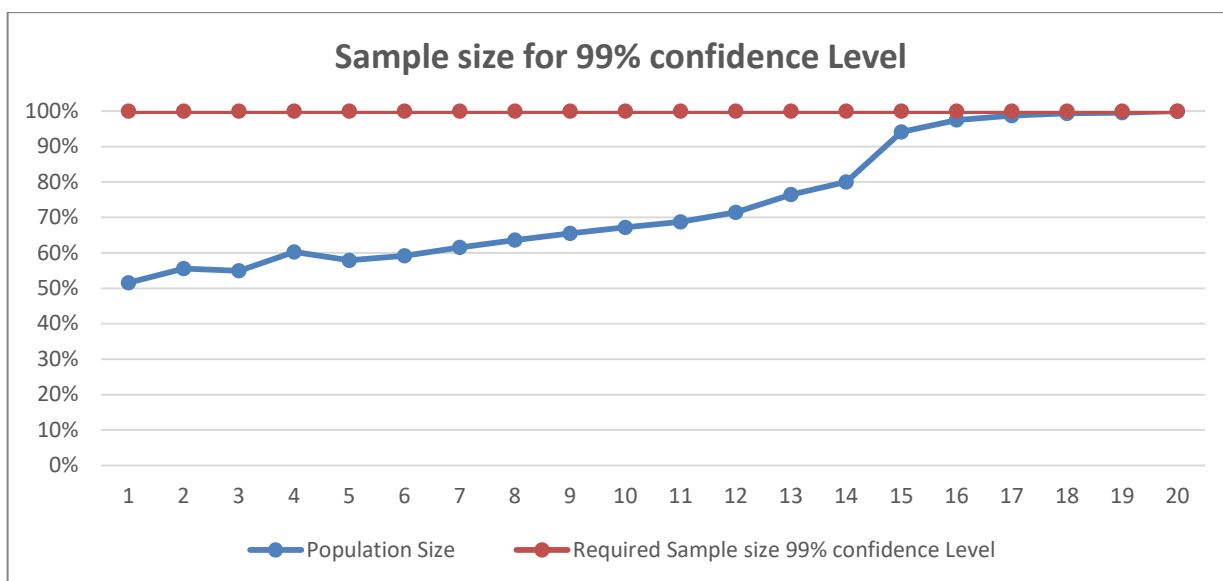
The sample size of any investigation is emphasized by the number of the representative variables. firstly the sample size of the population is represented. The specified confidence level affected the sample size .Other important aspect the investigation of margin of error smaller margin associated with broad sample to ensure finding the characteristics of the true parameters in a specified population. The variability within the population affected the sample size similiary the approach is used in the effect size and statistical power with smaller effect size and high level of power. Larger sample size detect the meaningful differences. The various factor if

considered with due consideration of investigator to such a sample size that is sufficient not only to support the objective of the investigation but the outcome is generated is statistically significant.

**Table 1: Required sample size for categorical data of various population size with distinct margin of error**

Population Size	Required Sample Size 95% confidence Level	Required Sample size 99% confidence Level
50	44	47
100	79	80
150	108	123
200	132	132
250	151	182
300	168	207
400	196	250
500	217	286
600	234	316
700	246	342
800	260	364
1000	278	400
1500	306	462
2000	322	500
10000	370	625
25000	378	635
50000	381	657
100000	383	662
150000	384	663
2000000	384	663

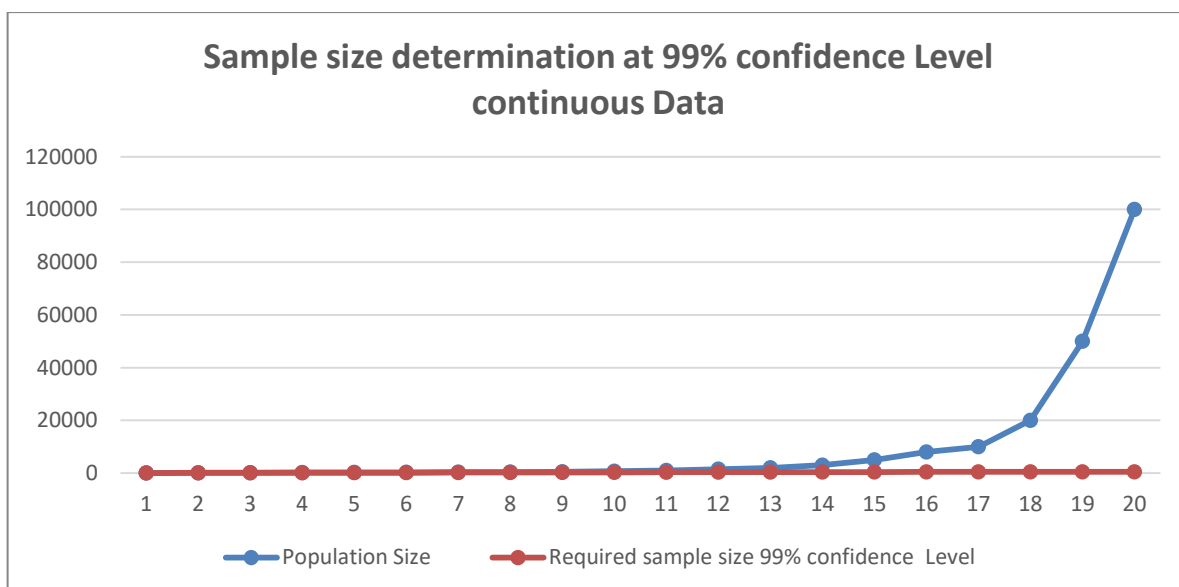


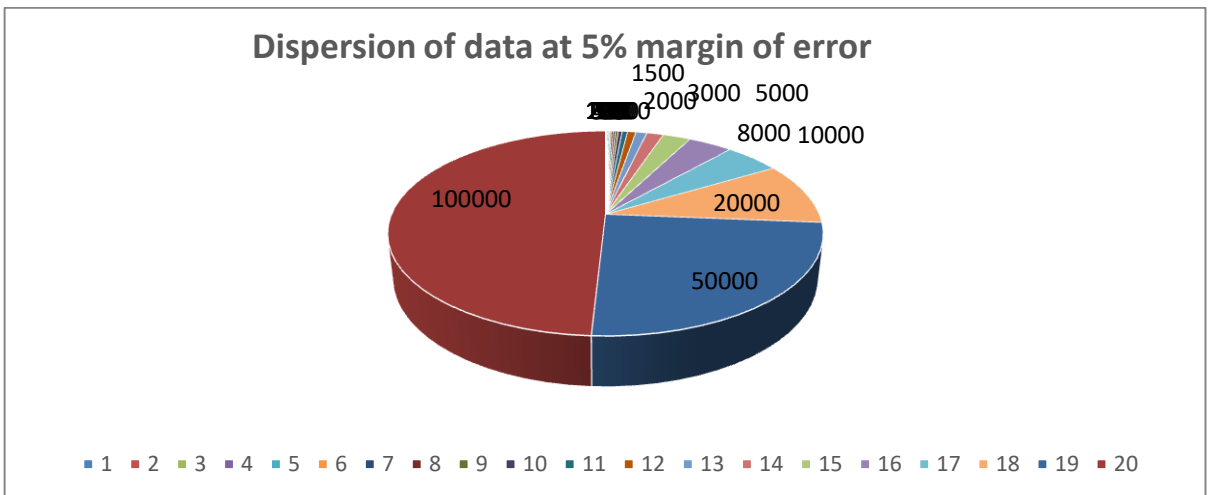
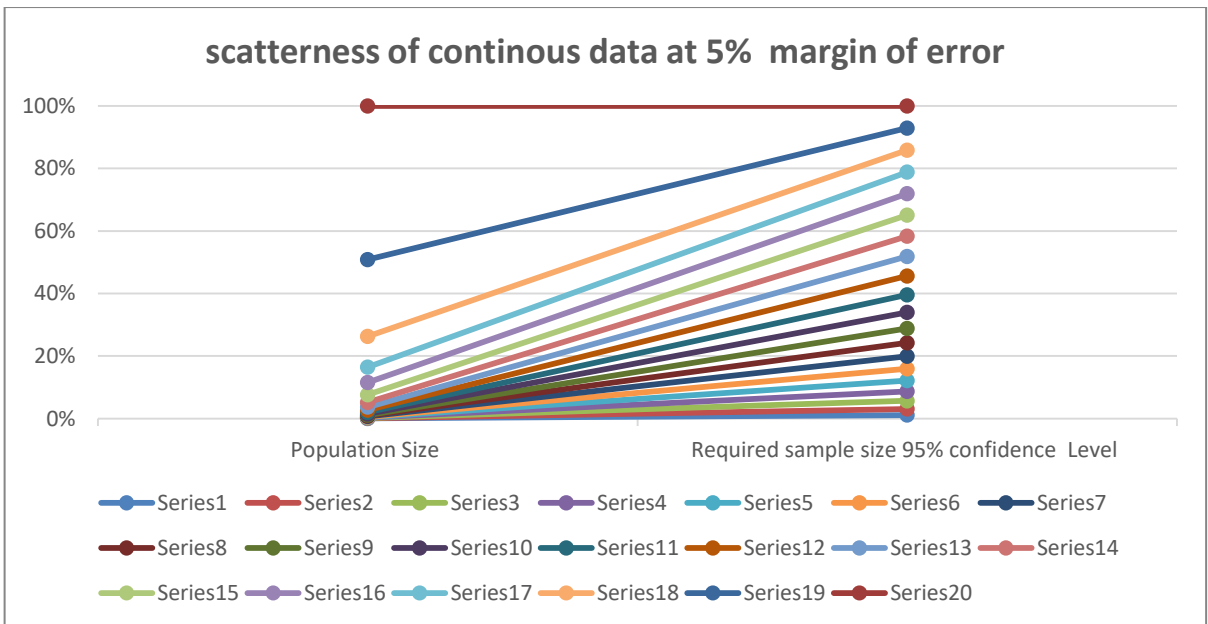
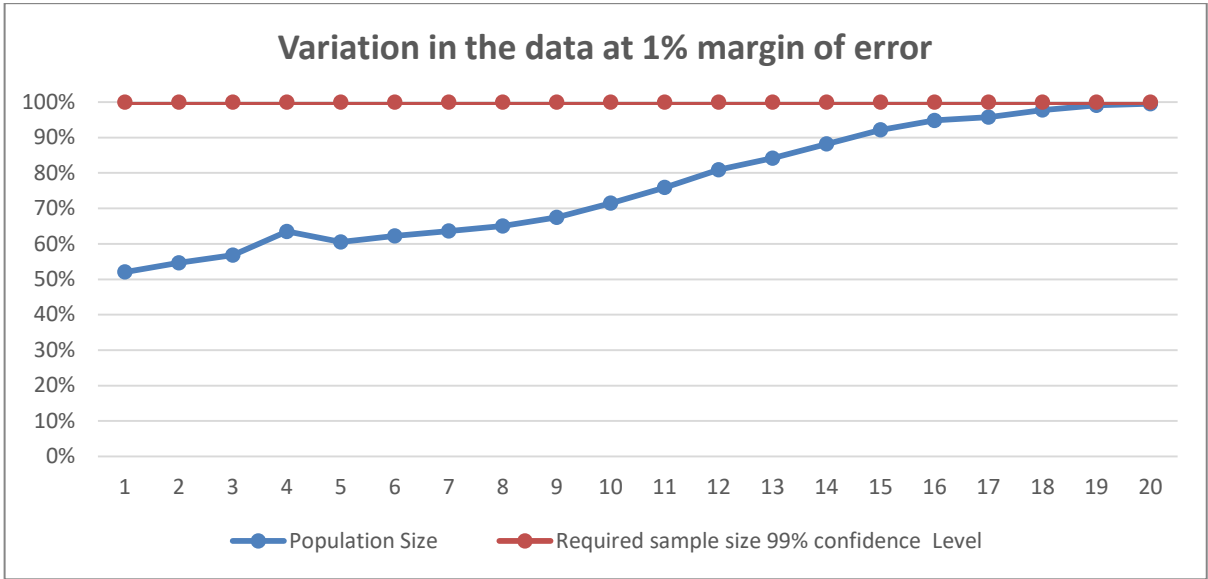


**Table 2: Required sample size for uninterrupted evidence of different population size with distinct margin of error**

Population Size	Required sample size 99% confidence Level	Required sample size 95% confidence Level
50	46	43
100	83	73
150	114	97
200	115	115
250	163	130
300	182	142
350	200	152
400	215	161
500	241	174
700	279	194
1000	317	211
1500	354	227

2000	376	236
3000	401	245
5000	424	254
8000	437	259
10000	442	260
20000	452	264
50000	459	266
100000	461	267





## 6.0 Conclusion

The perfection of the procedure in scientific investigation based on the nomination of proper sampling techniques. And determined the appropriate powered sample size. Probability sampling is thus fundamental in those investigation that tends to construct the remarkable the hypothesis about the broad community . By distinctiveness non probability sampling procedure provide rational drives when one conduct the empirical investigation when complete enumeration access can not be created. Precise computation of sample size is very remarkable step in any scientific investigation. This factor has a great applicability on the statistical dominion and the precision of conclusions that may be representative in population . Confidence level, margin of error, and size of effects that help the investigator ensures their research composition. Facilitate powerful holdings. In that regard the methodological instructions include the ways a investigator that beneficial to research phenomena. This inform that the these instructions optimistically sets a standard to conduct any scientific investigation by setting up works that the where the investigator takes work on the highest empirical and theoretical standards with an ultimate aim of universally increasing the collision of globally extend the commendation and conceivable of scientific investigation.

## References

1. Altman, D. G. (1990). *Practical statistics for medical research*. Chapman and Hall/CRC.
2. Andrade, C. (2021). The inconvenient truth about convenience and purposive samples. *Indian journal of psychological medicine*, 43(1), 86-88.
3. Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.
4. Das, S., Mitra, K., & Mandal, M. (2016). Sample size calculation: Basic principles. *Indian journal of anaesthesia*, 60(9), 652-656.
5. Daniel, W. W., & Cross, C. L. (2018). *Biostatistics: a foundation for analysis in the health sciences*. John Wiley & Sons.
6. . Elfil, M., & Negida, A. (2017). Sampling methods in clinical research; an educational review. *Emergency*, 5(1), e52.
7. Gill, J., & Johnson, P. (2010). *Research methods for managers*.

8. Kotrlik, J. W. K. J. W., & Higgins, C. C. H. C. C. (2001). Organizational research: Determining appropriate sample size in survey research appropriate sample size in survey research. *Information technology, learning, and performance journal*, 19(1), 43.
9. Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and psychological measurement*, 30(3), 607-610.
10. Rahman, M. M. (2023). Sample size determination for survey research and non-probability sampling techniques: A review and set of recommendations. *Journal of Entrepreneurship, Business and Economics*, 11(1), 42-62.
11. Shorten, A., & Moorley, C. (2014). Selecting the sample. *Evidence-based nursing*, 17(2), 32-33.
12. Suresh, K., Thomas, S. V., & Suresh, G. (2011). Design, data analysis and sampling techniques for clinical research. *Annals of Indian Academy of Neurology*, 14(4), 287-290.
13. Taherdoost, H. (2016). Sampling methods in research methodology; how to choose a sampling technique for research. *International journal of academic research in management (IJARM)*, 5.
14. Yamane, T. (1973). *Statistics: An introductory analysis*, second ed. New York.