International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Sample Size Determination and Sampling Technique

Ajithakumari G.

Professor, Department of psychiatric Nursing, Sree Balaji College of Nursing, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India

Abstract: ***Sampling is a process of selecting representative units from an entire population of study. When you conduct a research about a group of people, it's rarely possible to collect data from every person in that group. Instead, you select a sample. The sample is the group of individuals who will actually participate in the research. To draw valid conclusions from your results, you have to carefully decide how you will select a sample that is representative of the group as a whole. This is called a sampling method. There are two primary types of sampling methods that you can use in your research: Probability sampling involves random selection, allowing you to make strong statistical inferences about the whole group. Non - probability sampling involves non - random selection based on convenience or other criteria, allowing you to easily collect data.***

Keywords: Sampling, representative units, population, research, sampling methods, probability and Non - probability sampling

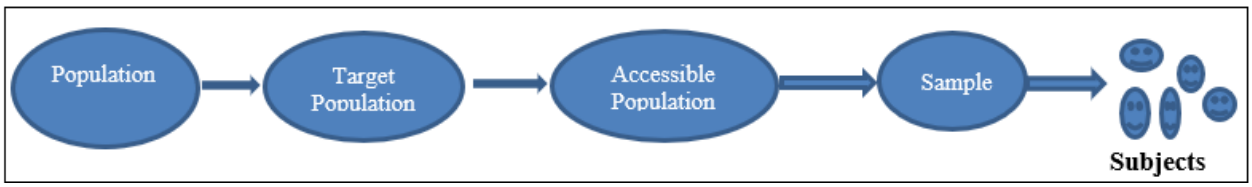
1. Introduction

Sampling is a process of selecting representative units from an entire population of study. Sampling is not a new development, but in recent times it is used by people in all fields, even in day to day life, to get an understanding about societies, opinions or situations. For example, a grain buyer takes handful of grains to get an idea about entire bag of grains; similarly, pathologists take only a few milliliters of blood to measure the hemoglobin of individual.

Similarly, in research studies it is not possible to study an entire population; therefore, the researcher draws a representative part of population through sampling process. In other words sampling is the selection of some part of an aggregate or a whole on the basis of which judgments or inference about the aggregate or mass is made.

Sampling Terminology

As with anything in life you have to learn the language of an area if you are going to use it forever. Now let us see some of the terminologies related to sampling



Population

Population is the aggregate of all the units in which the researcher is interested. In other words, a population is the set of people or entities to which the research are to be generalized. For example a researcher needs to study the problems faced by the post graduate nurses of India; in this the population will be all the post graduate nurses who are Indian citizens.

Target population

A target population consists of total number of people or objects which are meeting the designated set of criteria. In other words, it is the aggregate of all cases with certain phenomenon about which the researcher would like to make a generalization. For example the researcher interested in identifying the complication of diabetes mellitus type II among people who have migrated to south India. In this instance the target population is all the migrants in south India suffering with diabetes mellitus type II.

Accessible population

It is the aggregate of cases that conform to designated criteria and are also accessible as subjects for a study. For example the researcher is conducting a study on the registered nurses working in RGHUS. In this case the population of this study is all the RNs working in RGHUS but some of them may be on leave and may not be accessible for research study. Therefore accessible population of this study will be RNs who meet the designated criteria who are also available for the research study.

Sampling

Sampling is the process of selecting a representative segment of the population under the study.

Sample

Sample is the representative unit of a target population which is to be worked upon by researchers during their study. In other words sample consists of a subset of units which compromise the population selected by investigators or researchers to participate in their research projects.

Element

The individual entities that comprise the samples and population are known as elements and an element is an most basic unit about whom/which information is collected. An

element is also known as subject of the study. The most common element in nursing research is an individual.

Sampling frame
It is the list of all the elements or subjects in the population from which the sample is drawn. Sampling frame could be prepared by the researcher or an existing frame may be used. For example a researcher may prepare a list of all households of a locality which have pregnant women or may use the register of pregnant women for antenatal care available with the local anganwadi worker.

Sampling error
There may be fluctuations in the values of statistics of characteristics from one sample to another or even those drawn from the same population.

Sampling bias
Distortion that arises when a sample is not a representative of the population from which was drawn.

Sampling plan
The formal plan specifying a sample method a sample size and the procedure of selecting the subjects.

Purposes of Sampling

Economical
In most cases it is not possible and economical for researchers to study an entire population. With the help of sampling the researcher can save lots of time, money and resources to study a phenomenon. Therefore, sample provides an economical option for the researcher to generate empirical evidences.

Improve quality of data
It is a proven fact that when a person handles less number of people it is easier to ensure the quality of the outcome. Similarly, in research when a researcher is handling only a part of the population under study, it is easier to maintain the quality of the research work, which would not be possible in case the entire population was involved.

Quick study results
Studying an entire population itself will take a lot of time and generating research results of a large mass will be almost impossible as most research studies have time limits. But with a sample it is possible to generate study results faster which is one of the important objectives of every researcher.

Precision and accuracy of the data
Conducting a study on an entire population provides researchers with voluminous data and maintaining precision of that data becomes a major task while carrying a study on a part of the population helps the researcher to generate more precise data where formulation of the data becomes much easier. It is always easy to establish a better rapport with sample and thus to collect more accurate data. Thus a sample helps to generate precise and accurate data in research study.

Characteristics of Good Sample

Representative
A representative sample is one that the key characteristics of which are closely related to those of the population. Representativeness of the sample makes it possible to generalize the findings of the population.

Free from bias and error
A good sample is one which is free from deliberate selection of the subjects for study. Sample should be free from simple random sampling errors. O sampling bias.

No substitution and incompleteness
A sample is said to be good if once a subject is selected for the study neither is replaced nor it is incomplete in any aspect of researcher's interest.

Appropriate sample size
Generally, it is believed that in quantitative studies a larger the sample size better is the probability of the goodness of the sample however in qualitative studies this notion is not considered important.

Sampling Process
It is the procedure required right from defining a population to the actual selection of sample elements. There are seven steps involved in this process

Step: 1 Define the population
In research population is a precise group of people or objects that possesses the characteristics that is questioned in a study. To be able to clearly define the target population the researcher must identify the specific qualities that are common to all people ofr objects in focus. A population can be simple as all the citizens of India or it can be specific like all male 17 year old high school students with asthma who have been taking bronchodilators since 1 years of age.

It is the aggregate of all the elements defined prior to selection of the sample. It is necessary to define population in terms of
1) Elements
2) Sampling Units
3) Extend
4) Time

A few examples are given here

If we were to conduct a survey on the consumption of green tea among housewives in tamilnadu, then these specifications must be as follows
1) Elements: House wives
2) Sampling units: Households, then housewives
3) Extend: Tamilnadu state
4) Time1 - 10, 2019

It may be emphasized that all these specifications must be contained in the designated population. Omission of any of them would render the definition of population incomplete.

Step2: Identify the sampling frame
Identifying the sampling frame, which could be a telephone directory, a list of blocks and localities of a city, a map or any

other list consisting of all the sampling units? It may be pointed out that if the frame is incomplete or otherwise defective, sampling will not be able to overcome these short comings.

The question is - - - How to ensure that the frame is perfect and free from any defect. Leslie Kish has observed that a perfect frame is one where "every element appears on the list separately, once only once and nothing else appears on the list". This type of perfect frame would indicate one to one corresponds between frame units and sampling units. But such perfect frames are rather rare. Accordingly, one has to use frames with one deficiency or another, but one should ensure that the frame is not too deficient so as to be given up altogether.

This raises a pertinent question. What are the criteria for the suitable frame? In order to examine the suitability or otherwise of a sampling frame, a number of questions need to be asked. These are

1) Does it adequately cover the population to be surveyed?
2) How complete is the frame? Is every unit that should be included represented?
3) Is it accurate? Is the information about each individual unit correct? Does the frame as a whole contain units, which no longer exist?
4) Is there any duplication? If so, then the probability of selection is disturbed as a unit can enter the sample more than once.
5) Is the frame up - to - date? It could have met all the criteria when compiled but could well be deficient when it came to be used. This could well be true of all frames involving the human population as change is taking place continuously.
6) How convenient is it to use? Is it really accessible? Is it arranged in a suitable for sampling? Can it easily be rearranged so as to enable us to introduce stratification and to undertake multistage sampling?

These are demanding criteria and it is most unlikely that any frame would meet them all. Nevertheless they are the factors to be borne in mind whenever we undertake random sampling.

Step 3: Specify the sampling unit

The sampling unit is the basic unit containing the elements of the target population. The sampling unit may be different from the element. For example if one wanted a sample of house wives, it might be possible to have access to such a sample directly. However it is easier to select households as the sampling unit and then interview housewives in each of the households.

As mentioned in the preceding step, the sampling frame should be complete and accurate otherwise the selection of the sampling unit might be defective.

Step 4: Specify the sampling method

It indicates how the sample units are selected. One of the most important decisions in this regard is to determine which of the two - - - probability and non - probability sample is to be chosen.

In case of probability sample the probability or chance of every unit in the population being in the sample is known. Further the selection of specific units in the sample dependence entirely on chance. This means that no human judgment is involved in the selection of the sample. In contrast in a non - probability sample the probability of inclusion of any unit in the sample is not known. In addition the selection of units within a sample involves human judgment rather than chance.

In case of a probability sample, it is possible to measure the sampling error and thereby determining the degree of precision in the estimates with help of the theory of probability. It is not possible when a probability sample is used.

Step 5: Determine the sample size

In other words one has to decide how many elements of the target population are to be chosen.

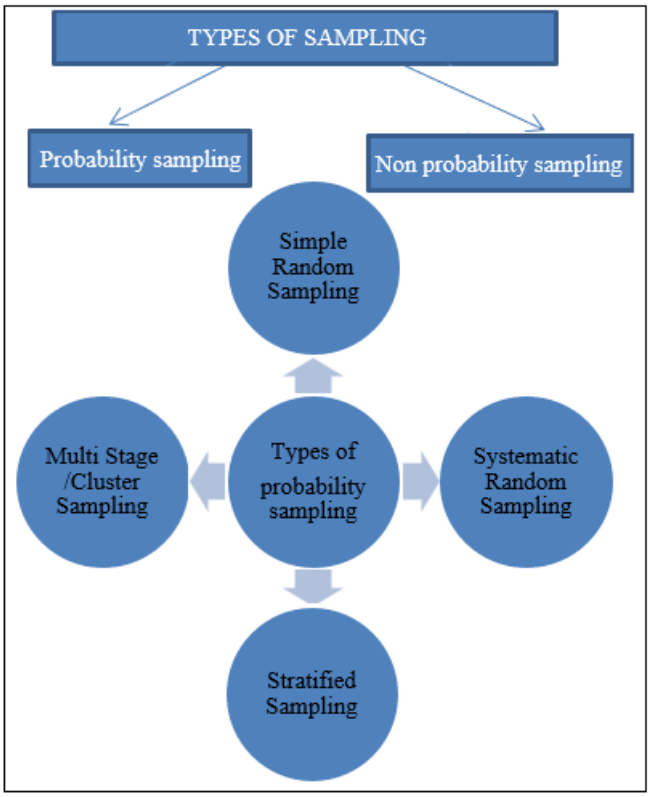
Step 6: Specify the sampling plan

This means that one should indicate how decisions made so far are to be implemented. For example, if a survey of households is to be conducted, a sampling plan should define a house hold, contain instructions to the interviewer as to how he should take a systematic sample of households, advise him on what he should do when no one is available on his visit to the household and so on. These are some pertinent issues in a sampling survey to which a sampling plan should provide answers.

Step 7: Select the sample

This is the final step in the sampling process. A good deal of field work is involved in the actual selection of the sampling elements. Most of the problems in this stage are faced by the interviewer while contacting sample respondents

Types of Sampling



Probability Sampling

In probability sampling every individual in the population have equal chance of being selected as a subject for the research

This method guarantees that the selection process is completely randomized and without bias.

The most basic example of probability sampling is listing all the names of the individuals in the population in separate pieces of paper and then drawing a number of papers one by one from the complete collection of names.

The advantage of using probability sampling is the accuracy of the statistical methods after the experiment. It can also be used to estimate the population parameters since it is representative of the entire population. It is also a reliable method to eliminate sampling bias.

Features of Probability Sampling

- Probability sampling is a technique where in the samples are gathered in a process that gives all the individuals in the population equal chance of being selected.
- In this technique the researcher must guarantee that every individual has an equal opportunity for selection and this can be achieved only if the researcher utilize randomization.
- The advantage of using a random sample is the absence of both systematic and sampling bias. If random selection is done properly, the sample is representative of the entire population.

1) Types of Probability Sampling

Random Sampling

This is one of the most popular type of probability sampling. There are many methods to do it. In this technique each member of the population has an equal chance of being selected as subject. The entire process of sampling is done in a single step with each subject selected independently of the other members of the population.

Methods of Random Sampling

The Lottery Method: The most primitive method. Each member of the population is assigned a unique number and then each number is placed in a bowl or hat and mixed thoroughly. The blind folded researcher picks the numbered tags from the hat. All the individuals bearing the numbers are the subjects for the study.

The Use of Table of Random Numbers: Random tables present several numbers in rows and columns. Researchers initially prepare a numbered list of the elements/numbers of the population and then with a blindfold choose a number from the random table. The same procedure is continued until thedecide number of the subject is achieved.

The Use of Computer: Now a days random tables may be generated from the computer and subjects may be selected as described in the use in random table. For population with a small number of members it is advisable to use the first method, but if the population has many members, a computer aided random selection is preferred.

2) Stratified Sampling Method:

Stratified sampling technique where in the researcher divides the entire population in to different sub groups or strata then randomly selects the final subjects proportionately from the different strata.

Stratified sampling is also known as proportional randomsampling. This is a probability sapling technique where in the subjects are initially grouped in to different classification such as age, socio economic status or gender.

Then the researcher randomly selects the final list of subjects from the different strata.

It is important to note that the strata must be non - overlapping.

The researcher must use simple probability sampling within the different strata.

Types of Stratified Sampling

a) Proportionate stratified random sampling

The sample size of each strata in this technique is proportionate to the population size of the stratum when viewed against the entire population.

A	B	C
100	200	300
½	½	½
50	100	150

b) Disproportionate stratified random sampling

In disproportionate stratified sampling the different strata have different sampling fractions.

3) Systematic Sampling

In this sampling the researcher first randomly picks the first item from the population. Then the researcher will select the n^{th} subject from the list. The process of obtaining the systematic sample is much like an arithmetic progression.

Starting number

The researcher selects an integer that must be less than the total number of population. . This integer will correspond to the first subject.

Interval

The researcher picks another integer which will serve as the constant difference between any two consecutive numbers in the progression.

For example, the researcher has a total population of 100 individuals and need12 subjects. He first his starting number 5 and he picks interval 8. The members of his sample will be individuals 5, 13, 21, 219, 37, 45, 53, 61, 69, 77, 85, 93.

Other researchers use modified systematic sampling, wherein they first identify the sample size. Then they divide the total number of population with the sample size to obtain the sampling fraction. The sampling fraction is then used as the constant difference between subjects.

4) Cluster Sampling

In this sampling technique the researcher takes several steps in gathering his sample population.

First the researcher selects groups or clusters and then from each clusters the researcher selects the individual subjects by either simple random or systematic random sampling.

For example a researcher wants to survey academic performance of high school students in India. He can dividethe entire population in to different clusters (cities), then from selected clusters the researcher can include all the high school students as subjects or he can select a number of subjects from each cluster through simple or systematic sampling random sampling. The important thing to remember is to give all the clusters equal chances of being selected.

Types of Cluster Sample

a) One stage cluster sample
   Recall the example given above; one stage cluster sample occurs when the researcher includes all the high school students from all the randomly selected clusters as sample.
b) Two stage cluster sample
   Two stage cluster sample is obtained when the researcher only selects a number of students' from each cluster by using simple or systematic random sampling.

5) Mixed/ Multistage Sampling

This probability sampling technique involves a combination of two or more sampling techniques enumerated above. For example consider the idea of sampling New york state residents for face to face interviews. Clearly we would want to do cluster sampling as the first stage of the process. We might sample township or census tracts throughout the state. But in cluster sampling we would then go on to measure everyone who is in the clusters we select. Even if we are sampling census tracts we may not be able to measure everyone who is in the census tract. So we might set up a stratified sampling process with in the clusters. In this case we would have two stage sampling process with stratified samples within cluster samples.

6) Sequential Sampling

In this technique sample size is not fixed. The investigator initially selects sample and tries out to make inferences; if not able to draw results he or she adds more subjects until clear cut inferences can be drawn. For example a researcher is studying association between smoking and lung cancer. Initially researcher takes small sample and tries to draw inferences. If unable to draw any inferences he or she continues to draw the sample until meaningful inferences are drawn.

Non- Probability Sampling

Introduction

Non probability samplings do not follow the theory of probability in the choice of elements from the sampling population. Non probability sampling designs are used when the number of elements in a population is either unknown or cannot be individually identified.

Features of Non- Probability Sampling

- Non probability sampling is a technique where in the samples are gathered in a process that does not give all the individuals in the population equal chance of being selected.
- Most researchers are bound by time, money and work force and because of these limitations, it is almost impossible to randomly sample the entire population.
- Subjects in a non - probability sample are usually selected on the basis of their accessibility or by purposive personal judgement of the researcher.
- The downside of this is that an unknown proportion of the entire population is not sampled. Therefore the results of the research cannot be used in generalization pertaining to the entire population.

Types of Non Probability Sampling

Non probability sampling techniques are classified as

- Purposive sampling or judgmental sampling
- Convenience or accidental sampling
- Consecutive sampling
- Quota sampling
- Snow ball sampling
- Theoretical sampling
- Voluntary sampling
- Model instance sampling
- Expert sampling
- Diversity samping
- Event sampling
- Time sampling

1) Purposive Sampling

Purposive sampling is more commonly known as judgmental or authoritative sampling. In this type of sampling subjects are chosen to be part of sample with the specific purpose in mind. In purposive sampling researcher believes that some

subjects are fit for research compared to other individuals. This is the reason why they are purposively chosen as subjects.

2) Convenient or Accidental Sampling
Accidental sampling is also based upon convenience in accessing the sampling population. Whereas quota sampling attempts to include people possessing an obvious /visible characteristic, accidental sampling makes no such attempt. You stop collecting data when you reach the required number of respondents you decided to have in your samples.

3) Consecutive Sampling
Consecutive sampling is very similar to convenience sampling except that it seeks to include all accessible subject as a part of the sample. This non probability sampling technique can be considered as the best of all non - probability samples because it includes all the subjects that are available, which makes the sample a better representation of the entire opulation.

4) Quota Sampling
Quota sampling is a non - probability sampling technique were in the researcher ensures equal or proportionate representation of subjects depending on which trait is considered as the basis of the quota. The bases of the quota are usually age, gender, education, race, religion and socio economic status.

For example if the basis of the quota is college level and the researcher needs equal representation with the sample size of 100, he must select 25 first year students, another 25 second year students, 25 third year students an 25 fourth year students. In this technique samples are selected without random process from different quotas.

5) Snow Ball Sampling or Chain Referral Sampling
This technique is widely used by the researchers to identify potential subjects in study where subjects are hard to locate such as commercial sex workers, drug abusers etc.

Types of snowball sampling
- Linear Snowball Sampling: In this each selected sample is asked to provide reference of only one similar subjects where a linear chain is created by the completion of desired sample
- Exponential No Discriminative Snow Ball Sampling: In this each sample is asked to provide reference of at least two similar subjects because of which the size of the sample grows exponentially.
- Exponential Discriminative Snow Ball Sampling: In this initially one sample is selected and asked for two references of similar subjects. out of which atleast one subject must be active to provide further references another could be non active in providing reference.

6) Theoretical Sampling
It is associated with qualitative research, primarily grounded theory method. As the study data are collected, coded and analyzed the researcher examines the emerging conceptual categories and themes and decides on further data collection procedures that have the potential to contribute to the developing theory.

7) Voluntary Sampling
Voluntary sampling is a type of non - probability sampling procedure in which the volunteers either offer or actively recruited to participate in a study.

8) Model Instance Sampling
Model instance sampling is a type of non - probability sample composed of subjects who represent the typical case that are constructed by the researcher for purpose of the study. The method draws its name from the mode the most frequently occurring score or value in a set of measurements.

9) Expert Sampling
The researcher select study participants based on the need to ascertain how experts in a field would react to or judge the phenomena of interest for the study. For example sample of nurse midwifery educators with expertise n curriculum development specific to the preparation of nurse midwife could be selected for the study proposing to determine the effectiveness of two different nurse midwife cuuricula.

10) Diversity Sampling (Heterogenecity Sampling)
Diversity is a non - probability sampling procedure used when the investigator require that subjects with wide variety of opinion and views included in the sample. To achieve diversity sampling the researcher would include individuals from all segment of the population without regard for representation of persons with these opinions and views as they occur proportionately in the population.

11) Event Sampling
Event sampling is a non - probability sampling procedure used by researchers who are concerned with sampling from those specific occurrences and events that are relevant to the study. For example research student who wrote the proposal compliance with universal precautions would collect her data about nurses using universal precautions would collect her data about nurses using universal precautions only when they were working with children.

12) Time Samping:
Time sampling is a non - probability sampling procedure used by researcher who are concerned with collecting data on activities that take place at specific times of the day or night. For example, a researcher who wanted to observe what was happening during meal time in an immediate care facility would collect data only at times when meals were being served.

Conclusion

The important factor in determining which sampling approach to use is consistent with the research problem and the purpose of the study.

Research Control
Control involves imposing conditions on the research situations. So that biases are maintained and precision and validity are maximised.

Research control most typically involves holding constant other influences on the dependent variable. So that the true relationship between the independent independent variable can be understood. In other words research control attempts to eliminate contaminating factors that might cloud the relationship between the variables that are of central interest.

Matching

The issue of contaminating factors or extraneous variables as they are called can be illustrated with an example. Suppose we are interested in studying whether teenage women are at high risk for having LBW infants than are older mother because of their age.

Two extraneous variables in our hypothetical study are the mothers nutritional habits and her prenatal care.
Mother's age - - - LBW
Mothers age - - - - - prenatal care - - - - - LBW
Mothers age - - - nutrition - - - - - LBW

To have a control the extraneous variables must be held constant

Age of mother	Nutritional percentage	No of visit
15 - 19	33% good 33% fair 33&good	33% 1 - 3 visits 33% 4 - 6 visits 33% >6 visits
25 - 29	33% good 33% fair 33&good	33% 1 - 3 visits 33% 4 - 6 visits 33% >6 visits

This matching approach involves deliberately selecting study participants in such a way that both older and younger mothers have similar eating habits and amounts of prenatal attention.

Randomness

For quantitative researcher a power tool for eliminating bias involves the concept of randomness having certain features of the study established by chance rather than by design or researcher preference. When people are selected at random to participate in the study. (eg) each person in the initial pool has an equal m probability of being selected. This in turn means that there no systematic biases in the makeup of the sample. Randomness is a compelling method of controlling confounding variables and reducing bias.

Masking /Blinding

Researchers want to improve patient care supported and study participants want to be cooperative and helpful and they also want to prevent themselves in the best light. These tendencies can lead to biases because they can affect what participants do and say and in ways that distort the truth.

A procedure known as masking is used in many quantitative studies to prevent biases stemming from awareness. Masking involves concealing information from participants, research agents such as data collectors, care providers or data analysts to enhance objectivity (eg) if study participants are not aware of whether they are getting an experimental drug or placebo then their outcomes cannot be influenced by their expectations of its efficacy.

Now a days the word binding is not used instead Masking is used. This term has fallen in to disfavour when it proves to be unfeasible or undesirable to use masking the study is sometimes called an open study in contrast to a closed study that results from masking.

When masking is used with only some of the people involved in the study ((e. g.) study participant) it is often called single blind study but when it is possible to mask with two groups (e. g. those delivering an intervention and those receiving it) it is called a double blind study.

Reflexivity

In qualitative study researcher often rely on reflexivity. Reflexivity is the process of reflecting critically on the self and of analysing and making note of personal values that could affect data collection and interpretation.

Schwant (1997) has described reflexivity as having two aspects. The first concern is the acknowledgement that the researcher is part of the setting control or social phenomenon under study. The second involves the process of self reflection, about one's own biases, preferences stakes in the research and theoretical inclination. Qualitative researchers are encouraged to explore these issues to be reflexes about every decision made during the inquiry and to record their reflexive thoughts in personal diaries and memos.

Reflexivity can be a useful tool in quantitative as well as qualitative research - - - self awareness and introspection can enhance the quality of any study

Generalizability and Transferability

Generalizability is the criterion used in quantitative studies to assess the extent to which the findings can be applied to other groups and settings. First and foremost they must design studies in reliability and validity. In selecting participants researchers must also give thought to the types of people to whom the results might be generalised and then select subjects in such a way that non biased sample is obtained.

Qualitative researchers do not specifically seek to make their findings generalizable. Nevertheless qualitative researchers often seek understandings that might prove useful in other situations.

LINCOLIN AND GUBA (1985) discuss the concept of transferability the extend to which qualitative findings can be transferred to other settings as another aspect of a study's trustworthiness. An important mechanism for promoting transferability is the amount of information qualitative researchers provide about the contexts of their studies. Thick description a widely used term among qualitative researchers. It refers to a research and thorough description of the research setting and of observed transactions and processes. Quantitative researchers like qualitative researchers need to describe their study participants and their research settings thoroughly. So that the utility of the evidence for others can be assessed.

Cross Over

Randomization in the context of a cross over design is an especially powerful method of ensuring equivalence between

groups being compared. When subjects are exposed to two different conditions they may be influenced in the second condition by their experience in the first. Because treatment are not applied simultaneously the best approach is to use randomized ordering

Homogeneity

When randomization and cross over are not feasible, alternative methods of controlling extraneous characteristics should be used. One such method is to use only subjects who are homogeneous with respect to both use variables. (e. g.) Oka and Sunder (2005) studied the effect of exercise on body composition and nutritional intake in patients with heart failure. Several variables are controlled through homogeneity including gender (only men were included) age (no one was older than 76 years and illness severity (only those with stable class II to III heart failure were included.

Blocking /Stratification

Another approach to control extraneous variable is to include them in the research design through blocking or stratification. (eg) if gender were the confounding variable we could build it in to the study in a randomized block design. in which elderly men and women would be randomly assigned separately to treatment groups. This can enhance the likelihood of detecting differences between our experimental and control groups because w can eliminate the effect of the blocking variable on the dependent variable.

Statistical Control

Another method of control is through statistical analysis rather than research design. We use analysis of covariance to control the research. Analysis of covariance controls by statistically removing the effect of extraneous variables on the dependent variable.

(e. g.) The effect of age - - - - - effectiveness of physical training programme on heart functioning.

Type I and Type II Error

Researchers decide whether to accept or reject a null hypothesis by determining how probable it is that observed results are due to chance. Researchers lack information about the population and so cannot know with certainty whether a null hypothesis is or is not true. Researchers can only conclude that hypothesis are probably true or probably false and there is always a risk for error.

The actual situation is that the null hypothesis is

		TRUE	FALSE
The Researchers calculates a test statistic and decides that the null hypothesis is	True (Null Accepted)	Correct decision	Type II Error (False negative)
	False (Null rejected)	Type I Error (False positive)	Correct decision

Researchers can make 2 types of Statistical error; rejecting a true null hypothesis or accepting a false null hypothesis. Researchers make a type I error by rejecting a null hypothesis that is in fact true.

For instance, if we concluded that a drug was more effective than a placebo in reducing cholesterol, when in fact the observed differences in cholesterol levels resulted from sampling fluctuation, we would be making a type I error –a false positive conclusion. Conversely if we concluded that group difference in cholesterol resulted by chance, when the drug did reduce cholesterol we would be committing a type II error – a false negative conclusion.

A type I error might allow an ineffective drug from coming in to the market.

Researchers control the risk for a Type I error by selecting a level of significance, which signifies the probability of incorrectly rejecting a true null hypothesis.

The two most frequently used significance levels (referred to as alpha or α) are 0.05 and.01

Lowering the risk for type I error increases risk for type II error.

The simplest way of reducing the risk for a type II error is to increase sample size.

Sample Size Determination

Power Analysis and Effect Size

Power analysis can be used to reduce the risk for Type II error by estimating in advance how big a sample is needed, There are four components in a power analysis, three of which must be known/estimated; power analysis solves for the fourth he four components are

1) The significance criterion α (probability of committing Type I error) as the level of significance.
2) The sample size (N), as sample size increases power increases.
3) The effect size (Es) is an estimate of how wrong the null hypothesis is that is how strong the relationship between independent variable and dependent variable is in population.
4) Power or1 - β. This is the probability of rejecting null hypothesis.

Researcher usually establish the risk for a type I error (α) as 0.05 and the conventional standard for 1 - β is 0.80. With power equal to 0.80 there is a 20% of risk for committing type II error.

The effect size must be estimated using available evidence. Some times from a pilot study. More often an effect size is calculated based on findings from earlier studies on a similar problem.

Es is often designated as Cohen's d, the formula

$$d = \frac{\mu_1 - \mu_2}{\sigma}$$

d is the difference between two population mean, divided by the population standard deviation (eg) Cranberry juice reduces urinary pH of diet controlled patients
α=0.05 & power = 0.80
Mean of control group μ_1= *5.70*
Mean of Experimental groupμ_2= *5.50*
SD = *0.50*

$$d= \frac{5.70-5.50}{0.50} = 0.40$$

From table Estimated n=98. . We find that estimated n to detect n to detect an effect size of 0.40 with power equal to 0.80 is 98 subjects. Total number of subjects needed in new study would be about 200 with half assigned to control and experimental group.

Formula for determining Sample size

$$n= (\frac{z\sigma}{d})^2$$

n= sample size
z= value at level of confidence
=standard deviation between population mean& sample mean (or) standard error of mean.
Value of Z at 1% level of confidence =2.58
Value of Z at 5% level of confidence =1.96
STD deviation=7
Population mean=50
Sample mean =45 and the desired degree of precision is 99%.

$$n = \left(\frac{2.58\ X\ 7}{50-45}\right)^2$$

$(3.612)^2$
=13.04

Therefore, the size of the sample is 13.

2. Importance of Statistic in Nursing

- Nurses can use the knowledge of statistics for seeing or forecasting the patterns of symptoms caused in the patients so that they can be well prepared for coping with the uncertain situations.
- The knowledge of statistics can prove to be highly beneficial for the people related to medical field. Nurses can take advantage from the previous data archives to study the patterns of diseases in the patients. They can use data sheets or frequency charts for remembering the timings of medication given to the patients. They can maintain the temporary record for the patient under her supervision who is admitted in the hospital. Since there are so many events which have reported death of patients because of small carelessness of nurses, it is very necessary for them to make use of the statistical techniques and data for the patients' dosages, drugs uses and side effects etc.
- Statistics plays a vital role in almost every field of study one way or another. It can be used in different applications. As in the field of nursing it can be used to calculate the arithmetic mean of the number of people who get examined each day and in the same manner the record can be maintained for a week, mint or even a year.
- Similarly data calculations can be used to have an analysis that how many patients suffer a particular kind of disease like HIV etc. and this data can be used further by the hospitals to take more measure against it. A percentage calculated by the nurses can be useful in this regard. Similar techniques can be used by nurses.
- Helps in mapping health problems in the community
- Helps in deciding about health care used in hospitals, health centres and nursing homes.
- Helps to case comparison of related events, disease and other things related to health and illness.
- Helps to check efficacy of health programmes at national, localand global levels.
- Helps to assess the healthcare utilization of government and private health organisations.
- Helps in planning, implementing and evaluating health services at all levels.

References

[1] Polit, Denin. F. And Cheryl Tatano Beck (2008), Nursing research principles and methods, 8th edition, Philadelpia, Lippincott publishers
[2] SundarRao. P. S. S and Richard (2007), Introduction To Biostatistics and research methods, 4th edition, NewDelhi, PHI Publication
[3] BT Basvanthappa (2010), "Nursing Research" 2nd edition, New Delhi, Jaypee bothers.
[4] Sharma, Suresh. K (2010), Nursing Research and statistics, 1st edition, Haryana, Elsevier publications
[5] Frederick. J. Grvetter, Larry B. Wallnau (1995), Essentials of statistics for the behavioural sciences, 2nd edition, New york, West Publications.