PriMera Scientific
Medicine and Public Health
Volume 2 Issue 2 February 2023
ISSN: 2833-5627

Choosing Appropriate Probability Sampling Designs in Research

Type: Review Article
Received: December 06, 2022
Published: January 28, 2023

Citation:
Lok Raj Sharma. "Choosing Appropriate Probability Sampling Designs in Research". PriMera Scientific Medicine and Public Health 2.2 (2023): 04-21.

Lok Raj Sharma*

Associate Professor of English, Makawanpur Multiple Campus, Hetauda, Nepal

***Corresponding Author:** Lok Raj Sharma, Associate Professor of English, Makawanpur Multiple Campus, Hetauda, Nepal.

Abstract

Sampling design is one of the most recurrently conversed and discussed topics in research studies. Which sampling design is preferable and suitable in research primarily depends on the characteristics of a population through which we form a sample for the study. A sampling design is distinct plan for attaining a sample from the given population. The principal objective of this article is to choose the appropriate probability sampling designs in research studies. The present study is qualitative in nature and the data are collected from the secondary sources. This article will no doubt be beneficial to the persons who are engrossed in research studies as it provides them with a clear concept of types of the probability sampling designs so that they can cull an apt probability sampling design to carry out the research studies.

Keywords: Probability; research; sample; sampling design

Introduction

Doing researches is one of the most esteemed scholarly tasks in this era. Research is a systematic process of collecting and analyzing information (Nanda & Khatoi, 2005) to increase our understanding of the phenomenon and to answer a particular question or problem under study. Best and Kahn (2010) consider research as really a cerebral and creative activity. It is disciplined inquiry (Dornyei, 2007). Research writing is typically an academic writing (Kothari & Garg, 2019) that must be systematic, consistent and accurate. Creswell (2012) takes research as a process of steps used to amass and analyze information to enhance our understanding of a topic or issue. It is the systematic investigation of a topic (Vogt, Gardner, & Haeffele, 2012). The step-wise process of research includes: formulating a research problem, conceptualizing a research design, constructing an instrument for collecting the data, selecting a sample through a sampling design, collecting data, interpretation and analysis of data and writing a research report. Design, which is the planning of strategy, explains how a researcher intends to conduct the research study. A sampling design is an essential aspect in the quantitative research study. A sampling design is an explicit plan for obtaining a sample from a given population. It is necessary for a researcher to make a careful selection of just a few elements from the population. There are primarily two sorts of sampling design: probability sampling design and non-probability sampling design. Probability sampling is generally accepted as the most appropriate method for making inference that can be generalized to a finite population (Brick, 2014). With a probability sample, every unit in the population has a known, non-zero chance of being sampled, and in the design-based framework these probabilities are the basis for the inferences (Hansen, Hurwitz, & Madow, 1953; Särndal, Swensson, & Wretman, 1992; Lohr, 2009).

The article writer has attempted to focus only on the probability sampling design by going through some books and journal articles written on this topic. Proper selection of sampling design is must for forming a sample size from which a researcher can draw conclusions to represent the population. Sampling is a highly talked and discussed topic in the circle of research scholars. Such discussions drove this writer to write this article on the probability sampling design in which simple random sampling, systematic sampling, stratified sampling, cluster sampling and multi-stage sampling designs have been mentioned. It is hoped that not only students but research scholars will also be benefitted from going through this journal article.

Literature Review

Sampling Procedures

Sampling procedures are the ways how researchers select the representative elements from the population. We should have the basic concepts of the following terms to study the sampling procedures:

Population / Universe: Population or universe is the entire collection of all items, elements, persons, or units of the interest for the research. Choosing the population depends mostly on the researcher's interests on the basis of information about cases that are relevant to the study (Henry, 2009) . We can name population in different ways:

Finite population: The finite population is defined as the population of all the individuals or objects that can be counted. The population in which number of units is finite and can be counted precisely is called finite population. For example, all the employees of a company.

Infinite population: The infinite population is also known as an uncountable population in which the counting of units in the population is not possible. For example, stars in space.

Existent population: The existing population is defined as the population of concrete individuals. If the units of population have physical existence, it is called existent population. For example, books, tables etc.

Hypothetical population: The population in which whose unit is not available in solid form is known as the hypothetical population. A population consists of sets of observations, objects etc that are all something in common. The population, units of which do not have physical existence but their existence is assumed or probability of their existence is found by statistical method is called hypothetical population. For example, the outcome of tossing a coin.

Homogeneous population: If all the units of population are identical or similar in terms of certain characteristic/s, it is called homogeneous population. For example, people of the Hindu religion living in a certain place.

Heterogeneous population: If all the units of population differ completely or in some aspects with one another, the population is called heterogeneous population. For example, people of the Hindu, Buddhist, Christian and Muslim religions living in a certain place.

The population, on the basis of characteristics, could be divided in to three types:

Univariate population: Population in which only one characteristic is considered, for studying at a time. For example, the characteristic may be age, income, sex, Television watching habit. listening habit, etc.

Bi-variate population: The population can be defined as a bi-variate type when we are measuring two characteristics simultaneously of each member. For example, sex and income of the population.

Multivariate population: A multivariate universe is the one in which we consider observations on three or more characteristics simultaneously. For example, sex, income, age etc. of the population.

Sample: Sample is a collection of representative items or elements from the population. It is a subset of the population. It is a small proportion of the population that is selected for observation and analysis (Best & Kahn, 2006). If a census is not needed, or not practical to carry out, a sample is the most appropriate (Kolb, 2011).

Sampling frame: A sampling frame contains the names of all items of the population or universe. It is the list of elements from which the sample is drawn (Zikmund, J. Babin, Carr, Adhikary, & Griff, 2010; Kumar, 2011). It is a complete listing of all the elements that make up a research population (Ruane, 2005).

Sampling unit: Sampling unit is each item from a sample. It is the unit that represents every character of population. It is a single or groups of elements subject to the selection in the sample (Zikmund, J. Babin, Carr, Adhikary, & Griff, 2010).

Margin of error: Margin of error is the degree of error in results received from random sampling surveys.

Sampling error: A sampling error is a statistical error that occurs when a researcher does not select a sample that represents the entire population. It is the difference between the sample estimate and the true population score (Cresswell, 2019).

Sample size: Sample size refers to the number or amount of items which is to be selected from the population. Samples should be as large as a researcher can obtain with a reasonable expenditure of time and energy (Fraenkel & Wallen, 1990).

The following formula can be used to calculate the sample size:

Sample size $(n) = \frac{N}{1+Ne^2}$

Where,

n = Required sample size
N= Population size
Confidence level: 95%
e = Margin of error
Source: (Yamane, 1967, p.886)

Sampling: The process of selecting a number of individuals from a population is sampling (Fraenkel & Wallen, 1990). It is the process of selecting certain units or items from the population. It is the selection of a subset of the population of interest in a research study. Sampling is a procedure, where in a fraction of the data is taken from a large set of data, and the inference drawn from the sample is extended to whole group (Raj, 1972). It can be used to make inference about a population or to make generalization in relation to existing theory. The process of selecting sample from population is called sampling. Sampling process takes place in the following way:

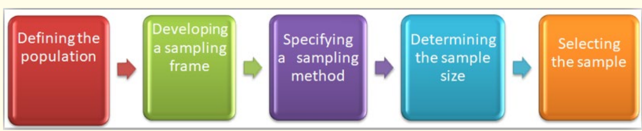
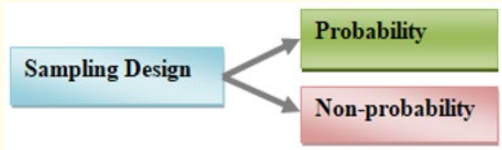


Figure 1: Sampling process.

Sampling Designs

Sampling design is a definite plan for obtaining a sample from the given population. It is a mathematical function that gives you the probability of any given sample being drawn. It refers to the techniques or the procedures the researcher adopts in selecting items for the sample. The researcher has to make cautious selection of the items or elements from the population, and study them deeply to reach the conclusion which can be applied or generalized to the entire population. Careful planning is essential in recruiting and accessing participants in the research process (Guthrie, 2010). Choosing a sampling technique depends greatly on the goal, and type of the research, what (Cohen, Manion, & Morrison, 2018). There are primarily two types of sampling design. They are: probability sampling design and non- probability sampling designs. In the probability sampling design, every item in the population has an equal chance of being included in a sample. This design is employed in a quantitative research, whereas a non-probability sampling is often associated with a qualitative research.



Probability Sampling Design

Probability sampling design is applicable and suitable in a quantitative research. The conclusion drawn from an intense study and analysis of a sample can be generalizable to the total population if the sample is formed through probability sampling design. Probability or random sampling has the greatest freedom from bias but may represent the most costly sample in terms of time and energy for a given level of sampling error (Brown, 1947). In probability sampling, each unit is drawn with known probability, (Yamane, 1967) or has a nonzero chance of being selected in the sample (Raj, 1972). Such samples are usually selected with the help of random numbers (Cochran, 1963). The major benefit of using random sampling is the liberty from human judgement bias and subjectivity (Taherdoost, 2016). Probability samples use some form of random sampling in one or more of their stages (Kerlinger, 2011).

Characteristics of Probability Sampling

The characteristics of probability sampling are as follows:

- Selection of subjects in an objective and unbiased way
- Each unit of population has certain probability to be selected in a sample.
- Researcher can select a sample by keeping in mind the size of the sample by applying a suitable method of probability sampling.
- A researcher's personal wish does not affect the selection of a certain subject.
- Every subject is selected independently.
- Selection of one subject does not affect the selection of another subject.
- It increases the possibility of selecting such a sample that represents the population completely. It is very helpful in determining sampling error.

There are five types of probability sampling designs: simple random sampling, systematic sampling, stratified sampling, cluster sampling and multi-stage sampling designs.

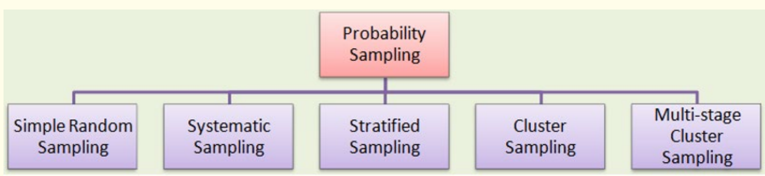


Figure 2: Probability sampling types.

Simple Random Sampling

Simple random sampling is the most recognized probability sampling design. It is a probability sampling design in which every element / unit has an equal and independent chance of being selected from the population of interest. It provides every element of the population with an equal probability of inclusion in sample. Disadvantages associated with simple random sampling include are necessity of complete frame, high standard errors of estimators and the high cost of obtaining a sample in case geographically scattered elements (Ghauri & Gronhaug, 2005). In a simple random sample, every member of the population has an equal chance of being selected. Your sampling frame should include the whole population (McCombes, 2022).

There are two types of simple random sampling

1. Simple random sampling with replacement (SRSWR): Simple random sampling with replacement is selecting “n” number of units out of “N” units one by one in such a way that at each stage of selection, the sample each unit has equal chance of being selected, i.e., 1/N.
2. Simple random sampling without replacement (SRSWOR): Simple random sampling without replacement is selecting “n” number of units out of “N” one by one at any stage of selecting a sample in such a way that anyone of the left units have the probability of being selected as a sample, i.e., 1/N.

Simple random sampling is chosen if

- The population is homogeneous.
- The whole population is available.
- The population contains a finite number of participants.
- Every participant needs to be studied exclusively.
- There is easy access to a complete and accurate sampling frame of the target population.
- It does not contain auxiliary information that may be used for the purposes of stratification.

Benefits of using the simple sampling design

- It does not have the sampling biasness.
- It is a good representative of the population.

Drawbacks of the simple random sampling

- It requires a lot of efforts especially for a large population.
- It is difficult to prepare an exhaustive list of participants.
- It is costly and time consuming.

Steps in selecting a simple random sample

1. Defining the target population
2. Identifying an existing sample frame of the target population
3. Evaluating the sampling frame
4. Assigning a unique number to each and every participant clearly.
5. Determining the sample size
6. Selecting the required participants randomly

Three techniques are typically used in selecting the required participants randomly: the lottery method, a table of random numbers, and randomly generated numbers using a computer program.

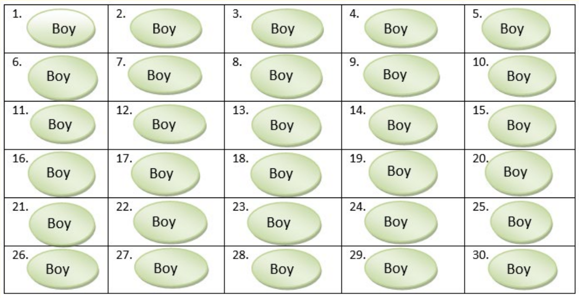


Figure 3: Specimen of homogeneous population.

If we need 5 boys for our sample, we can select any five boys through a lottery method from the list of 30 boys.

Systematic random sampling

Systematic random sampling is a method of selecting samples at a particular preset interval by selecting a random number starting point between 1 and the sampling interval. Systematic sampling is similar to simple random sampling, but it is usually slightly easier to conduct. Every member of the population is listed with a number, but instead of randomly generating numbers, individuals are chosen at regular intervals (McCombes, 2022). It involves selecting every nth member of the target group (Dornyei, 2007).

Systematic sampling should be used if

- The population is homogeneous.
- The whole population is available.
- The population contains a finite number of participants.
- Every participant needs to be studied exclusively.
- It is difficult to identify items using a simple random sampling method.
- A sampling frame is not available or impractical to organize.
- A stream of representative elements of the population is available.
- The listing of the population is essentially random or can be randomized.

Benefits

- It avoids biasness.
- Less costly and less time taking than the simple random sampling.

Drawbacks

- All the participants existing before the selected number are ignored.
- All the participants between the regular intervals are left.

The process involves

- Calculation of fix the sampling interval with dividing the population by the number of required sample size.
- Choice of a random starting point between 1 and the sampling interval.
- Repetition of the sampling interval to choose subsequent elements.

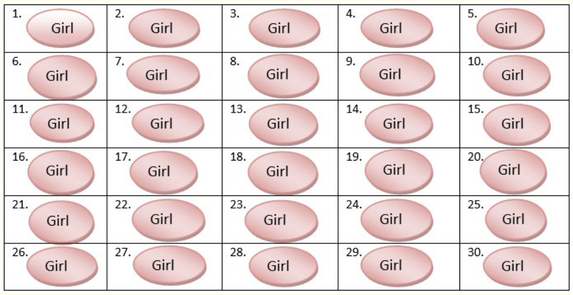


Figure 4: Specimen of homogenous population.

Selection of the sample size takes place in the following way:

Example-1

Population (N) = 30
Suppose, Required Sample Size (n) = 5
Interval (k) = N/n = 30/5 = 6
Suppose the number 4 is chosen via the lottery method from (1-6).
Random number (r) =4
Now, the selection takes place in this way:
The first number = r = 4
The second number = 2r = 2x4= 8
The third number = 3r = 3x4= 12
The fourth number= 4r = 4x4= 16
The fifth number= 5r = 5x4= 20
It means we should select 5 girls who are represented by the selected numbers.

Example-2

Population (N) = 30
Suppose, Required Sample Size (n) = 8
Interval (k) = N/n = 30/8 = 3.74 = 4
Suppose the number 4 is chosen via the lottery method from (1-4).
Random number (r) =4
Now, the selection takes place in this way:
The first number = r = 4
The second number = 2r = 2x4= 8
The third number = 3r = 3x4= 12
The fourth number = 4r = 4x4= 16
The fifth number = 5r = 5x4= 20
Sixth number = 6r = 6x4=24
Seventh number = 7r = 7x4= 28
Eighth number = 8r = 8x4= 32 (32 is beyond the limit).
We should not round up if the product of n and k is greater than N.
Selection will be done without rounding up. K = 3

Suppose the number 3 is chosen via the lottery method from (1-3).
The first number = r = 3
The second number = 2r = 2x3= 6
The third number = 3r = 3x3= 9
The fourth number = 4r = 4x3= 12
The fifth number = 5r = 5x3= 15
Sixth number = 6r = 6x3=18
Seventh number = 7r = 7x3= 21
Eighth number = 8r = 8x3= 24
It means we should select 8 girls who are represented by the selected numbers.

Linear systematic sampling

Linear systematic sampling is a systematic sampling method where samples aren't repeated at the end and 'n' units are selected to be a part of a sample from the population units. This process involves the following process:

- Arrangement of the entire population in a classified sequence (N).
- Determination of the sample size (n).
- Calculation of the sampling interval (k) = N/n.
- Selection of a random number between 1 to k (including k).
- Add the sampling interval (k) to the chosen random number. Then move ahead.

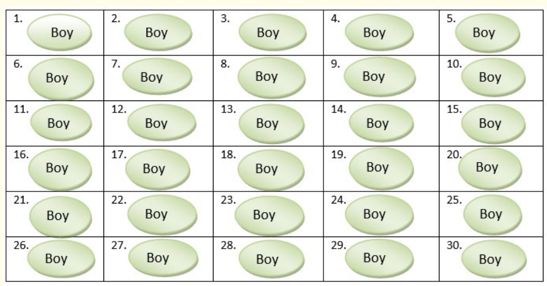


Figure 5: Specimen of homogeneous population.

Example-1

Population (N) = 30
Suppose, Required Sample Size (n) = 5
Interval (k) = N/n = 30/5 = 6
Suppose the number 4 is chosen via the lottery method from (1-6).
Random number (r) =4
Now, the selection takes place in this way:
The first number = r = 4
The second number = r +k = 4+6 =10
The third number = r +2k= 4+ 2x6= 4+12= 16
The fourth number= r +3k= 4+ 3x6= 4+18 = 22
The fifth number= r + 4k= 4+ 4x6 = 4+24=28
It means we should select 5 boys who are represented by the selected numbers.

Example-2

Population (N) = 30
Suppose, Required Sample Size (n) = 8
Interval (k) = N/n = 30/8 = 3.74 = 4
Suppose the number 4 is chosen via the lottery method from (1-4).
Random number (r) =4
Now, the selection takes place in this way:
The first number = r = 4
The second number = r +k = 4+4= 8
The third number = r + 2k = 4+ 2x4= 4+8= 12
The fourth number = r + 3k = 4+ 3x 4= 4+12= 16
The fifth number = r + 4k = 4+ 4x4= 4+16= 20
Sixth number = r+ 5k = 4 + 5x4=4+ 20 = 24
Seventh number = r + 6k = 4+ 6x4= 4+24 = 28
Eighth number = r + 7k = 4 + 7x4= 4+ 28 = 32 (32 is beyond the limit).
We should not round up if the product of n and k is greater than N.
Selection will be done without rounding up. K = 3
Suppose the number 3 is chosen via the lottery method from (1-3).
The first number = r = 3
The second number = r +k = 3+3= 6
The third number = r + 2k = 3+ 2x3= 3+6 = 9
The fourth number = r + 3k = 3+ 3x 3= 3+9= 12
The fifth number = r + 4k = 3+ 4x3= 3+12= 15
Sixth number = r+ 5k = 3 + 5x3=3+ 15= 18
Seventh number = r + 6k = 3+ 6x3= 3+18 = 21
Eighth number = r + 7k = 3+ 7x3= 3+ 21 = 24
It means we should select 8 boys who are represented by the selected numbers.

Circular systematic sampling

In circular systematic sampling, a sample starts again from the same point once again after ending. This sampling involves:

- Calculation of sampling interval (k) = N/n.
- Selection of the randomly number between 1 to N
- Creation of the samples by skipping through k units every time until we select members of the entire population.

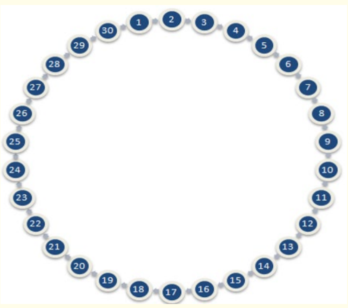


Figure 6: Circular systematic sampling.

Example-1

Population (N) = 30
Required Sample Size (n) = 6
Sampling Interval (k) = N/n= 30/6= 5
Suppose the random number selected through the lottery between 1 and N (may also include 1 or 30)
Suppose, (r) = 20
The first number = r = 20
The second number = r +k = 20+5 = 25
The third number = r + 2k = 20+ 2x5= 20+10 = 30
The fourth number = r + 3k = 20+ 3x 5= 20+15= 35 = 35-30 = 5
The fifth number = r + 4k = 20+ 4x5= 20+20= 40-30= 10
Sixth number = r+ 5k = 20 + 5x5=20+ 25= 45-30= 15
We can select 6 boys who are represented by the numbers.

Example-2

Population (N) = 30
Required Sample Size (n) = 11
Sampling Interval (k) = N/n= 30/11= 2.72= 3
Suppose the random number selected through the lottery between 1 and 30 (may include 1 or N) is (r) = 3
Suppose the random number selected through the lottery between 1 and N (may also include 1 or 30)
Suppose, (r) = 3
The first number = r = 3
The second number = r +k = 3+3= 6
The third number = r + 2k = 3+ 2x3= 3+6 = 9
The fourth number = r + 3k = 3+ 3x 3= 3+9= 12
The fifth number = r + 4k = 3+ 4x3= 3+ 12= 15
Sixth number = r+ 5k = 3 + 5x3=3+ 15= 18
Seventh number = r+ 6k = 3 + 6x3=3+ 18= 21
Eighth number = r+ 7k = 3 + 7x3=3+ 21= 24
Ninth number = r+ 8k = 3 + 8x3=3+ 24= 27
Tenth number = r+ 9k = 3 + 9x3=3+ 27= 30
Eleventh number = r+ 10k = 3 + 10x3=3+ 30= 33-N= 33-30=3 (Repeated)

In case of the repetition of the number, it is better not to round up or it is better to select the number next to the repeated number. Here, the number 4 can be selected.

Stratified Random Sampling

A stratified sampling is a probability sampling procedure in which the target population is first separated into mutually exclusive homogeneous segments (strata), and then elements are selected through a simple random sample from each segment (stratum). It is frequently used when there are variations in the traits of participants within a population. It is a combination of randomization and categorization (Dornyei, 2007). Its purpose is to ensure that every stratum is adequately represented (Ackoff, 1953). Stratified sampling involves dividing the population into subpopulations that may differ in important ways. It allows you draw more precise conclusions by ensuring that every subgroup is properly represented in the sample (McCombes, 2022).

Stratified sampling requires to be implemented if

- The population is heterogeneous.
- It is possible to divide a population into two or more homogeneous strata and construct a sampling frame for each stratum.
- There is access to a sampling frame of the target population that is complete and accurate and contains auxiliary information that may be used for stratification purposes.
- Some subgroups of the population are vastly different from other subgroups.
- It is very important to minimize sampling error.
- There is a concern about under representing smaller subgroups.
- There is a desire to use different selection methods for different strata.
- It is likely that answers to the research questions of a study are likely to be different for different subgroups.
- It is useful when each stratum needs to be reported separately.
- Comparative analysis of strata is desired.
- The whole population is available.
- The population contains a finite number of participants.
- Every stratum needs to be studied exclusively.

Benefit

- It produces a representative sample for a heterogeneous population.

Drawbacks

- It requires a lot of efforts for sampling.
- It is costly and time consuming.

Steps in selecting a stratified random sample

- Defining the target population
- Identifying stratification variable(s) and determine the number of strata to be used.
- Identifying an existing sampling frame
- Evaluating the sampling frame
- Dividing the sampling frame into strata
- Assigning a unique number to each element
- Determining the sample size for each stratum
- Selecting the targeted number of elements from each stratum randomly.

Proportionate stratified sampling is culled if:

- Sub-groups of approximately the same size are to be investigated or compared.

Disproportionate stratified sampling needs to be executed if

- Subgroups of vastly different sizes are to be investigated or compared.
- It is important to include a large number of elements from a small segment of the population.
- One is primarily interested in key similarities and differences among strata.
- Some observations are limited or hard to obtain.
- It is important to make statistically valid statements about subgroups.
- Subgroups of the population have different variances for the variables of interest.

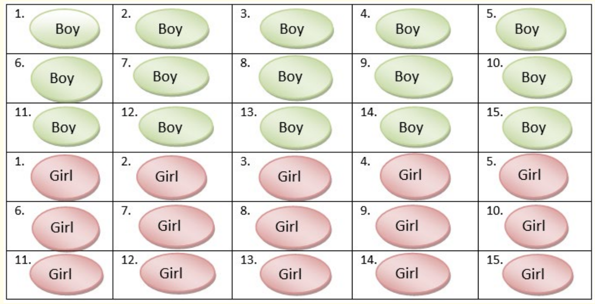


Figure 7: Specimen of heterogeneous population with two strata.

Example-1

Total Population (N) =30
Required sample size (n) = 6
We have,
Persons in the first stratum (S_1) = 15
Persons in the second stratum (S_2) = 15

Using Proportional Allocation

Required sample size from each Stratum:

$$\text{From}(S_1), (n_1) = \frac{S_1 \times n}{N} = \frac{15 \times 6}{30} = 3$$

$$\text{From}(S_2), (n_2) = \frac{S_2 \times n}{N} = \frac{15 \times 6}{30} = 3$$

Now selecting the required sample size from each stratum through SRT (Simple Random Technique).

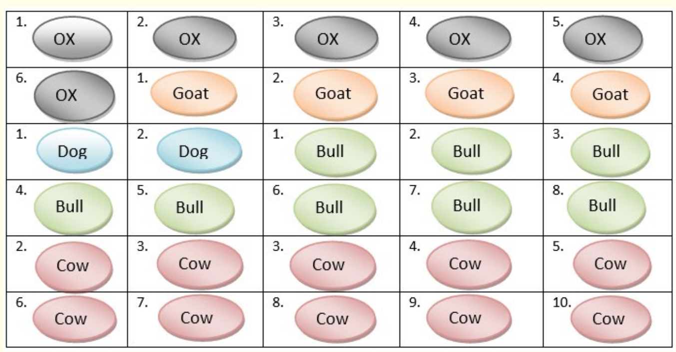


Figure 8: Specimen of heterogeneous population with five strata.

Total Population (N) =30

Suppose, the required sample size (n) = 12

We have,

Oxen in the first stratum (S_1) = 6

Goats in the second stratum (S_2) = 4

Dogs in the third stratum (S_3) = 2

Bulls in the fourth stratum (S_4) = 8

Cows in the fifth stratum (S_5) = 10

Using Proportional Allocation

Required sample size from each Stratum:

$$\text{From}(S_1),\ (n_1) = \frac{S_1 \times n}{N} = \frac{6\times12}{30} = 2.4 = 2$$

$$\text{From}(S_2),\ (n_2) = \frac{S_2 \times n}{N} = \frac{4\times12}{30} = 1.6 = 2$$

$$\text{From}(S_3),\ (n_3) = \frac{S_3 \times n}{N} = \frac{2\times12}{30} = 0.8 = 1$$

$$\text{From}(S_4),\ (n_4) = \frac{S_4 \times n}{N} = \frac{8\times12}{30} = 3.2 = 3$$

$$\text{From}(S_5),\ (n_5) = \frac{S_5 \times n}{N} = \frac{10\times12}{30} = 4$$

Now selecting the required sample size from each stratum through SRT (Simple Random Technique).

There are basically two types of stratified sampling designs. They are: proportionate and disproportionate sampling:

In proportionate stratified sampling, the number of elements allocated to the various strata is proportional to the representation of the strata in the target population.

Suppose that the population of the students from four sections was 327 and the required sample size was 180. The following tables can be used to illustrate proportionate and disproportionate stratified samples.

Sections as strata	***Population***		***Proportionate Stratified Sample***	
	Frequency	***Percent***	***Frequency***	***Percent***
A	84	26%	47	26%
B	88	27%	49	27%
C	80	24%	43	24%
D	75	23%	41	23%
Total	327	100%	180	100%

Table 1: Proportionate stratified sample.

Disproportionate stratified sampling is a stratified sampling procedure in which the number of elements sampled from each stratum is not proportional to their representation in the total population. Disproportionate stratified sampling can be determined by using unequal allocation. In this sampling, sufficient numbers of elements must be selected for each category.

Section	Population		Disproportionate Stratified Sample Using unequal Allocation	
	Frequency	Percent	Frequency	Percent
A	84	26%	52	29%
B	88	27%	50	28%
C	80	24%	47	26%
D	75	23%	31	17%
Total	327	100%	180	100%

Table 2: Disproportionate stratified sample.

Disproportionate stratified sampling can also be determined by using equal allocation. In this sampling, the equal number of elements is selected for each category.

Section	Population		Disproportionate Stratified Sample Using Equal Allocation	
	Frequency	Percent	Frequency	Percent
A	84	26%	45	25%
B	88	27%	45	25%
C	80	24%	45	25%
D	75	23%	45	25%
Total	327	100%	180	100%

Table 3: Disproportionate stratified sample.

Cluster Sampling

A cluster sampling is a probability sampling procedure in which elements of the population are randomly selected in naturally occurring groupings (clusters). A random sample is taken from these clusters, all of which are used in the final sample (Wilson, 2010). Cluster sampling is advantageous for those researchers whose subjects are fragmented over large geographical areas as it saves time and money (Davis, 2005). In cluster sampling, diverse parts of a population are taken as clusters, and persons from each cluster are selected (Bhardwaj, 2019). Cluster sampling also involves dividing the population into subgroups, but each subgroup should have similar characteristics to the whole sample. Instead of sampling individuals from each subgroup, the researchers randomly select entire subgroups (McCombes, 2022). It is used when the target population is extensively dispersed (Dornyei, 2007).

Cluster sampling is considered appropriate if

- The population is homogeneous.
- The participants of population are spread over a wide geographical area.
- The clusters are possible.
- Each and every cluster represents the population
- A sampling frame of individual population elements is not available but a sampling frame of clusters of elements is available.
- It is important to minimize data collection costs and there are substantial fixed costs associated with each data collection location.

Benefit

- It produces a representative sample for a homogeneous population spread over a large geographical area.
- It is less Costly and less time consuming.

Drawback

- If the clusters are not homogeneous among them, the sample may not be representative of the population.

Steps in selecting a cluster sample

- Defining the target population.
- Determining the desired sample size.
- Identifying an existing sampling frame or develop a new. .sampling frame of clusters of the target population.
- Evaluating the sampling frame.
- Determining the number of clusters to be selected.
- Selecting the targeted number of clusters Randomly.

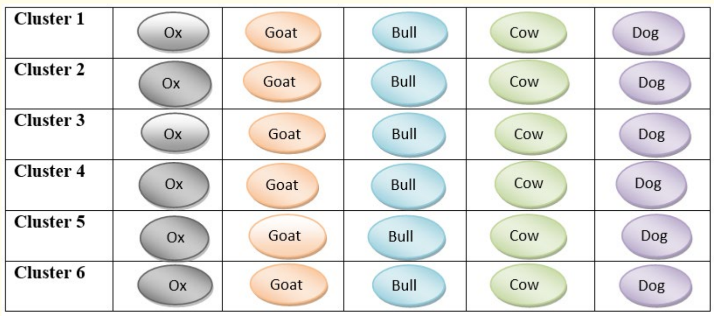


Figure 9: Specimen of homogeneous population among clusters.

There is the homogeneity of elements among the clusters. Here, every cluster has an Ox, a Goat, a Bull, a Cow and a Dog, but the homogeneity of elements is not found within each cluster, because there are different animals in each cluster. Each cluster is a representative of the population. In this sampling technique, the researcher selects the clusters through the simple random technique, and then all the elements (animals) are taken from the selected clusters under study.

Multi-stage Sampling

A multi-stage sampling is a probability sampling technique wherein he sampling is carried out at several stages such that the sample size gets reduced at each stage. Multi-stage sampling is a process of moving from a broad to a narrow sample, using a step by step process (Ackoff, 1953). Multistage sampling acts like a funnel where a broad population sample is narrowed down until it becomes the final unit for the investigation (Alvi, 2016).

Multi-stage sampling is put into practice if

- There is a need to identify and collect information from a subgroup of the population that is difficult to collect from the first stage of data collection.
- The population is so widely spread over a huge geographical area.
- The first four techniques of sampling are very difficult or impossible.
- The participants of population are spread over a wide geographical area.
- The huge geographical area can be divided into small areas as its parts.
- The parts can further be divided into small areas as their parts.

Benefits

- It produces a representative sample for a homogeneous population spread over a large geographical area.
- It saves our time and effort.

Drawbacks

- The selected clusters may not capture the characteristic diversity of the population.
- If the clusters are not homogeneous among them, the sample may not be representative of the population.

Steps in selecting a multi-stage sample

- Dividing the target population into clusters.
- Assigning a unique number to each and every cluster clearly.
- Selecting the required numbers of clusters randomly (First stage)
- The clusters are homogeneous among them, but may be heterogeneous inside.
- The selected clusters are further divided into small clusters.
- Selecting small clusters randomly (Second stage).
- The selected clusters are further divided into small clusters.
- The researcher selects the required number of clusters (Third stage).
- Investigating all the participants from the finally selected clusters.

Method and Materials

This article is grounded on the qualitative research in which terms related to sampling designs, their definitions and instances have been regarded as qualitative data. The qualitative data have been taken from the related books, journal articles and research articles. Such data have been described and analyzed by focusing their merits, demerits and the proper condition of their applications.

Analysis of Probability Sampling Designs

Simple random sampling provides representative samples, and permits the execution of inferential statistics to calculate margin of errors. However, it tends to have bigger sampling errors and less precision than stratified samples of the same sample size. If the target population is widely isolated, data collection costs might be higher for simple random sampling than those for other probability sample designs, such as cluster sampling. Stratified sampling allows the estimation of population parameters and within-strata inferences and comparisons across strata, tends to be more representative of a population, takes advantage of knowledge the researcher has about the population, possibly makes for lower data collection costs, and allows the researcher to use diverse sampling procedures within the various strata. On the other hand, unlike unstratified sampling, stratified sampling entails prior information on the stratification variables and more complex analysis measures. Normally, systematic sampling is easier, simpler, less time-consuming, and more economical than simple random sampling. If the ordering is disparate to the study variables, but randomized, systematic sampling will give results similar to simple random sampling. On the other hand, periodicity in the sampling frame is a constant anxiety in systematic sampling. Some of the strengths of cluster sampling when compared to simple random sampling include requiring less time, money, and effort; and allowing subsequent sampling and the inference characteristics of clusters as well as the target population. However, cluster sampling when compared to simple random sampling may not be as representative of the population as a simple random sample of the same sample size, and variances of cluster sampling are probable to be higher than those for simple random sampling.

Conclusion

The process of selecting samples from a population is called sampling. A plan used to select a sample is called a sampling design. Depending on research goals, researchers can use specific sampling methods from a variety of methods. There are 5 probability sampling designs: simple random sampling, systematic sampling, stratified sampling, cluster sampling, and multilevel sampling designs.

Simple random sampling is a probability sampling technique that equalizes the chances of being selected for all items in the target population and all possible samples of a particular size. Systematic sampling is a probabilistic sampling technique that randomly selects the first item to sample, and then selects subsequent items using fixed or systematic intervals until the desired sample size is reached. Systematic sampling has three subtypes: random systematic sampling, linear systematic sampling, and circular systematic sampling. Stratified sampling is a probabilistic sampling technique that first divides the target population into mutually exclusive uniform segments and then draws a simple random sample from each segment (layer). There are two main subtypes of stratified sampling: proportionate sampling and disproportionate sampling. In proportional stratified sampling, the number of items assigned to different strata is proportional to the representation of the strata in the target population. In disproportionate stratified sampling, the number of items assigned to different strata is not proportional to the stratum representation of the target population. Cluster sampling is a probabilistic sampling technique in which population members are randomly selected in naturally occurring clusters. Subtypes of cluster sampling may be classified on the basis of the number of sampling events like single-stage cluster sampling, two-stage cluster sampling, and multistage cluster sampling. This article will unquestionably be fruitful to new researchers as it provides them with a clear idea of using appropriate sampling designs in their research studies.

References

1. Ackoff RL. "The design of social research". Chicago: University of Chicago Press (1953).
2. Alvi M. "A manual for selecting sampling techniques in research". Munich Personal RePEc Archive (2016).
3. Best JW and Kahn JV. "Research in education". Boston: Pearson Education Inc (2006).
4. Best J and Kahn J. Research in education (10th ed.). New Delhi: PHI Learning Pvt. Ltd (2010).
5. Bhardwaj P. "Types of sampling in research". Journal of the Practice of Cardiovascular Sciences 5.3 (2019): 157-163.
6. Brick JM. "Explorations in non-probability sampling using the web". Proceedings of Statistics Canada Symposium (2014).
7. Brown GH. "A comparison of sampling methods". Journal of Marketing 6 (1947): 331-337.
8. Cochran WG. Sampling techniques (2nd ed.). New York.: John Wiley & Sons, Inc (1963).
9. Cohen L, Manion L and Morrison K. Research methods in education (8th ed.). New York: Routledge Palmer (2018).
10. Cresswell JW. Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.). Noida: Pearson India Education Services Pvt. Ltd (2019).
11. Creswell J. Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.). Boston: Pearson Publications (2012).
12. Davis D. "Business research for decision making". Australia: Thomson South-Western (2005).
13. Dornyei Z. "Research methods in applied linguistics". Shanghai: Oxford University Press (2007).
14. Fraenkel JR and Wallen NE. "How to design and evaluate research in education (7th ed.)". New York: McGraw-Hill (1990).
15. Ghauri P and Gronhaug, K. "Research methods in business studies". Harlow: Prentice Hall (2005).
16. Guthrie G. "Sampling in basic research methods: An entry to social science research". Delhi: SAGE Publications India Pvt Ltd (2010).
17. Hansen MH, Hurwitz WN and Madow WG. "Sample survey methods and theory". New York: Wiley 1 (1953).
18. Henry GT. Practical sampling. In L. Bickman, & J. D. Rog (Eds.), The SAGE handbook of applied social research methods. Delhi: SAGE Publications Ltd (2009): 29.
19. Kerlinger FN. "Foundations of behavioral research". Delhi: Surjeet Publications (2011).
20. Kolb B. "Determining probability samples in marketing research". Delhi: SAGE Publications (2011).
21. Kothari CR and Garg G. Researh methodology: Methods and technology (4th ed.). New Delhi: New Age International (P) Ltd (2019).
22. Kumar R. "Research methodology: A steo-by- step guide for beginners". New Delhi: SAGE (2011).
23. Lohr S. "Sampling: design and analysis. Cengage Learning". Boston: Cengage Learning (2009).
24. McCombes S. Sampling methods, types, techniques and examples (2022).
25. Nanda G and Khatoi P. "Fundamentals of educational research and statistics". New Delhi: Kalyani Publishers. New Delhi: Kalyani

Publishers (2005).

26. Raj D. The design of sample surveys. New York: McGraw-Hill Book Company (1972).
27. Ruane JM. "Essentials of research methods: A guide to social science research". Maiden: Blackwell Publishing (2005).
28. Särndal C-E, Swensson B and Wretman J. "Model assisted survey sampling". New York: Springe (1992).
29. Taherdoost H. "Sampling methods in research methodology: How to choose a sampling technique for research". International Journal of Academic Research in Management 5.2 (2016): 1-10.
30. Taherdoost H. "Sampling methods in research methodology: How to choose a sampling technique for research". International Journal of Academic Research in Management 5 (2020): 18-27.
31. Vogt WP, Gardner DC and Haeffele LM. "When to use what research design". New York: The Guilford Press (2012).
32. Wilson J. "Essentials of business research: a guide to doing your research project". SAGE Publication (2010).
33. Yamane T. "Elementary Sampling theory". Englewood Cliffs: Prentice-Hall, Inc (1967).
34. Zikmund WG., et al. "Business research method: A south-Asian perspective". Delhi: Cengage Learning India Pvt. Ltd (2010).