Sampling methods

Preprint · September 2018

DOI: 10.13140/RG.2.2.22856.57605

1 author:

Sanjoy Datta
West Bengal University of Animal and Fishery Sciences
37 PUBLICATIONS **28** CITATIONS

SEE PROFILE

Concept of sampling methods and different types of sampling

Introduction to Sampling

The way in which we select a sample of individuals to be research participants is critical. How we select participants (random sampling) will determine the population to which we may generalize our research findings. The procedure that we use for assigning participants to different treatment conditions (random assignment) will determine whether bias exists in our treatment groups (Are the groups equal on all known and unknown factors?).

If we do a poor job at the sampling stage of the research process, the integrity of the entire project is at risk. If we are interested in the effect of TV violence on children, which children are we going to observe? Where do they come from? How many? How will they be selected? These are important questions.

Distinguishing Between a Sample and a Population

Before describing sampling procedures, we need to define a few key terms. The term population means all members that meet a set of specifications or a specified criterion. For example, the population of the United States is defined as all people residing in the United States. A single member of any given population is referred to as an element. When only some elements are selected from a population, we refer to that as a sample.

Types of sampling

1) Probability sampling methods.
2) Non-probability sampling methods.

Probability Sampling Methods

> Probability sampling is also called as random sampling or representative sampling.
> In probability sampling every member of the population has a known (non zero) probability of being included in the sample.
> The probabilities can be assigned to each unit of the population objectively.
> These techniques need population to be very precisely defined.
> These techniques cannot be used for the population that is too general a category found almost everywhere in the world.
> For instance if our target population is defined as college students. It means person studying at any college of the world is an element of our population.
> In this case probability sampling can be done as the population is precisely defined and limited to an infinite number of elements.

Advantages:

- This sampling technique reduces the chance of systematic errors.
- The methods minimize the chance of sampling biases.
- A better representative sample is produced using probability sampling techniques.
- Inferences drawn from sample are generalizable to the population.

Disadvantages:

- The techniques need a lot of efforts.
- A lot of time is consumed.
- They are expensive.

Non-Probability Sampling Methods

> Probability sampling is also called as judgment or non-random sampling.
> Every unit of population does not get an equal chance of participation in the investigation. No random selection is made.
> The selection of the sample is made on the basis of subjective judgment of the *investigator*.
> These techniques need not population to be very precisely defined.
> These techniques can be used for both types of population: the population that is too general a category, and the population that is a specific category (precisely defined).
> For instance if our target population is defined as college students. It means person studying at any college of the world is an *element* of our population. It is too general a category consisting of infinite number *of elements.*
> Thus, non probability techniques make it possible to take a sample of population the elements of which are infinite in number.
> Non-probability sampling is well suited for exploratory research intended to generate new ideas that will be systematically tested later.
> Probability sampling is well suited for research that is intended to develop the understanding of a population.

> **Advantages:**

- The techniques need less effort.
- These techniques need less time to finish up.
- They are not much costly.

> **Disadvantages:**

- The sampling techniques are prone to encounter with *systematic errors* and *sampling biases.*
- The sample cannot be claimed to be a good representative of the population.
- Inferences drawn from sample are not *generalisable* to the population.

Types of probability sampling methods

Following methods are used for probability sampling:

1. Simple Random Sampling
2. Systematic Random Sampling
3. Stratified Random Sampling
4. Cluster Sampling
5. Multistage Sampling

1) Simple Random Sampling:

> In this type of sampling each and every element of the population has an equal chance of being selected in the sample.
> The population must contain a finite number of elements that can be listed or mapped.
> Every element must be mutually exclusive i.e. able to distinguish from one another and does not have any overlapping characteristics.
> The population must be homogenous i.e. every element contains same kind of characteristics that meets the described criteria of target population.

2) Systematic Random Sampling

> This type of sampling is also used for homogenous population.
> It is a bit different from simple random sampling.
> Unlike simple random sampling, there is not an equal probability of every element been included.
> In this type of sampling the elements are selected at a regular interval.
> The interval may be in terms of time, space or order. For instance, element appearing after every 30 minutes, or present at a distance of two meters, or every 5th element present on a list.
> Thus this regularity and uniformity in selection makes the sampling systematic.
> The list of elements may or may not be required before the conduction of research.
> Sometimes it is not even possible to create a list because of the nature of population. Say, if it is possible to tell who is going to visit the coffee shop today.

3) Stratified Random Sampling

> This type of sampling method is used when population is heterogeneous. i.e. every element of population does not matches all the characteristics of the predefined criteria.
> Instead the elements differ from one another on a characteristic.
> So the sub groups are formed that are homogenous i.e. all the elements within a group contains same kind of characteristics (keep in mind, those characteristics are to be taken into account that defines the target population).
> The sub groups are called as strata (single stratum).
> The topic and nature of the investigation tells on what criterion the strata are to be made.

- Common criterions used for stratification are gender, age, ethnicity, socioeconomic status. However, the criterion vary greatly investigation to investigation
- This formation of strata can also be called a mini reproduction of population as each stratum consists of elements that are different from other strata's element in some characteristics.
- For instance if an investigation is taking young adults into account, so this population may need to be divided (of course, on the basis of what the investigation is about) into sub groups like male young adults and female young adults, educated young adults and uneducated young adults, high income young adults and low income young adults etc. in this way each stratum is a different population
- The sample is selected from each stratum randomly.
- There are two techniques that are used to allocate sample from strata: proportional allocation technique and equal allocation technique.
- Using proportional allocation technique the sample size of a stratum is made proportional to the number of elements present in the stratum.
- Using equal allocation technique same number of participants is drawn from each stratum regardless of the number of elements in each stratum.

4) Cluster Sampling

- The group of elements residing in one geographical region is called as cluster.
- And sampling of clusters is called as cluster sampling.
- This sampling technique is used when the elements of population are spread over a wide geographical area.
- The population is divided into sub-groups called as clusters on the basis of their geographical allocation.
- Usually this division of population is similar to what the standard of division has been used yet.
- For instance population spread over a country is clustered up into cities, population spread over a city is clustered up into towns etc.
- The clusters ought to be homogenous among them on the characteristic variable of the research.
- However for being truly representative sample, the selected clusters must capture the heterogeneity of population.
- For instance if in the selection of towns only small towns are selected leaving behind the bigger towns, the sample is not going to be a true representative of the population.

5) Multistage Sampling

- It is a sampling technique where two or more probability techniques are combined.
- It is used when the elements of population are spread over a wide geographical region and it is not possible to obtain a representative sample with only one aforementioned technique.

> It can be described as sampling within the sample.
> The final unit or element of population which is used in investigation is obtained after sampling at several stages.

Random sampling methods

- **Simple Random Sampling:** Every member of the population is equally likely to be selected).
- **Systematic Sampling:** Simple Random Sampling in an ordered systematic way, e.g. every 100th name in the yellow pages.
- **Stratified Sampling:** Population divided into different groups from which we sample randomly.
- **Cluster Sampling:** Population is divided into (geographical) clusters - some clusters are chosen at random - within cluster units are chosen with Simple Random Sampling.

Terminologies

- **Element -** a single indivisible entity of a population.
- **Exhaustive list -** a list which contains each and every element in a way that not a single element is left.
- **Generalizability -** the extent to which inferences drawn from a sample are true for the population**.**
- **Heterogeneous population -** a population the elements of which are not similar to each other.
- **Homogenous population -** a population whose every element is similar regarding the variables required for the research.
- **Investigator/ Researcher -** one who is conducting a research.
- **Lottery method -** method of selecting elements by taking out the slips from a bowl, box etc.
- **Mutually exclusive category -** a category of sampling technique which is totally different and is not a sub or super category of other types.
- **Participant -** a person taking part in a research.
- **Sampling bias -** a situation where the selected sample does not truly reflect the characteristics of population.
- **Sub groups -** groups within a population that differ from one another on some variables.
- **Systematic errors -** the errors that are caused by over or under representation of some characteristics of population in the sample.

Difference between probability and non-probability sampling techniques

	Probability sampling Technique	**Non probability Sampling Techniques**
Requirement of resources	Require more resources in terms of time, cost and efforts.	lesser resources are required
Selection of sample	random; no subjective judgment of researcher is involved	non random; subjective judgments of researcher is involved
Quality of Inferences Drawn	generalizable to the population	not generalizable to the population
Best suited for researches	the goal of which is to understand a population	the goal of which is to understand or develop a concept or idea
Applicable to The kind of Population	elements of which are finite in number which is precisely defined and a specific category	elements of which are finite elements of which are infinite which is a too general category, not very precisely defined which is precisely defined
Chances of Error and Biases	sampling errors and systematic biases are less likely to occur	prone to encounter with systematic errors and sampling biases
Types	• Simple random sampling • Systematic random sampling • Stratified random sampling • Cluster sampling • Multistage sampling	• Volunteer sampling • Convenient sampling • Purposive sampling • Quota sampling (proportional and non proportional) • Snowball sampling • Matched Sampling • Genealogy Based Sampling

Name of Type	Characteristics	Method	Benefits	Draw backs/ Crucial Issues
Simple Random Sampling	Every element has an equal chance of been selected Exhaustive lists of elements are essential Elements are selected randomly	Exhaustive list of elements is produced Each element is allotted a number Numbers are randomly selected through lottery method or using computer generated random table	Omits the chance of systematic errors and sampling biases Representative sample is produced	Difficult for very large population
Systematic Random Sampling	elements are selected at a regular interval (may be time, order or space) Exhaustive list may or may not be required is used when a homogenous population is grouped within itself	in the case where the lists are available, the lists are compiled to form a single list each element is given a number to select an appropriate interval, N^{*} is divided by n^{**}; number obtained by this division (say k) is the size of interval first an element is selected at random and then every k^{th} element from the first selection is included in the sample	ensures the extension of sample to the whole population make it possible to get a probable sample where list of elements cannot be produced	If the existing grouping is biased in some way the sample may not be representative difficult for very large population
Stratified Random Sampling	is used when a population is heterogeneous	first the population is divided into homogenous sub groups called as strata then elements are randomly selected from each stratum	ensures a representative sample for a heterogeneous population	it requires more resources in terms of time and efforts if the variable used for making strata is not appropriate depending on the research, the whole working may go in vain

Cluster sampling	is used when the target population is homogenous but is spread over a wide geographical region instead of elements clusters are randomly selected a cluster is defined as a group of elements residing together in one geographical region	first the population is divided into clusters each cluster is allotted a number then the decided number of clusters are selected randomly	make the probability sampling possible for a large population	there is possibility of systematic error if the selected clusters fail to capture the characteristic diversity of the target population, the sample cannot be claimed to be representative of the population
Multistage sampling	it can be defined as sampling within the sample two or more probability sampling techniques are combined first a sample is extracted randomly and then from the selected sample another sample .	first the target population is divided into clusters clusters are randomly selected out of the selected clusters, there may be a formation of clusters or strata (in case of heterogeneity of population) now a random selection of clusters is done/ or there occurs a selection of elements from each strata the final units selected are investigated	a representative sample is produced for a population that is spread over a wide geographical region and is also heterogeneous	if the characteristic criterion used for the formation of strata at any stage is not appropriate, the sample cannot be representative of the population if also there occurs a systematic error in the selection of clusters, the results would not be able to generalize