STUDY POPULATION AND SAMPLING

By

David Onen (Ph.D.)

College of Education and External Studies

Makerere University

A Paper Presented on 6th January 2020 to the Participants at an Academic Staff Seminar held in the School of Research and Postgraduate Studies of Amoud University, Somaliland

What is a Study Population…?

- It refers to a complete set of elements (persons or objects) that possess some common characteristic defined by the sampling criteria established by the researcher.
- It is often composed of two groups - target population & accessible population.
- The target population (also called the universe/parent population) is the entire group of people or objects to which the researcher wishes to generalize the study findings.
- Such a population should meet set of criteria of interest to the researcher e.g.

➢ All the academic staff of universities in Somaliland.

➢ All school-age children with visual impairment studying in primary schools in Borama District.

What of the Accessible Population…?

- The accessible population refers to the portion of the target population to which the researcher has reasonable access and from which a sample can be drawn.
- It is considered as the subset of the target population.
- It may be limited to region, district, county, or institution – depending on a set criterion e.g.

➢ All the academic staff in Amoud and University X in Somaliland.

➢ All school-age children with visual impairment studying in primary schools in schools X, Y, Z in Borama District.

- Note: a researcher must always justify the choice of both the target and accessible populations and appropriately describe them.

Meanwhile, a Sample is…

- A sample is the selected elements (people or objects) procedurally chosen for participation in a study to represent the target or accessible population.
- A sample of people is referred to as subjects or participants.
- The researcher must determine his/her sample carefully considering factors like e.g.

➢ The sampling frame – which is the list of all the elements in the population from which the sample is drawn.

➢ Sometimes, the frame can be so large but the researcher needs to provide opportunity for everyone in the population to be identified so that they will have an equal opportunity for selection as a subject (element).

What is Sampling…?

- Sampling is the process of selecting a group of people, events, behaviors, or other elements with which to conduct a study.
- It is actually the process of selecting “few cases” from a target or accessible population in order to provide information that can be used to make judgments about a much larger number of cases.
- The “few cases” are called the sample and the whole group from which the sample has been drawn is referred to as the population or universe.
- In most cases, there are methodological, statistical, theoretical and practical limitations which make the study of whole populations difficult, if not impossible.
- This causes the need to use a part of the population (the sample) to gain sufficient insight into the characteristics of the population.

What is Sampling…?

- Sampling is the process of selecting a group of people, events, behaviors, or other elements with which to conduct a study.
- It is actually the process of selecting “few cases” from a target or accessible population in order to provide information that can be used to make judgments about a much larger number of cases.
- The “few cases” are called the sample and the whole group from which the sample has been drawn is referred to as the population or universe.
- In most cases, there are methodological, statistical, theoretical and practical limitations which make the study of whole populations difficult, if not impossible.
- This causes the need to use a part of the population (the sample) to gain sufficient insight into the characteristics of the population.

Why Sample…?

- To shorten the time that would otherwise be spent studying the whole population.
- To reduce costs since fewer cases and resources are required than if the whole population was used.
- To gain greater accuracy in data collection and analysis, since few people are involved either as subjects or as participants.
- To increase the scope and flexibility of coverage because of the ability to use limited time and other resources.
- To avoid wasting resources to study a whole population when the same information can be obtained from a part of the population

But Sampling has weaknesses too…

- Despite the usefulness of a sample in research, it suffers from two major weaknesses.

1. Firstly, it is merely an estimate and no matter how accurate and exhaustive, it can never comprehensively describe the population.
2. Secondly, in studies where the treatment leads to destruction of the sampling units, sampling becomes very uneconomical.

Common Terms Used in Sampling…

- Randomization refers to a scenario where each individual in the population has an equal opportunity to be selected for the sample.
- Representativeness refers to the situation where a sample must be as much like the population in as many ways as possible.
- Parameter refers to a numerical value or measure of a characteristic of the population; *remember P for parameter & population*.
- Statistic refers to the numerical value or measure of a characteristic of the sample; *remember S for sample & statistic*.
- Precision refers to the accuracy with which the population parameters have been estimated; *remember that population parameters often are based on the sample statistics*.

Sampling Strategies…

There are two main ways of selecting study samples from the target population: Probability and Non-probability sampling techniques.

Probability Sampling

- Probability sampling refers to a situation where the chance of each member of the population (or more specifically - the chance of each sampling unit) being included in the sample is known prior to drawing the sample.
- This is the only sampling process that ensures the selection of a representative sample.
- There are four main probability sampling techniques: simple random sampling stratified random sampling, area (cluster) sampling and systematic sampling.

Non-probability Sampling

- This is a sampling strategy in which the chance that an individual will be selected is unknown – hence the researcher cannot claim that the sample is representative of the larger population.
- They are techniques based on non-representativeness. Results from such studies lack external validity and are generally limited in terms of generalizabilty to the larger population.
- However, they are preferred because they are much less complicated and less expensive.
- They are often used in a combination and as supplements to probability sampling.
- The common non-probability sampling techniques include: convenience, purposive, quota and networking (snow-ball) sampling techniques.

Table 1: Summary of Common Sampling Techniques

Sampling Technique	Meaning	Purpose	Justification
Simple Random Sampling (SRS)	Select a sample without bias from the accessible population.	To select a random (representative) sample.	Ensures that each member of the target population has an equal and independent chance of being included in the sample.
Stratified Sampling	Identify sub-groups in the population and their proportions and select from each sub group to form a sample.	- To group population into homogenous subsets that share similar characteristics. - To ensure equitable representation of the population in the sample.	- Ensures that subgroups are proportionately represented. - Accounts for the difference in subgroup characteristics.
Systematic sampling	Arrange the study population in a certain order and select the sample population after a defined interval.	Used to select individual subjects to represent study population	Saves time and money.
Cluster Sampling	A random sampling process that involves stages of sampling The population is divided into relatively smaller groups (clusters) & part of the clusters randomly selected as the sample. All members of the chosen	Used to select group rather than individual members.	Saves time and money.

Table 4.1: Summary of Common Sampling Techniques

Snow-ball or Networking	Select a member and that member leads you to other members	Used for identifying subjects difficult to get e.g. prostitutes.	Easy to implement
Purposive Sampling	The researcher decides who to include in the sample.	Used to collect focused information.	Selects typical and useful cases only. Saves time and money.
Convenience Sampling	Selects those who happen to be available on first come first served basis.	Used for pilot or exploratory studies.	- Collects data at the spur of the moment. - Takes advantage of those who happen to be there.
Quota Sampling	Identify sub-groups in the population and their proportions and select from each sub-group, but not at random, to form a sample.	To group population into homogenous subsets that share similar characteristics.	Accounts for the difference in subgroup characteristics

The choice of a sampling technique is guided by:

- The objectives/hypotheses/questions of the study.
- The purpose of the sampling technique. Each sampling technique is intended to achieve a particular purpose. The purpose of the study must tally with the purpose of the sampling technique.
- Careful examination of the relationships between the variables in the study.

Techniques of Determining Sample Size

- **Non-Statistical Estimations**

The sample size is decided by looking at several factors in the study without applying any approved mathematical formula. Such issues that are considered include:

The Type of Research: Generally the minimum sample size is determined by the research design

For example, the following rule is conventionally applied:

Type of Research	**Sample Size**
Survey: - **Each major group.** - **Each minor subgroup.**	 100 20 – 50
Ex-post Facto/Experimental	15 in each condition
Correlation	30

Source: Adapted from Kathuri & Palls (1993).

Statistical Methods…

- There are mathematical formulae available for determining the required sample size.
- These formulae are based on:

1. The type of research being undertaken.
2. The knowledge of the population regarding required characteristics e.g. variability, level of confidence.
3. Whether the study is comparing the averages, frequencies, etc.

Determining Sample Size through Power Analysis

- Power analysis is a procedure for estimating either the likelihood of committing a Type II error or a procedure for estimating sample size requirements.
- To use it to determine sample size, you need to have the following data:
- Level of significance criterion = alpha a, use .05 for most social science studies and your calculations.
- Power = 1 - b (beta); if beta is not known standard power is .80, so use this when you are determining sample size.
- Population size effect = gamma g or its equivalent, e.g. eta squared h^2; use recommended values for small, medium, or large effect for the statistical test you plan to use to answer research questions or test hypothesis.

Use of Statistical Tables…

- There are tables available in books of statistics for determining the required sample size.
- However, these tables only deal with finite populations.
- As such, the researcher has to know the number of subjects in the population he/she intends to study

Table 1: Table for Determining Sample Size for a Finite Population

N	*S*	*N*	*S*	*N*	*S*
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357
100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384

Note.—*N* is population size. *S* is sample size.

Source: Krejcie & Morgan, 1970

Adapting Sample Size

- **Sample sizes can be adapted from studies whose credibility are proven.**
- **Studies whose results are published in refereed Journals can be considered to have employed credible methodologies; hence their sample sizes could be adapted for similar studies.**

But Avoid Sampling Error and Sampling Bias….

- Sampling error refers to the difference between the sample statistic (e.g. sample mean) and the population parameter (e.g. population mean) that is due to the random fluctuations in data that occur when the sample is selected.
- While sampling bias (also called systematic bias or systematic variance) refers to the difference between sample data and population data that can be attributed to faulty sampling of the population.
- The sampling errors result into Type 1 and 2 errors in data analysis.
- Type 1 error is based on the statistical analysis of data in which the researcher wrongly rejects a true null hypothesis; and therefore, accepts a false alternative hypothesis.
- While Type error 2 is based on the statistical analysis of data in which the researcher wrongly accepts a false null hypothesis; and therefore, rejects a true alternate hypothesis.

Sampling in Qualitative Studies…

- In qualitative research, there are no rigid rules on sample size determination.
- The common view rests on reaching the point of saturation.
- Saturation is defined by many as the point at which the data collection process no longer offers any new or relevant data.
- But since qualitative research involves an intensive study of individuals, a small sample is usually required.
- Many scholars recommend between 5 to 50 subjects. Others recommend a minimum of 25 to 30 while others recommend between 10 to 20 subjects!
- In most cases, the researcher does not determine the sample size in advance.
- Sample size will depend on reaching the point of saturation or by using a prior set criterion.

Writing the sampling techniques…

- A well-described sampling technique should:
  - Name the sampling technique.
  - State the purpose (use) of the technique in the study.
  - Describe the main features of the sampling technique(s).
  - Provide a justification for selecting the sampling technique for the study.

THANK YOU
LADIES AND GENTLEMEN!