"How To": Choose Sampling Techniques for Evaluations

By Tom Mwebaze
School of Economics, Makerere University

1. Define Sampling

Sampling is the process of selecting units (e.g. people, organizations, time periods) from a population of interest, studying these in greater detail, and then drawing conclusions about the larger population. The evaluator collects data from a subset of individuals – *a sample* – and uses those observations to make inferences about the entire population. The characteristics of the sample should reflect the characteristics of the population targeted by the intervention. In that case, the evaluator's conclusions from the sample are probably applicable to the entire population.

2. Techniques of Sampling during Evaluation Process

Sampling techniques can be classified into two major categories: Probability (Statistical Sampling/Proportional Sampling) and Non-probability (Non-statistical Sampling).

a) Probability/ Statistical Sampling/ Proportional Sampling

In probability sampling, the population to be sampled is defined by the number of items/units. It is usually used to reveal the presence (or absence) of a defined characteristic in a portfolio of items/units with similar characteristics. Probability sampling options use random or quasi-random options to select the sample, and then use statistical generalization to draw inferences about that population. To minimize bias, these options have specific rules on selection of the sampling frame, size of the sample, and managing variation within the sample. Consider why you want to study your population of interest and what you want to do with the information that you have gathered, before you choose your option.

The options include:

- Simple random: drawing a sample from the population completely at random.
- Systematic: selecting every n^{th} case from a list (e.g. every 10^{th} client).
- Cluster sampling in which larger clusters are further subdivided into smaller, more targeted groupings for the purposes of surveying.
- Stratified: splitting the population into strata (sections or segments) in order to ensure distinct categories are adequately represented before selecting a random sample from each.

Each of the options starts with a *sampling frame*, which can be thought of as a list of all elements in the population of interest (e.g. names of individuals, telephone numbers, house addresses, and census tracts). The sampling frame operationally defines the target population from which the sample is drawn and to which the sample data will be generalized.

Simple Random

This technique gives all units in the population an equal opportunity of being selected by using a method that will select units completely at random. To use this technique you will need a complete list of everyone in the population – a sampling frame. Sampling is done in a single stage with each element selected independently. You can randomly select by using a random number generator or by using a random number table, or by pulling out of a hat. It is very effective if you have a small or moderate sized population.

Systematic sampling

Samples are drawn by starting at a randomly selected element in the sampling frame and then taking every n^{th} element. Systematic sampling is easier to perform and hence is less subject to interviewer errors than simple random sampling.

Stratified Random sampling

This technique divides the population into meaningful homogenous or similar groups based on a certain characteristic (e.g. gender, race, socioeconomic status) called strata, and then selects a random sample from each group. Stratified sampling ensures greater representativeness on a characteristic of interest

within the population. This method allows you to study a wider range of the population without a larger sample size.

Cluster Sampling
The sample is selected in stages, first selecting groups of elements, or clusters (e.g. city blocks, census tracts, schools), and then selecting individual elements from each cluster (e.g. randomly or by systematic sampling). In case there is no list of every member of the population, it is still possible to choose a random sample by using cluster sampling. Cluster sampling divides your population into groups and a simple random selection of those groups is made. You then survey everybody within the selected groups.

Method	Best when
Simple random sampling	Whole population is available
Stratified sampling (random within target groups)	There are specific sub-groups to investigate (e.g. demographic groupings)
Systematic sampling (every n^{th} person)	When a stream of representative people are available (e.g. in the street)
Cluster sampling (all in limited groups)	When population groups are separated and access to all is difficult, e.g. in many distant cities

b) Non-probability/ Non-statistical Sampling
The difference between non-probability and probability sampling is that non-probability sampling does not involve *random* selection and probability sampling does. With non-probability samples, we may or may not represent the population well, and it will often be hard for us to know how well we've done so. In general, researchers prefer probabilistic or random sampling methods over non-probabilistic ones, and consider them to be more accurate and rigorous. Non-probability sampling can be classified into purposive, snowball, convenience sampling or quota sampling.

Purposive Sampling (or Purposeful)
This is sampling with a *purpose* in mind. We usually would have one or more specific predefined groups we are seeking. Purposive sampling can be very useful for situations where you need to reach a targeted sample quickly and where sampling for proportionality is not the primary concern. The variables to which the sample is drawn up are linked to the research question**.** Purposive sampling options study information-rich cases from a given population to make analytical inferences about the population. Units are selected based on one or more predetermined characteristics and the sample size can be as small as one (n=1). To minimize bias, this cluster of options encourages transparency in case selection, triangulation, and seeking out of disconfirming evidence. Only people with specified characteristics are selected. These characteristics can represent the range of characteristics in your population of interest, or you can focus on some characteristics in particular.

For example, if you are evaluating the attitudes of drivers towards speeding, you may want to only sample those who have got penalty points. Or, you could sample those with extreme characteristics by only selecting drivers who have been disqualified as a result of multiple offences.

Snowball sampling
A type of purpose sampling where existing participants recruit future subjects from among their acquaintances. Thus the sample group appears to grow like a rolling snowball. Using this approach, a few potential respondents are contacted and asked whether they know of anybody with the characteristics that you are looking for in your evaluation. For example, if you wanted to interview a sample of vegetarians/cyclists/people with a particular disability/people that support a particular political party or the homeless, your initial contacts may well have knowledge of other. In these cases you are not likely to be able to find good lists of people of these characteristics within a specific geographical area. However, if you go to that area and identify one or two, you may find that they know very well who the others in their vicinity are and how you can find them. This makes it easier to contact people you might

not otherwise have any way of getting in touch with. However, it does introduce a lot of bias into your results because your samples are all likely to know each other and to have similar opinions. With this method you survey whoever you happen to have access to. It is also called 'opportunity sampling'. This is commonly used in market research.

Quota
This is another method commonly used in market research. With quota sampling, you divide the population (e.g. older drivers) into distinct parts (strata). You then decide how many of each stratum you want to have in the total sample. For example, you could stand outside an out-of-town supermarket in the day time. You have decided that you want to survey 50 male drivers aged over 65. Every time you see a male older driver enter the store you ask if he is aged over 65 and you continue until you reach your quota of 50 completed surveys. This is the non-random form of stratified sampling.

Method	Best when
Quota sampling (get only as many as you need)	You have access to a wide population, including sub-groups
Proportionate quota sampling (in proportion to population sub-groups)	You know the population distribution across groups, and when normal sampling may not give enough in minority groups
Non-proportionate quota sampling (minimum number from each sub-group)	There is likely to be a wide variation in the studied characteristic within minority groups
Purposive sampling (based on intent)	You are studying particular groups
Expert sampling (seeking 'experts')	You want expert opinion
Snowball sampling (ask for recommendations)	You seek similar subjects (eg. young drinkers)
Modal instances sampling (focus on 'typical' people)	When sought 'typical' opinion may get lost in a wider study, and when you are able to identify the 'typical' group
Diversity sampling (deliberately seeking variation)	You are specifically seeking differences, e.g. to identify sub-groups or potential conflicts

3. Sampling contributes to a quality evaluation

Sample size
Decisions on sample size depend on whether your data is quantitative or qualitative. For quantitative data, you use statistical calculations to determine how large your sample should be.

Response rates
Once a sample is selected, an attempt is made to collect data (e.g. through interviews or questionnaires) from all of its members. In practice, researchers never obtain responses from 100% of the sample. Some sample members inevitably are travelling, hospitalized, incarcerated, away at school, or in the military. Others cannot be contacted because of their work schedule, community involvement, or social life. Others simply refuse to participate in the study, even after the best efforts of the researcher to persuade them.

Unless every person in a sample is interviewed, the sample will not be a true random sample. Non response must therefore be kept to negligible proportions if the results of the survey are to be valid. If a respondent is not interviewed when first called upon, it will be necessary to follow-up with a second call. The simplest way of dealing with non-response is to treat the non-respondents as being similar to the respondents. This treats the people who do not respond at all as if they are similar to the initial non-respondents who subsequently respond, but takes into account differences between those who respond initially and those who do not.

It is useless to substitute the person next door. Suppose, for example, an enquiry is made as to the size of family. Then people without children are more likely to be out than those who have children. If the people without children are not interviewed because they are not at home, and the interviewer turns to other people who are at home, the results will obviously be biased and inaccurate.

Sampling Errors

It has already been mentioned that even when a correct process is employed in choosing a sample, the sample cannot be exactly representative of the aggregate. Thus errors will arise. These inevitable errors arising from the fact that the sample will vary from the aggregate are called random sampling errors. The amount of the error will depend upon:

(a) the size of the sample;
(b) the extent of the variability of the material;
(c) the sampling procedure employed, and;
(d) the way in which the results are calculated.

However, given a proper method of selecting the sample, the probability of errors of any given size can be calculated from the detailed results of the actual sample. By calculating the errors of different sampling methods which can be used on the same material, it is possible to plan further surveys more efficiently. The use of appropriate sampling methods and an adequate response rate are necessary for a representative sample, but not sufficient. In addition, the sample size must be evaluated. The sampling error is a number that describes the precision of an estimate from any one of those samples. It is usually expressed as a *margin of error* associated with a statistical level of confidence.

4. Conclusion

In conclusion, the following items need to be considered when selecting the sampling technique:

- A combination of different sampling methodologies at different stages in evaluation.
- The representation of different stakeholders.
- Correction of inherent biases.

References

- Cochran W. G. (1963). *Sampling Techniques*. (2nd ed.). New York: John Wiley and Sons, Inc.
- Fowler, F. J. (1988). *Survey Research Methods*. (rev. ed.). *Applied Social Science Research Methods Series, Volume 1*. Newbury Park, CA: Sage Publications.
- Fruend, J.E. and Walpole, R.E., Mathematical Statistics, Prentice Hall, Inc.
- Mendenhall W. et al, Mathematical Statistics with Applications. PWS-Kent Publishing Company.
- Morris H. (1991). Statistical Analysis for Decision Making, (5th Ed).
  http://changingminds.org/explanations/research/sampling/choosing_sampling.htm