# UNDERSTANDING SURVEY SAMPLING

WHY IS THIS IMPORTANT?

In social research, sampling is the process of selecting a group of subjects from the larger population that is being studied. You can make a sample from households, organizations, a value chain or any other unit of analysis. Drawing up a sample is an essential step in undertaking most types of survey (Oxfam's research guideline on Planning Survey Research outlines other key steps). Information gathered from a sample can be used to draw inferences about the whole population, a relationship schematized in the diagram below. The more representative a sample is, the more accurate inferences based on it are likely to be. If you do not have a sample and simply collect data from anyone, you will not be able to explain what the data refers to.

Sampling

Inference
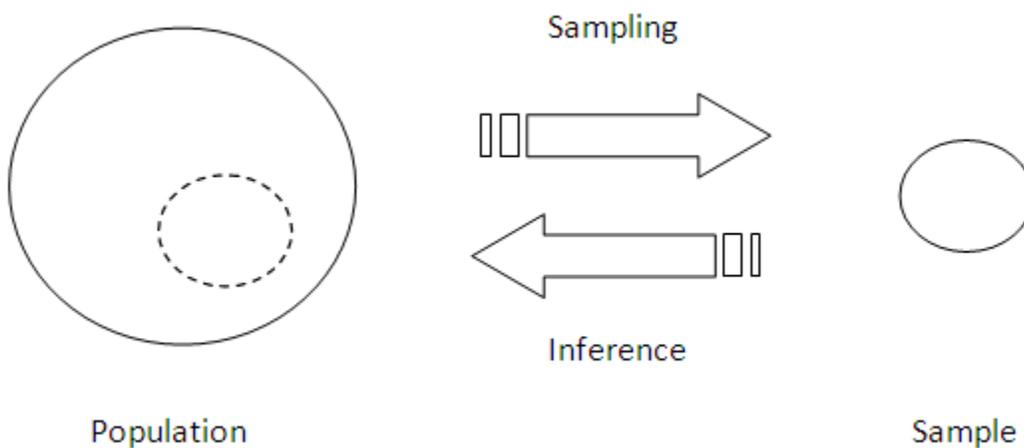
Population

Sample

Sampling is different from the procedure used in administering censuses, when the whole population is surveyed. Unless the population being studied is very small, sampling is much quicker and cheaper, and generally the only feasible way to undertake a survey and generate credible results about the population as a whole. While survey researchers should know how to draw up different kinds of samples, understanding this process is also important for others involved in commissioning surveys and interpreting the results. This guideline therefore provides an introductory overview of sampling and different kinds of samples.

RESEARCH GUIDELINES

KEY TERMS USED IN SAMPLING

This guide uses key terms from sampling theory and practice to ensure clarity in communication and alignment with the state of the knowledge about sampling.

- **Sampling units**: the type of entity on which data are sought (individuals, households, schools, etc.).
- **Population** (universe): the full set of units of analysis about which you want to infer conclusions.
- **Sampling frame**: a list of all units of analysis in a given population.
- **Census**: the collection of data from all units of analysis in a given population.
- **Sample:** a subset of units drawn from the sampling frame or sometimes directly from the population.
- **Sampling error:** mistakes in estimation that reflect the difference between a sample and the population from which it is drawn.

KEY STEPS IN THE SAMPLING PROCESS

The sampling process can be divided into three key steps: definition of the population of interest, identification of the sampling frame, and selection of the sample.

1. Determine the population of interest

The first stage is to identify the population of interest for research. This depends on the research question(s) the survey is trying to answer. In some cases the definition of the population is obvious, for example when the aim is to collect descriptive statistics about people in a particular geographic area (with defined boundaries). In other cases the population of interest may be harder to define. For example when evaluating impacts of development projects, it is important to sample both project participants (referred to as the 'intervention group') and non-project participants (the 'comparison group') whose selection may depend on a range of factors. Time spent carefully defining the population of interest is seldom wasted time, and it often helps to raise questions and ambiguities that become relevant later.

2. Identify the sampling frame

This entails finding the most comprehensive list that can be obtained of the sampling units (usually households or individuals) in the population of interest. A list of all the households or residents in a location, for example, can sometimes be obtained from local authorities or a previous census. This must be as accurate and complete as possible; in particular you should make sure that portions of the population with particular characteristics are not systematically excluded.

Where it is not possible to obtain a full and complete list of sampling units, innovative techniques are now available that use satellite images in order to map settlements on the ground and so construct sample frames (Shannon et al. 2012).

3. Select the sample

This entails selecting the units for data collection. The sample can be selected from the sampling frame using either probability or non-probability sampling techniques:

- **Probability samples**, sometimes called **random samples**, are those in which every unit in the population has the possibility (a 'nonzero' probability) of being selected. It is not necessary for each element to have the same chance of selection, but it must at least have some chance, and this chance needs to be known.
- **Non-probability samples**, including quota and accidental samples (defined below), are not representative of the whole population. These sampling strategies can be used for testing questionnaires or for ad hoc research that does not require making inferences or generalizing estimates to the entire population.

Only probability samples allow conclusions to be drawn about the whole population, test for statistical significance and compute for confidence intervals, and preclude the possibility of consciously or unconsciously introducing biases in the sample. They are therefore preferable to non-probability samples.

DIFFERENT KINDS OF SAMPLES

1. Probability samples

The following are the most common kinds of probability sampling:

Simple random sampling

In simple random sampling (SRS) the units to be surveyed are randomly selected from the sampling frame. This means that each unit (individual, household, etc.) will have the same probability of selection. If the sample frame comprises a list of all the households in a location, then SRS will entail picking a random selection of these households for data collection.

SRS is relatively easy and straightforward to implement. It can, however, have some limitations if the researcher is investigating questions specific to particular subgroups of the population.[1] Stratified sampling (discussed next) addresses this weakness of SRS.

Stratified random sampling

Stratified random sampling (SS) is an approach that makes it easier to compare and contrast different recognized subgroups in a population. Examples of such subgroups are different socio-economic classes or castes, religious denominations and genders. SS entails first dividing the population in the sampling frame into distinct categories called 'strata' based on the subpopulations of interest. Each stratum is then sampled as an independent subpopulation in which individuals or other survey units can be randomly selected.

For example, suppose you are interested in studying a population of 10,000 which is divided into one group of 8,000 people (A) and a second group of 2,000 (B), while you have resources sufficient for a sample of 100, and your priority objective is to generalize results to all 10,000 people in the population. For a **proportional stratified random sample**, you would randomly draw 80 people from A and then 20 from B to guarantee the proportion of people in each stratum is the same as in the population. Note that sampling 100 people using simple random sampling might yield a sample with 88 from A and only 12 from B, or perhaps 75 from A and 25 from B. Proportional stratified random

---

[1] The concern is that with smaller subgroups there might be some random error which can increase the standard errors, reducing power for comparison.

sampling ensures that will not happen. On the other hand, if your primary purpose was to test whether group B has different levels of a certain outcome compared with A, then it would be best to sample 50 observations from A and 50 observations from B. This would increase the precision of inferences made about group B, as the sample would be much larger. This is called a **disproportional stratified random sample**.

SS can increase the precision of inferences made to the full population or increase inferences made to comparison among strata, but it cannot do both at the same time. Increasing precision of inferences to the full population is done by keeping the ratio of sample size and population size equal to all subpopulations (proportional stratified random sample). Increasing precision of inferences to comparisons among strata is done by taking approximately equal-sized samples for each group (disproportional stratified random sample).

Multi-stage sampling

For practical purposes it is sometimes more cost-effective to choose the sample in a process of two or more successive stages, which is called multi-stage sampling. **Primary sampling units** (PSU) are units selected at the first stage, **secondary sampling units** (SSU) are those selected at the second stage, and so on. The most common multi-stage sampling strategy is **two-** (or three-) **stage cluster sampling**.

A cluster consists of the units in well-defined geographic areas (such as a village, city block, school, etc.). The two-stage cluster involves selecting the clusters in the first stage. The clusters are usually selected using a **probability proportional to size** (PPS) method, where larger clusters have a greater probability of being selected for the sample. Within each selected cluster a list of all SSU is then drawn up. In the second stage a fixed (or proportional) number of SSU is then selected from each selected cluster using simple random sampling.

Using two-stage cluster sampling has the advantage of reducing the cost of travel and survey administration, and it guarantees a representative sample of the target population when there is a list of all clusters to be sampled but not a list of all target individuals within each cluster. However, this sampling approach is generally not advisable if there are only a small number of clusters in the population, or only a small number of units within many of the clusters.

2. Non-probability samples

There are three common non-probability sampling techniques.

Convenience samples

A convenience sample is a non-probabilistic technique that involves selecting a sample on the basis of its accessibility or other factors of convenience. As a result some categories of people are likely to be overrepresented and others underrepresented or excluded, making it impossible to generalize with any accuracy from this kind of survey. This kind of sample should only be used as a last resort.

Snowball samples

Snowball sampling is used when the boundaries of the population are unknown, and it is impossible to obtain the sample frame. In this type of sample each suitable unit which is identified is included in the sample, and then used to identify other appropriate units that are also included. This method should

only be employed when the sample frame is unknown.

Purposeful samples

Purposeful (or 'judgement') samples are samples in which the units have been selected to meet the specific purposes of the study, but not as a basis for generalization to the whole population. Some examples of purposeful samples are:

- **Typical cases (or median) samples**: where units are drawn deliberately in order to represent the middle range of some characteristics of interest.
- **Maximum variation (heterogeneity) sample**: where units are drawn deliberately in order to represent the full range of characteristics of interest.
- **Extreme cases sample**: where units are drawn deliberately in order to represent some extreme characteristics of interest or outliers (for example, the oldest and youngest, or poorest and richest).
- **Quota samples**: where units are selected using fixed quotas based on some characteristics of interest.

DETERMINING SAMPLE SIZE

One frequent question is how big a sample should be for the results of a survey to be statistically valid. The sample size is particularly important for any study in which the goal is to make inferences about the sampled population. For many practical purposes sample size depends on a number of variables, the most important being:

- Time and budget available.
- The size of the population of interest.
- Precision.
- Statistical power.

While the first two are self-explanatory, the third and the fourth of these points require more explanation. Precision is a measure of how close a sample estimate is likely to be to the actual population parameter. It may be expressed as standard error or absolute error. Statistical power refers to the probability of detecting a certain statistical effect, given that the effect exists in the population. Statistical power is largely determined by four factors: sample size, effect size, variance of the variable of interest and level of significance. For this reason, before conducting a survey, a statistical power analysis is required in order to determine the size of the sample and its allocation across clusters and strata. Readers unfamiliar with statistical techniques involved should seek technical advice and/or consult one of the available guidelines, such as the [Institute of Fiscal Studies' practitioner's guide](#) (McConnell and Vera-Hernández 2015) or the [3ie manual on power calculation](#) (Djimeu and Houndolou 2016). As a rule of thumb, sample sizes allowing for robust comparisons in means are usually in the hundreds (between 400 and 900 in the case of most [Oxfam Effectiveness Reviews](#). Smaller sample sizes are more likely to have insufficient power, meaning that they may not able to find a statistical difference when in fact there is a difference between groups.

REFERENCES

All links last accessed March 2019.

Djimeu, E.W. and Houndolou, D.-G. (2016). *Power calculation for causal inference in social science: sample size and minimum detectable effect determination.* 3ie Working Paper 26. New Delhi: International Initiative for Impact Evaluation (3ie). Retrieved from http://www.3ieimpact.org/evidence-hub/publications/working-papers/power-calculation-causal-inference-social-science-sample

Marsden, P.V. and Wright, J.D. (eds.). (2010). *Handbook of Survey Research* (2nd edition). Bingley: Emerald Group Publishing.

McConnell, B. and Vera-Hernández, M. (2015). *Going beyond simple sample size calculations: a practitioner's guide*, IFS Working Paper 15/17. London: Institute for Fiscal Studies (IFS). Retrieved from http://www.ifs.org.uk/uploads/publications/wps/WP201517_update_Sep15.pdf**Glennerster, R. and Takavarasha, K.** (2013). *Running Randomized Evaluations: A Practical Guide.* Princeton, New Jersey: Princeton University Press.

Oxfam GB. (2016). *How Are Effectiveness Reviews Carried Out?* Oxford: Oxfam. Retrieved from http://policy-practice.oxfam.org.uk/publications/how-are-effectiveness-reviews-carried-out-594353

Shannon, H.S., Hutson, R., Kolbe A., Stringer, B. and Haines, T. (2012). *Choosing a Survey Sample When Data on the Population are Limited: A Method Using Global Positioning Systems and Aerial and Satellite Photographs. Emerging Themes in Epidemiology* 9 (5). Retrieved from http://ete-online.biomedcentral.com/articles/10.1186/1742-7622-9-5

LINKS

Oxfam Research Guidelines: https://policy-practice.oxfam.org.uk/our-approach/research/research-guidelines

Planning Survey Research: http://policy-practice.oxfam.org.uk/publications/planning-survey-research-578973