Sampling of Sampling

Source: <http://www.stats.gla.ac.uk/steps/glossary/sampling.html>

Retrieved February 12, 2017 Spring 2017

|  |
| --- |
| STEPS Statistics Glossary |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Sampling

|  |  |
| --- | --- |
| [**Target Population**](http://www.stats.gla.ac.uk/steps/glossary/#targpop)[**Matched Samples**](http://www.stats.gla.ac.uk/steps/glossary/#matsamp)[**Independent Samples**](http://www.stats.gla.ac.uk/steps/glossary/#indepsamp)[**Random Sampling**](http://www.stats.gla.ac.uk/steps/glossary/#randsamp)[**Simple Random Sampling**](http://www.stats.gla.ac.uk/steps/glossary/#srs)[**Stratified Sampling**](http://www.stats.gla.ac.uk/steps/glossary/#stratsamp)[**Cluster Sampling**](http://www.stats.gla.ac.uk/steps/glossary/#clustsamp) | [**Quota Sampling**](http://www.stats.gla.ac.uk/steps/glossary/#quotasamp)[**Spatial Sampling**](http://www.stats.gla.ac.uk/steps/glossary/#spatsamp)[**Sampling Variability**](http://www.stats.gla.ac.uk/steps/glossary/#samplingvar)[**Standard Error**](http://www.stats.gla.ac.uk/steps/glossary/#stanerr)[**Bias**](http://www.stats.gla.ac.uk/steps/glossary/#bias)[**Precision**](http://www.stats.gla.ac.uk/steps/glossary/#prec) |

|  |
| --- |
| [Main Contents page](http://www.stats.gla.ac.uk/steps/glossary/index.html) | [Index of all entries](http://www.stats.gla.ac.uk/steps/glossary/alphabet.html)  |

steps**Target Population** The target population is the entire group a researcher is interested in; the group about which the researcher wishes to draw conclusions.*Example* Suppose we take a group of men aged 35-40 who have suffered an initial heart attack. The purpose of this study could be to compare the effectiveness of two drug regimes for delaying or preventing further attacks. The target population here would be all men meeting the same general conditions as those actually included in the study.steps**Matched Samples** Matched samples can arise in the following situations: 1. Two samples in which the members are clearly paired, or are matched explicitly by the researcher. For example, IQ measurements on pairs of identical twins.
2. Those samples in which the same attribute, or variable, is measured twice on each subject, under different circumstances. Commonly called repeated measures. Examples include the times of a group of athletes for 1500m before and after a week of special training; or the milk yields of cows before and after being fed a particular diet.

Sometimes, the difference in the value of the measurement of interest for each matched pair is calculated, for example, the difference between before and after measurements, and these figures then form a single sample for an appropriate statistical analysis.steps**Independent Sampling** Independent samples are those samples selected from the same population, or different populations, which have no effect on one another. That is, no correlation exists between the samples.steps**Random Sampling** Random sampling is a sampling technique where we select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has a known, but possibly non-equal, chance of being included in the sample.By using random sampling, the likelihood of bias is reduced.Compare [simple random sampling](http://www.stats.gla.ac.uk/steps/glossary/#srs).steps**Simple Random Sampling** Simple random sampling is the basic sampling technique where we select a group of subjects (a sample) for study from a larger group (a population). Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample. Every possible sample of a given size has the same chance of selection; i.e. each member of the population is equally likely to be chosen at any stage in the sampling process.Compare [random sampling](http://www.stats.gla.ac.uk/steps/glossary/#randsamp).steps**Stratified Sampling** There may often be factors which divide up the population into sub-populations (groups / strata) and we may expect the measurement of interest to vary among the different sub-populations. This has to be accounted for when we select a sample from the population in order that we obtain a sample that is representative of the population. This is achieved by stratified sampling.A stratified sample is obtained by taking samples from each stratum or sub-group of a population.When we sample a population with several strata, we generally require that the proportion of each stratum in the sample should be the same as in the population.Stratified sampling techniques are generally used when the population is heterogeneous, or dissimilar, where certain homogeneous, or similar, sub-populations can be isolated (strata). Simple random sampling is most appropriate when the entire population from which the sample is taken is homogeneous. Some reasons for using stratified sampling over simple random sampling are: 1. the cost per observation in the survey may be reduced;
2. estimates of the population parameters may be wanted for each sub-population;
3. increased accuracy at given cost.

*Example* Suppose a farmer wishes to work out the average milk yield of each cow type in his herd which consists of Ayrshire, Friesian, Galloway and Jersey cows. He could divide up his herd into the four sub-groups and take samples from these.steps**Cluster Sampling** Cluster sampling is a sampling technique where the entire population is divided into groups, or clusters, and a random sample of these clusters are selected. All observations in the selected clusters are included in the sample.Cluster sampling is typically used when the researcher cannot get a complete list of the members of a population they wish to study but can get a complete list of groups or 'clusters' of the population. It is also used when a random sample would produce a list of subjects so widely scattered that surveying them would prove to be far too expensive, for example, people who live in different postal districts in the UK.This sampling technique may well be more practical and/or economical than simple random sampling or stratified sampling.*Example* Suppose that the Department of Agriculture wishes to investigate the use of pesticides by farmers in England. A cluster sample could be taken by identifying the different counties in England as clusters. A sample of these counties (clusters) would then be chosen at random, so all farmers in those counties selected would be included in the sample. It can be seen here then that it is easier to visit several farmers in the same county than it is to travel to each farm in a random sample to observe the use of pesticides.steps**Quota Sampling** Quota sampling is a method of sampling widely used in opinion polling and market research. Interviewers are each given a quota of subjects of specified type to attempt to recruit for example, an interviewer might be told to go out and select 20 adult men and 20 adult women, 10 teenage girls and 10 teenage boys so that they could interview them about their television viewing.It suffers from a number of methodological flaws, the most basic of which is that the sample is not a random sample and therefore the sampling distributions of any statistics are unknown.steps**Spatial Sampling** This is an area of survey sampling concerned with sampling in two (or more) dimensions. For example, sampling of fields or other planar areas.steps**Sampling Variability** Sampling variability refers to the different values which a given function of the data takes when it is computed for two or more samples drawn from the same population. steps**Standard Error** Standard error is the standard deviation of the values of a given function of the data (parameter), over all possible samples of the same size.steps**Bias** Bias is a term which refers to how far the average statistic lies from the parameter it is estimating, that is, the error which arises when estimating a quantity. Errors from chance will cancel each other out in the long run, those from bias will not.The following illustrates bias and precision, where the target value is the bullseye:

|  |  |  |
| --- | --- | --- |
|  | **Precise** | **Imprecise** |
| **Biased** | load images | bi |
| **Unbiased** | up | ui |

*Example* The police decide to estimate the average speed of drivers using the fast lane of the motorway and consider how it can be done. One method suggested is to tail cars using police patrol cars and record their speeds as being the same as that of the police car. This is likely to produce a biased result as any driver exceeding the speed limit will slow down on seeing a police car behind them. The police then decide to use an unmarked car for their investigation using a speed gun operated by a constable. This is an unbiased method of measuring speed, but is imprecise compared to using a calibrated speedometer to take the measurement.See also [precision](http://www.stats.gla.ac.uk/steps/glossary/#prec).steps**Precision** Precision is a measure of how close an estimator is expected to be to the true value of a parameter.Precision is usually expressed in terms of imprecision and related to the standard error of the estimator. Less precision is reflected by a larger standard error.See the illustration and example under [bias](http://www.stats.gla.ac.uk/steps/glossary/#bias) for an explanation of what is meant by bias and precision. |