

Population Sampling: Probability and Non-Probability Techniques

Samuel J. Stratton, MD, MPH 

Stratton SJ. Population sampling: probability and non-probability techniques. *Prehosp Disaster Med.* 2023;38(2):147–148.

Keywords: convenience sampling; population sampling; purposeful sampling; random sampling; study design

doi:[10.1017/S1049023X23000304](https://doi.org/10.1017/S1049023X23000304)

© The Author(s), 2023. Published by Cambridge University Press on behalf of the World Association for Disaster and Emergency Medicine.

This discussion is a brief overview of population sampling in research. A target population is considered the entire group of subjects that are of interest for a study. It is rare that an investigator has access to an entire study target population. To address the issue of valid research findings for a target population, researchers often use one of a number of population sampling techniques or resort to analysis of a database that is assumed to be or known to be an accurate reflection of the study population. Sampling is the selection of a group of individual population members that represent the entire target population. Important is that the intent of sampling is to select study subjects that represent the entire study target population. The most valid form of sampling is one in which each individual member of the entire population has the same probability or likelihood for being selected as a member of the sample.

Typically, research and statistics textbooks include only a few pages of an entire text for discussion of sampling techniques. Ironically, a most common reason for having an otherwise well-conducted population-based study rejected for publication is lack of use of sampling techniques that limit unmeasurable process error and lack of representative validity of a target population.

A sample will likely represent an entire target population if two conditions exist: (1) a large enough sample; and (2) a sample that is formed using a random sampling technique. Random sampling means that every individual in a target population has the same probability of being selected into the study sample. There are a number of methods used for random sampling including systematic sampling, stratified sampling, random cluster analysis, computer generated subject assignment, and others that are beyond the scope of this discussion. Important is that only by the use of a random sampling method (probability sampling) can a researcher draw conclusions regarding an entire target population. Random sampling is often difficult, but efforts should be made to select study samples in a random manner such that non-participation bias and other measures of error and bias can be determined and factored into any findings of a research project. Random sampling is important as a method to assure researcher objectivity and lack of bias in selection of a study sample.

A common non-random population sampling method is the non-probability method called convenience sampling. Convenience sampling is the selection of study subjects because they are accessible for one reason or another to the researcher. This common form of sampling is often used for rapid market analysis or projection of political elections (both of which are most often invalid when final results are available). An example of convenience sampling is the stationing of a research associate at a train station with a survey conducted of passer by volunteer participants. In this example, the sample is not drawn at random from the target population, but by those who are conveniently at the train station and conveniently agree to participate as a sample member. Therefore, the findings of this form of research cannot be generalized to a cohort or population other than the sample group. A further issue with convenience sampling is that there is an inability to determine potential sampling error, selection bias, or motivation bias of participants to engage with the study. Convenience sampling occurs in a number of ways with a type encountered for submissions to *Prehospital and Disaster Medicine* being collection of data in a medical evacuation site, emergency department, or the field by researchers during convenient times of day as opposed to all hours of a day when potential subjects come and go and have different temporal attitudes.

A special form of convenience sampling is quota sampling in which sample subjects are selected from a baseline convenience sample that meet demographic characteristics of the target population. An example is forming a sample from an original convenience sample

group that matches the gender proportions of the study target population. Basically, the original sample is adjusted to match the desired gender categories by excluding those that are not of the gender desired once that gender category is filled. While quota sampling appears to be representative of a target population, it is based upon original convenience sampling technique and compounds selection bias. Therefore, quota sampling cannot be considered to represent a target population.

Another form of non-probability sampling is voluntary response sampling in which possible members of the target population are invited to participate in the research (usually online). Voluntary response sampling adds to the limitations of other forms of convenience sampling by incorporating self-selection bias. Voluntary sampling is often confused with census studies. Census research, which is actually not sample-based research, includes all members of the target population in the study. On the other hand, voluntary response sampling does not directly assure that all those in the target population receive the investigation material or are motivated to participate as a sample member. Most often, an organization member list is used to send invitations for study participation. Volunteer response sampling also occurs when supervisors in a target population group are asked to distribute materials or recruit study sample participants from within their organization.

Purposeful sampling is often used in phenomenon-based qualitative research. In this type of sampling, the researcher uses their own expertise or judgement to select a sample that may represent a target population. Unless the participants in a purposeful sample are selected using random sampling, purposeful sampling is a form of non-probability research. A common term to identify non-random selected purposeful samples is “subjective purposeful sampling.” Common errors in use of subjective purposeful sampling include researcher bias in selection of sample subjects, convenience selection of sample subjects, and selection of researcher colleagues for the sample.

As a non-probability form of sampling, subjective purposeful sampling cannot be considered representative of an entire target population and the results can only apply to the sample participants. To show possible generalizing of purposeful sampling to a target population, there must be explicit inclusion and exclusion

criteria used for the selection of the sample. The number of those who decline the invitation to participate should be reported to determine non-participant error. To aid readers of the research to appreciate the potential validity of a purposeful sample, a list of the sample members should be included with their education background, experience history, gender, age, and location or area of residence, as well as other appropriate indicators.

Often, purposeful or convenience sampling groups are expanded by snowball sampling. In this form of sampling, those who have been selected to participate in the sample are asked to recruit other potential sample members that they may be affiliated with. Snowball sampling is at high risk for selection bias and will further extend any sampling bias already present. Whether for expanding purposeful samples or convenience samples, the addition of snowball sampling adds further limitations to those that already exist for the original sampling technique.

Multistage sampling techniques include use of multiple sampling strategies to form a study sample group. Often, the original pool is selected by purposeful or convenience sampling and then randomly assigned to one or more study samples. These sample groups are subject to the same limitations as the non-probabilistic method originally used in the process. Basically, randomization of a non-random selected larger group is not a true random sampling method.

If non-probability sampling must be used for situations where there is not a form of random sampling possible, the non-probabilistic method used must be carefully selected. Comparison of the demographics of a non-probabilistic sample such as a purposeful sample to the known demographics of the target population is a technique used to associate a non-random sample with the overall target population. But, no matter how comparisons are made, the final research results of a non-random (non-probabilistic) selected sample cannot be assumed or stated to be representative of the target population and only represents the sample group.

This discussion presents a brief summary of population sampling techniques. Multiple resources are available that explore sampling in depth.¹⁻³ The most important concept presented in this discussion is that for a sample to represent a target population, the sample must be selected using a random sampling technique.

Additional Reading

1. Taherdoost H. Sampling methods in research methodology: how to choose a sampling technique for research. Social Science Research Network (SSRN); Elsevier. Published April 10, 2016. <http://doi.org/10.2139/ssrn.3205035>. Accessed February 24, 2023.

2. Etikan I, Sulaiman AM, Rukayya SA. Comparison of convenience sampling and purposive sampling. *Am J Theoretical and Applied Stats*. 2016;5(1):1–4.

3. Sharma G. Pros and cons of different sampling techniques. *International J Applied Research*. 2017;3(7):749–752.