The SAGE Encyclopedia of Social Science Research Methods

Representative Sample

http://dx.doi.org/10.4135/9781412950589.n853
The purpose of a sample design is to select a subset of the elements or members of a POPULATION inaway that STATISTICS derived from the SAMPLE will represent the values that are present in the population. In order to have a representative sample, a PROBABILITY method of selection must be used, and every element in the population has to have a known, nonzero chance of selection. When such a method is employed well, then the researchers can be confident that the statistics derived from the sample apply to the population within a stated range and with a stated degree of confidence.

The relevant population is theoretically defined on the basis of a specific research hypothesis. It consists of a set of elements that comprises the entire population, such as "adults residing in the United States." The easiest way to draw a probability sample is to have a list of all of the elements in the population, called a SAMPLING FRAME, and then apply a probability design to the sampling frame to produce a representative sample.

For many sampling problems, a list of all of the elements in the population exists, or at least a very good approximation of it does. In order to estimate the proportion of drivers who had an accident in the past year, a researcher might want to interview drivers and ask them whether they had an accident in the preceding 12 months. In order to produce a sample of drivers, for example, a researcher could start with a list of all individuals with a driver's license. The list might actually include all of those with licenses as of a recent date, and the list would have to be updated to include those who had received licenses since that date.

In another instance, a researcher might be interested in estimating how many American adults have a credit card. There is no existing list of all adults in the United States, so it has to be developed or approximated in the course of the study itself. This is a typical problem for those who conduct TELEPHONE SURVEYS. In this case, it is common to use a MULTISTAGE SAMPLING process that begins with a sample of telephone numbers. This is commonly generated by a computer to account **[p. 964 ↓]** for unlisted numbers and newly assigned ones. These numbers are called, and when a residence is contacted, all of the adults residing there are listed. Through this multistage process, a CLUSTER SAMPLING of adults residing in telephone households is produced; and

in the final step, one is selected at random as the respondent. That person is asked whether he or she has a credit card.

The key question for the researcher is how well the sample estimate (proportion of drivers who had a traffic accident in the past year or proportion of adults who carry a credit card) represents the population value. This is a function of two properties of the sample: its CONFIDENCE INTERVAL and confidence level. The primary advantage of using a probability method is that it allows researchers to take advantage of the NORMAL DISTRIBUTION that results from drawing repeated samples with a given characteristic from the same population. The properties of this SAMPLING DISTRIBUTION include information about the relationship between the standard deviation of the distribution and the size of the samples, as well as the distribution of the resulting sample estimates.

Given these properties, if we estimated from interviews with a simple random sample of 1,000 drivers that 14% of them had an accident in the past year, we would be 95% confident that the "true value" in the total population of all drivers ranged between 11% and 17% (the 14% estimate +/- the three-percentage-point margin of error). The application of the probability design to a good frame ensures that the result is a representative sample with known parameters that indicate its relative precision.

Michael W. Traugott

Kish, L. (1965). Survey sampling. New York: Wiley.

Levy, P. S. (1999). Sampling of populations: Methods and applications. New York: Wiley.

http://dx.doi.org/10.4135/9781412950589.n853

SSAGE researchmethods