

Encyclopedia of Survey Research Methods

Snowball Sampling

Contributors: James R. Chromy

Editors: Paul J. Lavrakas

Book Title: Encyclopedia of Survey Research Methods

Chapter Title: "Snowball Sampling"

Pub. Date: 2008

Access Date: October 14, 2013

Publishing Company: Sage Publications, Inc.

City: Thousand Oaks

Print ISBN: 9781412918084

Online ISBN: 9781412963947

DOI: <http://dx.doi.org/10.4135/9781412963947.n535>

Print pages: 824-825

This PDF has been generated from SAGE Research Methods. Please note that the pagination of the online version will vary from the pagination of the print book.

<http://dx.doi.org/10.4135/9781412963947.n535>

Snowball sampling is a technique that can be applied in two survey contexts. The first context involves surveying members of a rare population. The second involves studying mutual relationships among population members. In both cases, respondents are expected to know about the identity of other members of the same population group.

Studying Rare Populations

In this context, snowball sampling is a nonprobability sampling technique. The general objective is to identify members of the rare population. It involves identifying one or more members of a rare population and asking them to name other members of the same population. These additional persons are then contacted and asked to name additional persons in the rare population; and so forth. The process continues until an adequate sample size has been obtained or until no new names are elicited from the process.

If terminated when adequate sample size is obtained, the method yields a sample, but not a probability sample.

If the population can be restricted in some way, say to a limited geographic area such as a county, snowball sampling may be successful as a rare population frame-building technique. To be successful, several rounds of the process must be conducted, and the initial sample should be large and adequately distributed among the rare population members. Within this restricted population, the identified rare population members can then be sampled using probability sampling techniques, or a complete enumeration (census) may be conducted. If the limited geographic areas are first-stage units in a multi-stage probability sample design, the approach can yield an estimate for a larger target population.

If some members of the targeted rare population are isolated from the remainder of the population, they are not likely to be named even after several rounds of enumeration. Serious coverage problems may remain even if the process is carried out diligently.

[p. 824 ↓]

Studying Relationships

In the early 1960s, sociologist Leo Goodman proposed a probability sample-based method for studying relationships among individuals in a population. An initial zero-stage (Stage 0) probability sample is drawn. Each person in the sample is asked to name k persons with some particular relationship; example relationships are best friends, most frequent business associate, persons with most valued opinions, and so on. At Stage 1 these k persons are contacted and asked to name k persons with the same relationship. The Stage 2 sample consists of new persons named at Stage 1, that is, persons not in the original sample. At each subsequent stage, only newly identified persons are sampled at the next stage. The process may be continued for any number of stages, designated by s .

The simplest relationships involve two persons where each names the other. If the initial sample is a probability sample, an unbiased estimate of the number of pairs in the population that would name each other can be obtained. More complex relationships such as "closed rings" can be studied with more stages of sampling. For example, person A identifies person B; person B identifies person C; and person C identifies person A.

If the initial sample is drawn using binomial sampling so that each person has probability p of being in the sample and $s = k = 1$, an unbiased estimate of the number of mutual relationships in the population designated by M

11

is

$$\hat{M}_{11} = \frac{y}{2p},$$

where y is the number of persons in the Stage 0 sample who named a person who also names them when questioned either in the initial sample or in Stage 1.

The theory for estimating the population size for various types of interpersonal relationships has been, or can be, developed assuming binomial sampling and may

apply, at least approximately, when using other initial sample designs more commonly applied in practice, for example, simple random sampling (without replacement).

James R. Chromy

<http://dx.doi.org/10.4135/9781412963947.n535>

See also

Further Readings

Goodman L. A. Snowball sampling. *Annals of Mathematical Statistics* vol. 32 (1961) pp. 148–170.

Kalton G. and Anderson D. W. Sampling rare populations. *Journal of the Royal Statistical Society, Series A* vol. 149 (1986) pp. 65–82.