

LECTURE 22: QUOTA SAMPLING

Definition. In quota sampling, specified numbers (quotas) of particular types of population elements (subpopulations or strata) are required in the final sample. The choice of sampled elements is often entirely at the discretion of the interviewer, so that a sample of convenience is often chosen within each subpopulation.

Example. To obtain a quota sample of size 3000, it is specified that the sample contain 1000 white males, 1000 white females, 500 men of color, and 500 women of color, but no further instructions are given about how these quotas to be filled.

Thus, quota sampling is not a form of probability sampling – we do not know the probabilities with which each individual is included in the sample. It is often used when probability sampling is impractical, overly costly, considered unnecessary, or when people designing the survey just don't know any better.

Famous Example. In the 1948 U.S. presidential elections, all of the major polls printed just a few days before the election predicted that Dewey would defeat Truman by landslide. In fact, of course, Truman won the election. One of the problems of those polls was that they all used quota sampling.

“Dewey Defeats Truman” was a famously incorrect banner headline on the front page of the Chicago Tribune on November 3, 1948, the day after incumbent United States President Harry S. Truman beat Republican challenger and Governor of New York Thomas E. Dewey in the 1948 presidential election in an upset victory.



Harry S. Truman

“This is one for the books,” an elated **Thomas Edmund Dewey**
President Harry S. Truman proclaims
while holding up the Chicago Tribune
in this famous photograph.

ESTIMATING POPULATION SIZE

One of the most widely used method of estimating the size of a population is called two-sample capture-recapture estimation.

Example. Suppose we want to estimate the number of fish in a lake. The capture-recapture algorithm works as follows. Catch randomly and mark 200 fish in the lake, and then release them. Allow the marked and released fish to mix with the other fish in the lake. Then take a second random sample of 100 fish. Suppose that 20 of the fish in the second sample are marked. Then since 20% of the fish in the second sample are marked, we estimate that 20% of the fish in the lake were marked. Therefore, the 200 tagged fish in the original sample represent approximately 20% of the population of fish. The population size is then estimated to be approximately 1000 fish.

The two-sample capture-recapture method relies on the following assumptions:

1. The population is closed – no elements enter or leave the population, or have time to replicate.
2. Both samples are SRSs and are independent.
3. The tagged elements don't lose their markings.

SAMPLING RARE POPULATIONS

Sometimes investigators are interested in sampling rare or illusive (mobile) populations. For example, they might be interested in interviewing millionaires, or homeless people, or illegal drug users, or rare disease patients, or female smokers with two classes of education.

Several sampling approaches are commonly used in this case.

1. One of the approaches to sampling rare populations, which is costly and not very practical, is to take a very large sample. Nonresponse, though, is a special hazard in this case, because if the members of the population with the rare characteristic are more likely to be nonrespondents than members without the rare characteristic, then obtained estimates will be very biased. For example, in some health surveys, the characteristic itself can lead to nonresponse – a survey of cancer patients may have nonresponse because the illness prevents persons from responding. It is therefore important to try to minimize nonresponse for any survey of a rare population.
2. Sometimes strata can be constructed so that the rare characteristic is much more prevalent in one of the strata. Then a stratified sample in which the sampling fraction is higher in that stratum can give a more accurate estimate of the prevalence of the rare characteristic in the general population. This method employs stratified sampling with disproportional allocation.
3. Network sampling may be used, in which each sampled element can provide information also about its network. For example, each sampled household may provide information about the adult siblings of adult household members.
4. Snowball sampling is based on the premise that members of the rare population know one another, and will refer interviewers. For example, each interviewed homeless person may identify several other homeless people who will be subsequently interviewed.
5. Location sampling refers to methods used to sample individuals who visit specific locations such as night clubs, rock concerts, libraries, museums, shopping centers.