Sociology Shortcuts

M15. Representative Sampling

1. Simple Random Sampling

This type of sampling is similar to a lottery and is based on the probability the random selection of names from a sampling frame will produce a representative sample.

For a truly random sample everyone in the target population must have an equal chance of being chosen.

2. Systematic Sampling

This variation on the above - often used when the target population is very large - selects names systematically from a sampling frame. For example, choosing every 4th name on the frame.

This is sometimes called *quasi-random* sampling because, technically, not every name on the sampling frame has an equal chance of being chosen. It is, however, "random enough" for most samples.

Advantages

Some general advantages of both types:

- Relatively quick and easy and time-efficient ways to select a sample, especially when the target population is very large.

- Compared to other types, both are relatively inexpensive to create using a sampling frame accurate for the target population.

- Other than a way of identifying people in the target population (such as a name), the researcher doesn't require any other knowledge about this population.

- They produce random or near-random samples, based on *chance* (the sample cannot be accidentally biased by researcher selections).

Disadvantages

- An accurate, up-to-date, sampling frame may not be available for a target population.

- The time and cost involved in creating some sampling frames - and contacting those selected - may be prohibitive. This is particularly the case with large, diverse, target populations dispersed across a wide area.

- If the target population is *not homogeneous* (it doesn't consist of people who share whatever characteristics are important to the research) a biased sample can easily occur.

- Sampling based on simple chance may not produce a representative sample. Sample selection bias a particular problem with diverse target populations containing a mix of large and small groups.

- Simply selecting people from a sampling frame doesn't guarantee they will appear in your research. The random selection of doctors from an accurate frame, for example, may be unrepresentative if those selected can't be contacted or if they are unwilling to be part of the sample.

3a. Stratified Random Sampling

One way around the *diversity problem* that occurs in some simple random samples - where large groups in a target population may be *over-represented* and smaller groups *under-represented* - is use stratified random sampling.

This keeps the idea of sample selection based on chance and resolves the diversity problem by dividing (*stratifying*) the target population into groups whose characteristics are *known* to the researcher (such as "males and females" or different age groups). *Each group* is then treated as a *random sample in its own right.*

For example, a stratified sample based on *gender* might involve:

- a target population of 100 people (80 females and 20 males)
- a 10% sample
- two groups (males and females)
- select 8 women from female-only group and 2 men from male-only group
- combine the two samples to get a final random sample fully representative of the target population.

3b. Stratified Quota Sampling

This variation involves the same basic technique as stratified random sampling, but with two main differences:

1. With *stratified random sampling* the researcher needs a sampling frame from which to identify, select and contact people in the sample.

While this is always useful, it's *not strictly necessary* for quota sampling. It's enough to know the characteristics of a target population in order to construct a sample. In the above example, we just need to know the female-to-male ratio in our target population is 80%-20%.

2. Sample selection is done on an opportunity basis; the researcher works their way through the group of 20 males, asking each *in turn* to be part of the sample. Once 2 have agreed, the *male-only quota* for the sample is complete and no further males are selected.

Advantages

Known differences in the target population will be accurately reflected in the sample; we can, therefore, be sure our sample will be broadly representative of the target population.

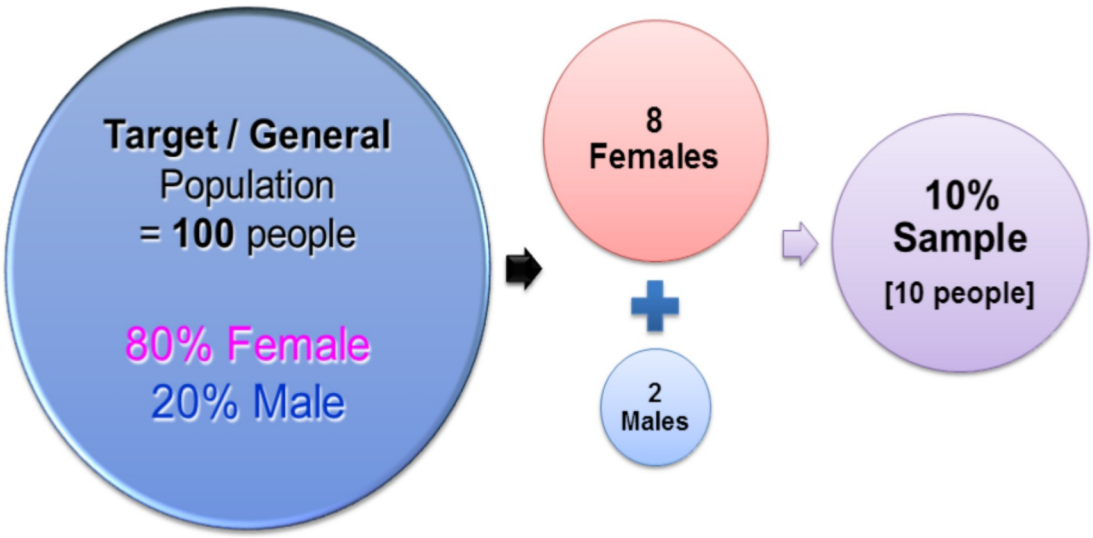
Given the high level of representativeness it's possible to generalise from the sample to the target population, even when - as with stratified quota sampling - the sample is generally small.

Stratified samples can be relatively small precisely because it's possible to ensure an accurate reflection of the target population.

The researcher can focus their sample on relevant distinctions in the target population (age, gender, class, ethnicity, etc.) and ignore irrelevant factors.

In terms of resources, quota samples are usually cheap and quick to construct accurately (which is partly why they're used by political polling organisations).

Where sample selection is based on a probabilistic method (the increased probability of creating a representative sample) it increases the likelihood that any statistical conclusions drawn from the sample will be valid.



Disadvantages

One potential drawback is that this technique isn't truly random in its sample selection: not everyone in the target population has an equal chance of being selected. The first person asked, for example, has a greater chance than the last.

A more-serious problem with stratified samples is that they are critically dependent on the accuracy of information about the target population. If the target population isn't accurately modeled (the weighting given to categories like gender or ethnicity, for example, is incorrect) the sample cannot be representative.

Even in situations where accurate information is available, this information may be outdated by the time the research is actually done. This is especially true where the sample is large and complex or where the composition of the target population may change rapidly – age groups, for example, may change on a daily basis.

When using a team of researchers to construct a quota sample you can't be certain they have correctly placed everyone in the right quota category.

If, for example, your research assistant cannot find "100 men over the age of 65" to fill their quota, there may be a temptation to fill it using men under that age - which would again affect the representativeness of the sample.

Stratified sampling can be hugely-expensive, time-consuming and mathematically-complex if the researcher is dealing with large, diverse, target populations. In our previous example we had a small target population and two clearly-delineated gender groups which makes for relatively simple sampling.

Now imagine adding variables such as class, age and ethnicity into the gender mix: not only are these *complex to define*, they also multiply the number of different sampling groups involved. This adds to both sample size and overall research costs.

As with simple random and systematic sampling, where a sampling frame is required it may be difficult to:

- find - a list encompassing all a researcher's target population may not exist.
- compile - where target populations are large, diverse and complex creating a sampling frame may, as we've suggested, be difficult, expensive and out-of-date by the time it is applied.
- access - there may be complex ethical issues surrounding access to and use of a target population list.