

1.4: Sampling Methods

Objectives

- Identify biased samples
- Distinguish between methods of sampling
- Distinguish between random sampling and random assignment

Why We Sample

Sampling plays a significant role in inferential statistics. Keeping in mind that our goal is to use data from a sample to infer about the larger population, we must ensure that our sample is representative by selecting it to be sufficiently large and without any systematic biases. There are many ways to sample; some are better than others.

Simple Random Sampling

Researchers adopt a variety of quality sampling strategies. The most straightforward is **simple random sampling**. Such sampling requires every member of the population to have an equal chance of being selected for the sample. In addition, the selection of one member must be independent of the selection of every other member. That is, choosing one member from the population must not increase or decrease the probability of picking any other member (relative to the others). In this sense, we can say that simple random sampling chooses a sample by pure chance.

? Text Exercise 1.4.1

What is the population? What is the sample? Was the sample picked by simple random sampling? Is the sample biased?

A research scientist is interested in studying the experiences of twins raised together versus those raised apart. She obtains a list of twins from the National Twin Registry and selects two subsets of individuals for her study. First, she chooses all those in the registry whose last name begins with *Z*. Then she turns to all those whose last name begins with *B*. Because there are so many names that start with *B*, our researcher decided to incorporate only every other name into her sample. Finally, she mails out a survey and compares characteristics of twins raised apart versus together.

Answer

The population consists of all twins recorded in the National Twin Registry. It is important that the researcher only make statistical generalizations to the twins on this list, not to all twins in the nation or world. That is, the National Twin Registry may not be representative of all twins. Even if inferences are limited to the Registry, a number of problems affect the sampling procedure we described. For instance, choosing only twins whose last names begin with *Z* does not give every individual an equal chance of being selected into the sample. Moreover, such a procedure risks over-representing ethnic groups with many surnames that begin with *Z*. There are other reasons why choosing just the *Z*'s may bias the sample. Perhaps such people are more patient than average because they often find themselves at the end of the line! The same problem occurs with choosing twins whose last name begins with *B*. An additional problem for the *B*'s is that the "every-other-one" procedure (called systematic sampling) disallowed adjacent names on the *B* part of the list from being both selected. Just this defect alone means the sample was not formed through simple random sampling.

Sample Size Matters

Recall that the definition of a simple random sample is a sample in which every member of the population has an equal chance of being selected. The sampling procedure defines what it means for a sample to be random, not the results. Random samples, especially if the sample size is small, are not necessarily representative of the entire population. For example, if a simple random sample of subjects was taken from a large enough population with an equal number of males and females, it would be about 10 times more likely that the sample consisted of 80% women if we only sampled 10 people as opposed to 20 people (4.3945% vs 0.4621%). A sample consisting of 80% women would not be representative, although the sample would be drawn randomly. Large sample sizes make it more likely that our sample is close to representative of the population. For this reason, inferential statistics takes into account the sample size when generalizing results from samples to populations. In later chapters, we will see what kinds of mathematical techniques ensure this sensitivity to sample size.

Other Sampling Methods

Our goal in constructing a sample is to arrive at a representation that yields accurate inferences regarding the population. The simplest way to guarantee that we are not systematically biased in our sampling methodology is to use a simple random sample. At times, we are aware that our population has distinct groups that differ from each other in a significantly relevant way to the topic at hand. If this were the case, we would want to ensure that each of these distinct groups would be represented in our sample. How could we guarantee appropriate representation without systematically biasing our sample?

Stratified random sampling can be an effective sampling method to guarantee the representation of different groups in a population that has natural differences. These distinct groups, known as strata, are each randomly sampled so that their sizes in the sample are proportional to their sizes in the population.

? Text Exercise 1.4.2

Suppose we were interested in views of capital punishment at an urban university. We have the time and resources to interview 200 students. The student body is diverse with respect to age; 30% of students are older people who work during the day and enroll in night courses (average age is 39), while 70% of students are younger students who generally enroll in day classes (average age of 19). It is possible that night students have different views about capital punishment than day students. How could we use stratified sampling to get a random sample?

Answer

Since 70% of the students are day students, it makes sense to ensure that 70% of the sample consists of day students. Thus, our sample of 200 students would consist of $140 (.7 \cdot 200 = 140)$ day students chosen at random and $60 (.3 \cdot 200 = 60)$ night students also chosen at random. The proportion of day students in the sample and in the population (the entire university) would be the same. Inferences to the entire population of students at the university would be more secure.

Simple random sampling ensures that any bias is due to random chance, but every possible sample is equally likely to occur. This means that we could end up with a sample that is difficult to collect, which makes data collection quite costly and time-consuming. We would save a lot of time and money if we could easily get a representative, random sample. In general, this isn't possible, but if we have more information about our population, we may be able to devise a better sampling strategy than simple random sampling. What would be necessary about our population if such a sampling method were to be effective? We would need the population to be well-mixed. This means that we can divide the population into sections such that each section does not have any significant differences from other sections regarding the topic at hand.

Cluster random sampling is a method that is used when a population is divided naturally into smaller groups (clusters), and each group does not have any significant differences from the others. Once these groups are created, we randomly select a set of clusters. We differentiate **two types of cluster sampling** based on how the clusters are studied once randomly selected. In **single-stage cluster sampling**, every member of each of the randomly selected clusters is studied. In **double-stage cluster sampling**, a simple random sample is taken from each randomly selected cluster. Double-stage cluster sampling aids efficiency and cost management, but single-stage cluster sampling is preferred since it includes more members.

Both cluster and stratified sampling divide the population into groups and select from those groups. The difference between them is that in a stratified sample, every group is selected, whereas in a cluster sample, only some of the groups are selected.

? Text Exercise 1.4.3

Suppose we are interested in voters' views regarding a school bond for the local municipal high school. Is the use of cluster sampling to obtain a random sample appropriate? If so, how could single-stage cluster sampling be implemented?

Answer

The population would be all voters registered to vote in the city. A natural partitioning of the population would be the municipal voter precincts. Since there is just one municipal high school, there would not be competition between parents of different schools vying for more money for their particular high school. We might expect tension regarding a school bond to depend on the age and presence of children. If this highlights a major difference between precincts, to cluster using them as sections would be inappropriate. While it is true that there are different types of neighborhoods, we would expect

families at different stages of life to live in each precinct. From this, there are no major differences between precincts regarding the school bond; we can implement the method.

We could randomly select 10 precincts from the total precincts in the municipality and then survey all voters in each precinct. This would be faster and much more efficient than randomly selecting voters from every precinct (stratified sampling) and randomly selecting voters regardless of any other factors (simple random sampling).

? Text Exercise 1.4.4

We are interested in determining the average height of students at our local high school. Despite being able to measure every student, we deem conducting a census not practical. Consider both methodologies (stratified and cluster sampling) for this population. Express your thoughts.

Answer

Note answers may vary. There are two natural groupings that immediately come to mind when thinking of high school students: class rank and gender. Are there major differences in heights between men and women? Yes, men tend to be taller than women on average. This means that cluster sampling by gender would be inappropriate. Similarly, there are major differences in class rank because students generally continue to grow throughout all of high school. Thus cluster sampling by class rank would be inappropriate. As such, it seems that we have identified eight natural strata (each class split by gender), meaning stratified sampling would be appropriate.

Perhaps the high school has a "homeroom" system that groups students across ages and genders into similar groups. Such homerooms could be fine candidates for clusters, depending on how they were constructed.

While our goal is to get a representative sample, the best we can do is to guarantee that any bias in the sample is due to random chance. Inferential statistics is built upon this framework. We must be wary of sampling methods that admit systematic bias. We have already encountered several of them in the course of this book: **voluntary response** (coach with cartwheels), **convenience** (students in the front row), and **systematic** (choosing every other last name starting in B). Note that in systematic it need not be every other member of the population but every k^{th} member.

? Text Exercise 1.4.5

Construct definitions for the voluntary response and convenience sampling methods. Explain how they are related and how to distinguish between them. Discuss why the methodologies produce biased samples most of the time.

Answer

Voluntary Response: A form of sampling in which a mass request is sent out or posted asking for participation in the sample. Any member of the population who receives the request and volunteers for the study will be included in the sample.

Convenience: A form of sampling in which population members are identified and selected for the sample simply because some aspect makes collecting data easier.

Both sampling methods are based on ease of access. Convenience sampling has a much broader application than voluntary response because convenience sampling may have pretty much anything as the population. Voluntary response sampling requires that the population consists of people. However, there are convenient samples of people that fail to be voluntary response. Consider a psychology professor conducting a survey of his psychology students as a sample of the student body instead of soliciting survey participants via school-wide emails. The latter is a voluntary response, while the former is merely convenience. Also, consider studying the use of turn-signals by standing at a busy intersection and counting instances over a given period of time. It is convenient to sample drivers at a single location, and participation was not voluntary. The key to remember is that voluntary response requires self-selection on the participant's part.

Voluntary response sampling generally produces biased samples because people who feel strongly, either positively or negatively, are more prone to respond to requests.

Convenience sampling generally produces biased samples because the sample is convenient for a reason; some characteristics are common among them. Thus, the subjects that do not possess that characteristic are most likely underrepresented.

In both of these methodologies, the sample produced may be representative. We cannot confirm when this is the case, and we cannot assert that any bias is due to random chance because of the previously mentioned reasons. These methods are often used, and sometimes for good reason. Doing rigorous sampling can be costly or logistically impossible. However, skepticism is justified when assessing conclusions drawn about a population based on a convenience sample.

Random Assignment in Medical Trials

In experimental research, populations are often hypothetical. For example, in an experiment comparing the effectiveness of a new anti-depressant drug with a placebo (fake treatment), there is no actual population of individuals taking the drug. In this case, a specified population of people with some degree of depression is defined and a random sample is taken from this population. The sample is then randomly divided into two groups; one group is assigned to the treatment condition (taking the drug) and the other group is assigned to the control condition (taking the placebo). This random division of the sample into two groups is called **random assignment**. Random assignment is critical for the validity of an experiment.

For example, consider the bias that could be introduced if the first 20 subjects to show up at the experiment were assigned to the experimental group and the second 20 subjects were assigned to the control group. It is possible that subjects who show up late tend to be more depressed than those who show up early, thus making the experimental group less depressed than the control group even before the treatment was administered.

In experimental research of this kind, failing to assign subjects randomly to groups is generally more serious than having a non-random sample. Failure to randomize (the former error) invalidates the experimental findings, while a non-random sample (the latter error) simply restricts the degree to which the results are generalizable.

1.4: Sampling Methods is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by The Math Department at Fort Hays State University.

- **1.4: Inferential Statistics** by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.