Terminology

Statistics is the science of collecting, summarizing, analyzing, and drawing conclusions from data.

One of the main objectives of statistics is to make an inference about a population of interest based on information obtained from a sample of measurements from that population.

The population in a statistical study is the entire collection of individuals about which we want information.

Individuals are the objects described by a set of data. Individuals may be people, but they may also be animals or things.

A sample is a part of the population from which we actually collect information, used to draw conclusions about the whole.

Example: What is the average height for the population of all female students at the university of Idaho.

More Terminology

Possible Statistical Studies

EXAMPLE A biologist observes each of a number of beetles and records the sex, weight, length, width, and other characteristics of a species of beetle in order to identify it and distinguish it from other species.

Example A political scientist records the attitudes of voters concerning some local issues in order to predict upcoming voting results.

An individual is the object that we observe.

An observation is the information recorded for each individual.

Example The biologist observes each of a number of beetles and records the sex, weight, length, and width for each. A single beetle is an individual. The information recorded for each beetle is the observation.

Example The political scientist interviews a sample of 1,000 people from a local electorate and records the opinion of each person concerning an issue on the ballot. A single person is an individual and the person’s opinion is the observation.

A characteristic of the individuals that varies from one individual to another is called a variable.

Example – weight of a beetle
Example – sex of a beetle
Example – voter opinion on a particular issue
A collection of observations on one or more variables is called data.

A categorical variable places an individual into one of several groups or categories. Observations made on a qualitative variable are called categorical data.

**Example** A beetle’s sex is a qualitative variable.

**Example** A person’s political affiliation is a qualitative variable.

A quantitative variable takes numerical values for which arithmetic operations such as adding and averaging make sense. Observations made on a quantitative variable are often called quantitative data.

**Example** The weight, length, and width of a beetle are examples of quantitative variables.

### Collecting Data

In an experiment, the researcher controls or manipulates the environment of the individuals. The intent of most experiments is to study the effect of changes in one variable (such as hours without sleep) on another variable (such as reaction time).

**Example** A psychologist conducted a study of the effect of propaganda on attitude toward a foreign government. There were 100 subjects available. 50 of the subjects were randomly assigned to group 1 and the remaining 50 were assigned to group 2. All 100 subjects were tested for attitude toward the German government. The 50 subjects in group 1 then read German propaganda regularly for several months, while the 50 subjects in group 2 were instructed not to read German propaganda; otherwise both groups went about their normal lives. Then all 100 subjects were retested. Group 1 showed a more positive change in attitude toward Germany between test and retest. This change can be attributed to reading the propaganda.

<table>
<thead>
<tr>
<th>Random Allocation</th>
<th>Test of attitude</th>
<th>Reading of Propaganda</th>
<th>Retest of attitude</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test of attitude</td>
<td></td>
<td>Retest of Attitude</td>
</tr>
</tbody>
</table>

**Randomization:** The purpose of randomization is to create groups that are equivalent prior to the experiment. Many factors (sex, age, race, religion, political opinion) may influence a subjects’ reaction to German propaganda. Randomization produces groups that are similar with respect to these extraneous factors as well as other extraneous factors we have not thought of.

**Example** Suppose we randomly assign each of a number of new babies into one of two groups. Without an intervention, we should expect both groups to gain about the same amount of weight, on average. If we then expose one group to the sound of a heartbeat and that group gains significantly more weight than the other group, we can be reasonably certain that the weight gain was due to the sound of the heartbeat.

**Example** It has been claimed, most notably by the double Nobel laureate Linus Pauling, that large doses of Vitamin C will prevent colds. An experiment to test this claim was performed in Toronto in
the winter of 1971-1972. About 500 volunteers were randomly allocated to each of two groups. Group 1 received 1 gram per day of Vitamin C, and 4 grams per day at the first sign of cold. Group 2 served as a control group, and received a placebo pill identical in appearance to the vitamin C capsules. At the end of the winter, the percentage not having a cold was determined for each group.

Example Does aspirin prevent heart attacks? In 1988, the Steering Committee of the Physicians’ Health Study Research Group released the results of a 5-year experiment conducted using 22,071 male physicians between the ages of 40 and 84. The physicians had been randomly assigned to two groups. One group took an ordinary aspirin tablet every other day while the other group took a “placebo”, a pill designed to look just like an aspirin but with no active ingredients. Neither group knew whether or not they were taking the active ingredient.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Heart Attack</th>
<th>No Heart Attack</th>
<th>Attacks Per 1,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>104</td>
<td>10,933</td>
<td>9.42</td>
</tr>
<tr>
<td>Placebo</td>
<td>189</td>
<td>10,845</td>
<td>17.13</td>
</tr>
</tbody>
</table>

The rate of heart attacks in the group taking aspirin was only 55% of the rate of heart attacks in the placebo group. Because the men were randomly assigned to the two conditions, other factors such as the amount of exercise should have been similar for both groups. The only substantial difference in the two groups should have been whether they took the aspirin or the placebos. Therefore, we can conclude that taking aspirin caused the lower rate of heart attacks for that group.

Example Experiments that study the effectiveness of medical treatments on actual patients are called clinical trials. The clinical trial that made gastric freezing a popular treatment for ulcers had this design:

```
Impose treatment          Measure Response
Gastric Freezing       → Reduced Pain?
```

The patients did report reduced pain, but we can’t say that gastric freezing caused the reduced pain. It might be just a placebo effect. A placebo is a dummy treatment with no active ingredients. Many patients respond favorably to any treatment, even a placebo. With the design above, the placebo effect was confounded with any effect gastric freezing might have.

Better Design: A second clinical trial, done several years later, divided ulcer patients into two groups. One group was treated by gastric freezing as before. The other group received a placebo treatment in which the solution in the balloon was at body temperature rather than freezing. The results: 34% of the 82 patients in the treatment group improved, but so did 38% of the 78 patients in the placebo group. This and other properly designed experiments showed that gastric freezing was no better than a placebo, and doctors abandoned it.
Summary of Experimentation

A researcher wants to know the effect of a treatment, like the Salk vaccine, on a response, like getting polio. To find out, he or she can compare the responses of a treatment group, which gets the treatment, with those of a control group, which doesn’t.

If the treatment group is just like the control group, apart from the treatment, then an observed difference in the responses of the two groups is likely to be due to the effect of the treatment. However, if the treatment group is different from the control group with respect to other factors as well, the observed difference may be due in part to these other factors. The effects of these other factors are confounded with the effect of the treatment.

The best way to try and make sure that the treatment group is like the control group is to assign subjects to treatment or control at random.

Whenever possible, in a well-designed experiment the control group is given a placebo, which is neutral but which resembles the treatment. This is to make sure that the response is to the treatment itself rather than to the idea of the treatment.

A well-designed experiment is run double-blind whenever possible. The subjects do not know whether they are in the treatment or the control group. Neither do those who evaluate the responses. This guards against bias, either in the responses or the evaluation.

Experiments are not Always Possible

Example Suppose we want to study the effect of cigarette smoking on lung cancer and cardiovascular disease. One possible experimental design would be to randomize a fixed number of individuals (say, 1000) to each of two groups – one group would be required to smoke cigarettes for the duration of the study (say, 10 years), while those in the second group would not be allowed to smoke throughout the study. At the end of the study, the two groups would be compared for lung cancer and cardiovascular disease.

Problems:
1. Not ethical
2. It would be difficult to follow all participants and make certain that they follow the study plan.
3. It would be difficult to find nonsmoking individuals willing to take the chance of being assigned to the smoking group.

Another possible study would be to sample a fixed number of smokers and a fixed number of nonsmokers to compare the groups for lung cancer and for cardiovascular disease. This is an example of an observational study.

In an Observational Study, the researcher observes individuals and measures variables of interest but does not attempt to manipulate the environment of the participants. All observational studies share the principle “observe but don’t disturb.”

What has been sacrificed with an observational study?
The fundamental difference between an observational study and an experiment lies in the inference(s) that can be drawn.

**Drawing Conclusions in an Experiment:** For an experiment comparing smokers to nonsmokers, assuming the two groups of individuals follow the study plan, the observed differences between the smoking and nonsmoking groups could be attributed to the effects of cigarette smoking because individuals were randomized to the two groups; hence, the groups were assumed to be comparable at the outset.

**Drawing Conclusions in an Observational Study:** This type of reasoning does not apply to the observational study of cigarette smoking. Differences between the two groups could not necessarily be attributed to the effects of cigarette smoking because, for example, there may be hereditary factors that predispose people to smoking and cancer of the lungs and/or cardiovascular disease. Thus, differences between the groups might be due to hereditary factors, smoking, or a combination of the two.

Typically, the results of an observational study are reported by way of a statement of association. For example, if the observational study showed a higher frequency of lung cancer and cardiovascular disease for smokers relative to nonsmokers, it would be stated that this study showed that cigarette smoking was associated with an increased frequency of lung cancer and cardiovascular disease. It is careful rewording in order not to infer that cigarette smoking causes lung cancer and cardiovascular disease.

**Example** An article in a women’s magazine reported that women who nurse their babies feel warmer and more receptive toward the infants than mothers who bottle-feed. The author concluded that nursing has desirable effects on the mother’s attitude toward the child.

**Problem with this conclusion:** Women choose whether to nurse or bottle feed, and it is possible that those who already feel receptive toward the child choose to nurse, while those to whom the baby is a nuisance choose the bottle.

The effect on the mother’s attitude of the method of feeding is **confounded** with the already existing attitude of the mother toward the child.

The effects of two variables on another are said to be **confounded** when they cannot be distinguished from each other.

**SAMPLE SURVEY**

Sample surveys are used to describe a population.

In a sample survey, we use information contained in a sample to draw conclusions about the population from which the sample was obtained.

**Example** The student senate at a university with 15,000 students is interested in the proportion of students who favor a change in the grading system to allow + and – grades. Two hundred students are interviewed to determine their attitude toward this proposed change.

**Population:** All 15,000 students
Sample: The 200 students interviewed
Example: Representatives of the insurance industry wished to investigate the monetary loss due to damage to single-family dwellings in Pasadena, California, resulting from an earthquake that occurred on December 3, 1988. One hundred homes were selected for inspection from the set of all single-family homes in Pasadena, California.

Population: All single family homes in Pasadena California.
Sample: The 100 homes that were selected for inspection

Example: A building inspector has a chance to buy an odd lot of 5,000 used bricks at an auction. She is interested in determining the proportion of bricks in the lot that are cracked and therefore unusable for her current project, but she does not have enough time to inspect all 5,000 bricks. Instead, she checks 100 bricks to determine whether each is cracked.

Population: All 5,000 bricks in the lot.
Sample: The 100 bricks selected for inspection.

How should we select a sample?

Should We Use Voluntary Response?

Voluntary response to a mail or television questionnaire elicits the opinions of those who feel strongly enough to respond. The opinions of this group are often systematically different from those of the population as a whole.

For example, a 1972 local television station asked viewers to send postcards indicating whether they favored or opposed President Nixon’s decision to mine Haiphong harbor. Of those who responded 5157 agreed with the decision and 1158 disagreed. It later developed that Nix campaign workers had mailed about 2000 responses agreeing with his decision.

To avoid introducing bias, random sampling is used.

A simple random sample (SRS) of size n consists of n individuals from the population chosen in such a way that every set of n individuals has an equal chance to be the sample selected.

Choosing a Simple Random Sample

Step 1: Label. Assign a numerical label to every individual in the population. Be sure that all labels have the same number of digits.

Step 2: Table. Use random digits to select labels at random.

Example: A firm wants to understand the attitudes of its minority managers toward its system for assessing management performance. Below is a list of all the firm’s managers who are members of minority groups. Use Table B at line 139 to choose 6 to be interviewed in detail about the performance appraisal system.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Agarwall</td>
<td>08</td>
<td>Dewald</td>
<td>15</td>
<td>Huang</td>
</tr>
<tr>
<td>22</td>
<td>Puri</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example The student senate at a university with 15,000 students is interested in the proportion of all students at the school who favor a change in the grading system to allow + and – grades. To estimate the proportion of all students in favor of the change, two hundred students are randomly selected and interviewed to determine their attitude toward the proposed change. Of the 200, 120 (60%) said they were in favor. We can use 60% as an estimate of the proportion of all students at the school who favor the proposal.

In the above example, the proportion of all students at the school who favor a change in the grading system is a parameter.

A parameter is a number that describes the population. A parameter is a fixed number, but in practice we do not know its value.

In the above example, 60%, the proportion of the students in the sample who favored a change in the grading system, is a statistic.

A statistic is a number that describes a sample. The value of the statistic is known when we have taken a sample, but it can change from sample to sample. We often use a statistic to estimate an unknown parameter.

Example “Are you afraid to go outside at night within a mile of your home because of crime?” When the Gallup Poll asked this question, 45% of the people in the sample said “Yes”. The number 45% is a statistic. This statistic can be used to estimate the parameter of interest, the proportion of people in the population who are afraid to go outside within a mile of their home because of crime.

Example The mean height of the population of all female students at the University of Idaho is a parameter that we would like to estimate. To do so, we take a random sample of 100 female students from the population of all female students at the university. The average for the sample is 66.4 inches. 66.4 is a statistic.

Other Sampling Methods

Taking a simple random sample may not always be the best approach to sampling. The following are a few other sampling methods, all of which are good substitutes for simple random sampling in most situations. In fact, they often have advantages over simple random sampling.

Stratified random sampling: A stratified random sample is collected by first dividing the population of individuals into groups (strata) and then taking a simple random sample from each group. For example, public opinion pollsters often take separate samples from each region of the country so they can spot regional differences as well as measure national trends.
One advantage of using stratified random sampling over simple random sampling is that it allows us to gather separate information about each stratum.

**Systematic Sampling**: Suppose you had a list of 5,000 names and telephone numbers from which you wanted to select a sample of 100. This means you would want to select 1 out of every 50 people on the list. The first idea that might occur to you is to simply choose every 50th name on the list. If you did so, you would be using a systematic sampling plan.

To take a systematic sample, divide the list into as many consecutive segments as you need, randomly choose a starting point in the first segment, then sample at that point in each segment. In the above example, you would randomly choose a starting point in the first 50 names, then sample every 50th name after that. When you were finished, you would have selected one person from each of 100 segments, equally spaced throughout the list.

Problem with systematic sampling is that it can sometimes lead to a biased sample.

**Example** Suppose you were doing a survey of potential noise problems in a high-rise college dormitory. Further, suppose a list of residents was provided, arranged by room number, with 20 rooms per floor and two people per room. If you were to take a systematic random sample of, say, every 40th person on the list, you would get people who lived in the same location on every floor and thus a biased sampling of opinions about noise problems.

**Cluster Sampling**. Cluster sampling is often confused with stratified sampling, but it is actually a radically different concept. To take a cluster sample, divide the population into groups, called clusters, but rather than sampling within each group, we select a simple random sample of clusters and measure all individuals in those clusters.

**Example** Suppose we wanted to sample students living in the dormitories at a college. If the college had 15 dorms and each dorm had six floors, we could consider the 90 floors to be 90 clusters of units. We could then randomly select the desired number of floors and measure everyone on those floors. It would probably be much cheaper and more convenient than obtain a simple random sample of all dormitory residents.

One advantage of cluster sampling is that you only need a list of clusters, instead of a list of all individuals.