Sampling and Experimentation: Planning and Conducting Experiments

You can collect meaningful data in many ways. Statisticians use vocabulary that might sound similar but might have very different meanings. The meaning of common terminology is often confused. Use statistical vocabulary with care.

The two general types of statistical studies are **observational studies,** as discussed in the previous chapter, and **experiments.** When using an observational study, a researcher makes observation of the participants or subjects in the study. The researcher does not ask the subjects to change their behavior or try to control them in any way. The researcher simply observes and records those observations. The researcher then uses the data obtained from observing the sample to say something about the population from which the sample was drawn. In experiments, researchers have control over something, and they then try to measure the effect of what they are trying to control. A study actually becomes an experiment if the researchers do something to or for the subjects and observe the response.

Characteristics of a Well-Designed and Well-Conducted Experiment

When conducting observational studies, you must take care to limit bias. When conducting experiments, you must make sure that the response being measured is really caused by what the researcher is controlling. A well-designed experiment isolates that which is being studied through proper randomization techniques. Depending on the circumstances, researchers might choose between observational studies and experiments to collect their data. Many times, one method is preferred over the other. Sometimes, experiments are not possible.

EXAMPLE:
Consider the task a researcher might have to compare the effectiveness of two different diets. The researcher could survey a sample of people using diet number one and another sample of people using diet number two. The researchers could collect data from each sample concerning participants' starting weight, ending weight, and length of time on the respective diet. Comparisons then could be made and conclusions drawn from the data.

The drawback to this study is that there might be some unmeasured factors that will affect one set of dieters more than the other. Possibly one of the diets tries to control sugar intake more than the other diet. Therefore, more subjects who are diabetic might use the sugar restrictive diet than the other diet. Diabetics might have an easier, or harder, time losing weight on any diet, thus invalidating the results.

Proper experimental design could help eliminate this problem. Subjects who want to go on a diet could be randomly assigned to one of the two diets. By appropriate randomization techniques, the fact that some dieters are diabetic would not affect one diet more than the other; thus, this variable could be controlled.

EXAMPLE:
Suppose that a researcher is asked to determine whether more traffic accidents are caused by people who drink beer or people who drink wine. It would be difficult, if not illegal, to construct an experiment in which you ask some of your subjects to drink wine and some subjects to drink beer and then send them out driving so that you could tabulate the number of accidents they cause. In this case, as observational study would be more appropriate. Data about the type of alcohol consumed by drivers who caused accidents while driving under the influence of alcohol could be collected and analyzed.

Treatments, Control Groups, and Experimental Units

Researchers use different terms to describe the subjects of an experiment. The most common term is **experimental unit,** which represents a single entity. It can be a person, an animal, or a thing. When experimental units are people, the term subject or participant is usually used. The specific experimental condition that is being applied to these experimental units is called the treatment. Experimental units are divided into groups. Groups that receive the treatment are called **treatment groups.** Researchers create **control groups,** which are treated identically to all other groups with the exception that they do not receive the actual treatment.

EXAMPLE:
A researcher wants to determine whether listing the calorie content of desserts on the menu would influence a diner's dessert selection. Half of the parties (the treatment group) dining at an upscale restaurant are given menus with the dessert calorie content listed. Half of the parties (the control group) are given traditional menus. Appropriate random selection is used to determine which parties (experimental units) go into each group. By using good experimental design, other variables, such as the type of dinner ordered, can be controlled so that they do not affect one group more than the other group.

Random Assignments and Replication

Good experimental designs control variables that might have an influence on the outcome of the experiment. Good control is essential to be sure that any differences between the control group and the treatment group are caused only by the treatment variable and not any other variable. **Explanatory variables** explain or cause differences in the **response variable.** Two techniques to help control explanatory variables that are not treatment variables are random assignment and **replication.** Randomization is used to help balance groups so that the effect of these lurking variables is minimized. Random selection is used to determine which group, treatment, or control an experimental unit is placed in. Some individual experimental units are affected more by these lurking variables than others. By using random selection techniques, the overall groups can be balanced with respect to all variables except the treatment variable. Even with randomized selection, variation can occur. By using replication and repeating the experiment with as many experimental units as possible, these variations can be minimized.

EXAMPLE:
A teacher wants to determine whether a particular teaching approach will help improve student learning. The amount of previous knowledge a student has about a topic would be considered a lurking variable when using testing as a method to determine effectiveness. Replication could be used to balance these differences.

Researchers randomize with respect to *treatment type* and with respect to *treatment order*. If one treatment is being tested, then the experimental units need to be placed in either the treatment group or the control group.

The experimenter cannot be allowed to personally choose in which group to place each experimental unit. Some experimenters feel that if they equally divide experimental units with respect to as many lurking variables as possible and place half in the treatment group and half in the control group that they have done a good job of balancing. Good experimental design requires that random techniques, not experimenter choice, be used to determine the assignment of groups. The use of random number tables or statistical calculators with random number generators are the preferred way of choosing which experimental units get assigned to each group. If all subjects of an experiment are to receive all treatments, then the order of the treatments should be determined by randomization. If each subject in an experiment is to determine their preference for five different flavors of herbal tea, the order of tasting of the teas must be randomized.

Sources of Bias and Confounding, Including Placebo Effect and Blinding

When the response in an experiment is due to variables in addition to the treatment variable and the effect of these variables cannot be separated from the effect of the treatment variable, then these variables are referred to as **confounding variables.** That is, a confounding variable is a variable that has an effect on the response variable and is also related in some way to the explanatory (treatment) variable. The main problem with confounding variables is that they might have more of an effect that the treatment variable. A **lurking variable** looks like a confounding variable—that is, it could have an effect on the results of the experiment, but is not measured or used in the experiment. Proper randomization of experimental units should minimize or eliminate the effect of lurking variables. Proper experimental design helps control the effects of confounding variables. It is harder to control these variables in observational studies.

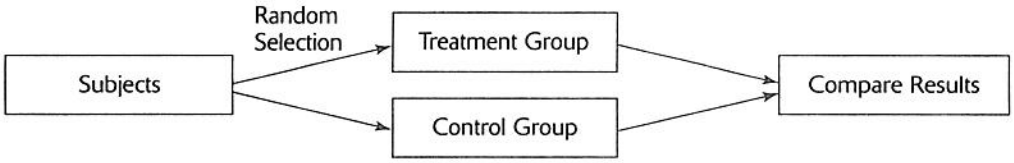
EXAMPLE:
A company that specializes in helping students prepare for the ACT advertises that, "If you took the ACT once and you did not like your score, take our course, and your score will improve." If you want to conduct an experiment to test this claim, you must be careful of a confounding variable. Suppose a student takes the ACT, does poorly, takes the preparation course this company is offering, repeats the test, and the score improves. Remove the confounding variable. Is the increase in score due to the preparation course or just the fact that these people have taken the test before? An increased comfort level with taking the test might contribute to an increased score. The experiment could be structured to control for this confounding variable.

Many statistical studies involve testing the effectiveness of drugs. A **placebo** looks identical to the actual drug but contains no active ingredient and so has no real physical effect. Humans want to be helped by the medication that is administered to them. If they think they are receiving a drug to help their condition, they tend to improve even if it turns out that the drug is a placebo. This is called the **placebo effect.** Since patients tend to respond to any treatment administered by their doctors, even a placebo, researchers must design experiments to control the placebo effect. Researchers randomly divide the patients into two groups. One of the groups receives the real medication and is referred to as the treatment group. The other group receives the placebo and is referred to as the control group. The patients are not told into which group they have been placed so they will not be influenced by their group association.

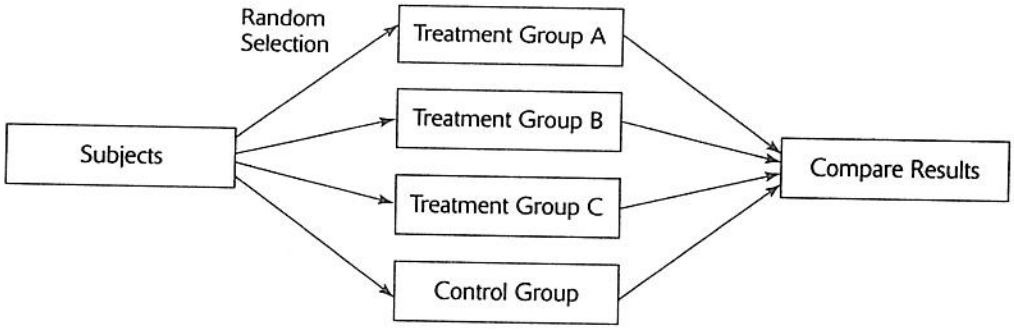
The two main sets of participants in experiments are the subjects, or experimental units, and the researchers. Well-designed experiments are **double-blind.** In a double-blind experiment, neither the subjects nor the researchers know to which group, treatment, or control, subjects have been assigned. If a researcher knows that a subject is in the control group, they do not expect a treatment effect, and their measurement of a response might be understated. If a researcher knows that a subject is in the treatment group, they might overstate a response simply because they expect it. An experiment might also be single-blind. In this case, only one of the participants, either the subjects or the researchers, knows to which group the subjects have been assigned.

Completely Randomized Design

If all the experimental units (subjects of the experiment) are randomly assigned to either the control group or to the treatment group, then the experiment has a **completely randomized design.** The following diagram illustrates a completely randomized design with one treatment group and one control group.



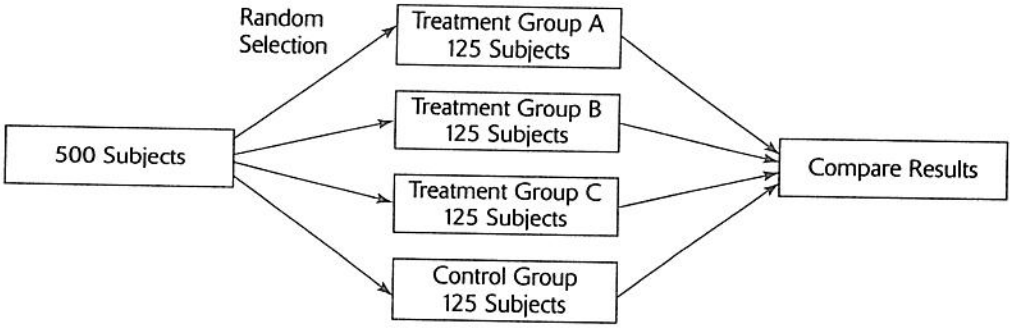
If the experiment involves more than one treatment group—for example, three treatment groups—the diagram may be modified as follows:



It is important to remember that the control group, receiving no treatment or a placebo that looks like a treatment, is actually a treatment group.

EXAMPLE:
A study is being conducted to compare the effectiveness of three medications to reduce cholesterol levels in adults. Five hundred subjects have volunteered to participate in this study. Design a double-blind, completely randomized experiment for testing the effectiveness of these three medications.

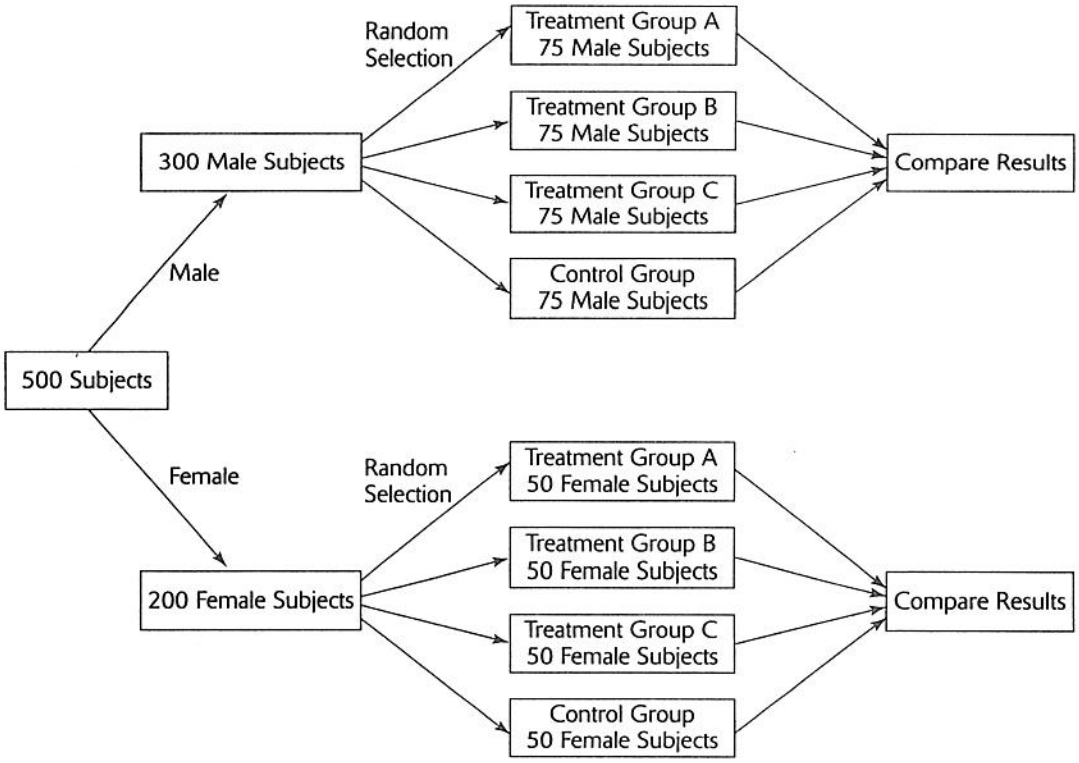
First, you must decide whether to include a control group in which volunteers will be given a placebo instead of one of the three medications being tested. If only one medication was being tested, a placebo-based control group would be necessary in order to be able to make a comparison. In this case, you could compare the effectiveness of the three medications to each other without a control group. We will choose to include a placebo-based control group. Medical personnel will be in contact with the subjects to run blood tests to monitor their progress. To be a double-blind experiment, both the subjects and the medical personnel must not know into which group the subjects have been placed.



Randomized Block Design, Including Matched Pairs Design

Completely randomized designs work well when variability among experimental units is unknown or minimal. If known variability exists among the subjects of the experiment and the effect of these variables cannot be separated from the treatment variables, then these variables are known as confounding variables. **Blocks** should be used to isolate these sources of variability among the subjects when these confounding characteristics can be identified for each subject.

EXAMPLE:
The experimental design in the preceding example does not take into account any known variability among the subjects. If, for example, it is known that the medications are more effective in men, then gender becomes a confounding variable and can be controlled for with the use of **blocking** since the gender of each subject can be determined. Men and women subjects are separated into two blocks, and a completely randomized design is used for each block. If the 500 subjects consisted of 300 men and 200 women, the following block design might be appropriate:

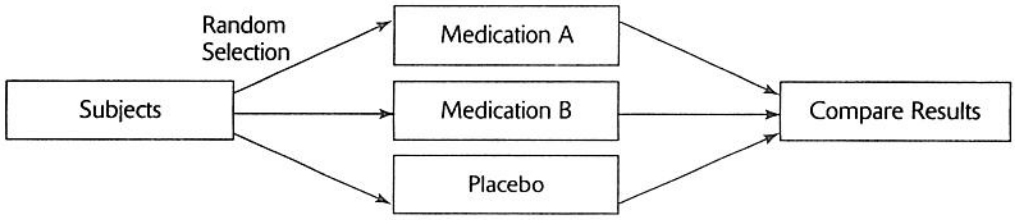


If an experimental design uses closely matched pairs of subjects, uses the same subject for each of two treatments, or uses each subject in a before-and-after experiment, it is called a **matched pairs** design. When the design applies two treatments to each subject, the *order* of the treatments must be randomized. In this type of matched pairs design, it is extremely important that a double-blind randomized assignment be used.

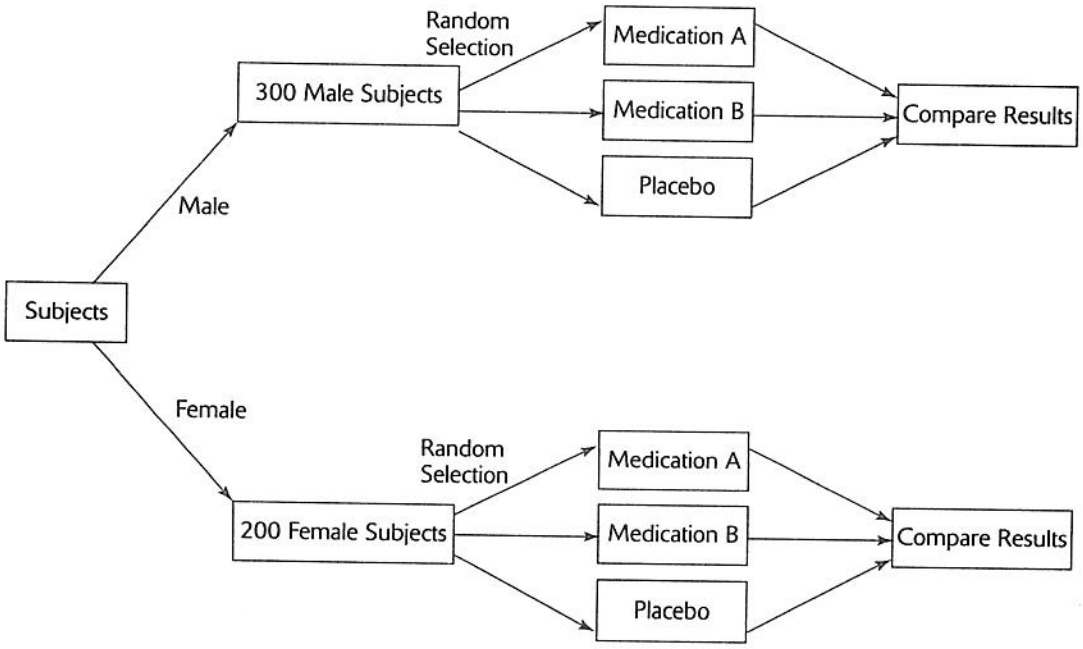
EXAMPLE:
A golf club manufacturer has designed a new style of golf club that it claims will help golfers hit the ball farther. Fifty golfers are randomly selected, and the hitting distance is determined for each with several types of clubs. Then the golfers are each given a set of the new clubs, and they are asked to practice with them for six months. After the practice period, their hitting distances are again determined. A comparison is made for each golfer between their hitting distance before receiving the new clubs and after practicing with them for six months. This design is a match pairs design since before and after distances are paired for each golfer.

EXAMPLE:
Two different pain medications are to be tested to determine their effectiveness. Each medication uses a different mechanism for blocking pain, and the two medications do not affect each other's effectiveness. Compare three different experimental designs for testing these medications.

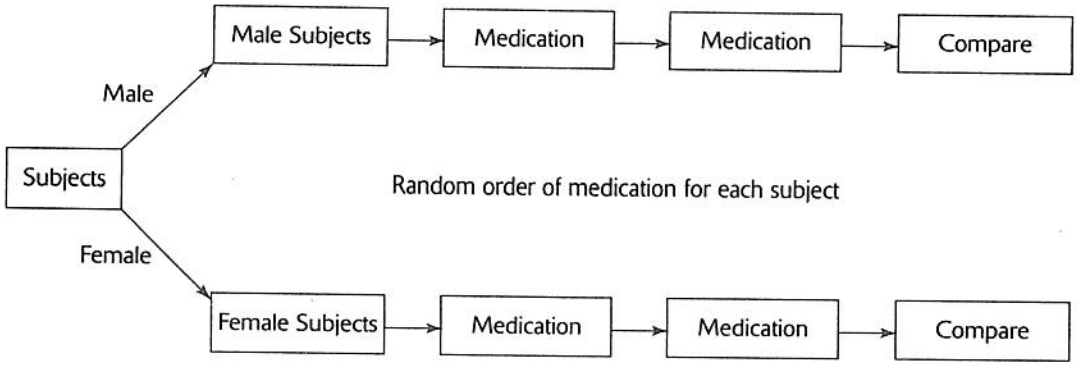
For the first method use a completely randomized design. In this design, no control exists for known differences in the subjects. Randomization is used to control unknown differences.



For the second method use a block design. In this design, blocks are used to control known differences that can be separated from the treatment. When dealing with human subjects, gender and age are classic differences that can be controlled with blocking. This example uses blocking to control for gender.



For the third method use a matched pairs design. Matched pairs is a type of block design. In this case, each subject will be given each medication separated by a period of time so as not to confuse effectiveness. The order that the two medications are administered for each subject is randomly selected. This design also uses blocking to control for gender, and each subject receives both treatments in a random order.



Generalizing Results from Observations, Experiments, and Surveys

The intent of collecting data from a sample is to be able to make generalizations about the population from which the sample is taken. This is true for observations, surveys, and experiments. Sample size plays a direct role in the accuracy of the generalization. Almost without exception, larger samples yield more accurate generalizations about populations. In experiments, replication helps control unknown lurking variables, and blocking helps control known variables in the sample. In surveys and observations, stratification helps make samples more representative of the population. Clearly specialized samples should be avoided no matter what method of data collection is used.