

Unit 15: Designing Experiments

SUMMARY OF VIDEO

Statistics helps us figure out the story hidden in a mound of data. Using statistics we can describe distributions, search for patterns, or tease out relationships. However, the reliability of our conclusions depends on the quality of the data collected.

One method of producing data is from an observational study. For the first story, we follow a team of marine scientists investigating how human populations affect coral reef ecosystems. They set up an observational study at four atolls in the remote Line Islands, each having a different history of human habitation. Kingman Reef has never had a human population; Palmyra was home to a military base during World War II but is no longer inhabited; Tabuaeran has a growing population of around 2,500; and Christmas Island has a population over 5,000. The research team recorded the size and quantity of predator fish, collected samples from each ecosystem, and took photographs of the coral. The scientists did not try to influence reef health – they simply observed it. They observed healthier ecosystems in areas with fewer humans. However, the problem with observational studies is they can't prove anything about cause and effect. So, while the scientists observed less healthy ecosystems in areas with human population, they could not state that humans caused the damage to the coral reefs.

In order to establish causal relationships, researchers rely on experimental studies. An experiment imposes some treatment on its subjects to see how they respond. The second story focuses on a study of how certain dietary supplements affected the pain of osteoarthritis. Here researchers set up a double-blind randomized comparative experiment. Participants were randomly assigned to one of five treatment groups: Glucosamine, Chondroitin, combination of Glucosamine and Chondroitin, Celecoxib, and placebo. The latter two groups were control groups. The Celecoxib group received a standard prescription medication and the placebo group received a dummy pill. The response variable was the reported decrease in knee pain. When researchers calculated the mean reduction in pain after six months for each treatment group, it turned out that they all had fairly similar outcomes. So, the dietary supplements were no worse or better than the prescription medication, or even the placebo.

The osteoarthritis study was a well-designed experiment – researchers randomly assigned their subjects to treatments, the treatments included control groups, and the number of subjects was large. Next, we visit Dr. Confound as he collects data for his study on mood-altering medication. The video clip with Dr. Confound focuses on two hypothetical patients –

his last two for his experiment, subjects 7 and 8. The patients get treated differently based on their initial mood – the one in a terrible mood is allowed to sit while the one in a good mood must stand. The doctor decides which medication to give each participant and he interacts with participants – adjusting one participant’s response and sympathizing with another’s. So this is a lesson in what not to do!

STUDENT LEARNING OBJECTIVES

- A. Distinguish experiments from observational studies.
- B. Recognize confounding in simple situations and the consequent weakness of uncontrolled studies.
- C. Outline the design of a randomized comparative experiment to compare two or more treatments.
- D. Understand how well-designed experiments can give good evidence for causation.
- E. Use a table of random numbers or calculator/computer to generate random numbers in order to carry out a random assignment of subjects to treatment groups.

CONTENT OVERVIEW

In an **observational study** we observe subjects and measure the variables of interest, explanatory variables and responses. For example, in studying the effects that humans have on the coral reefs, we go to several locations where the level of human habitation varies and then observe the conditions of the coral reefs in those areas. Even though we might observe that coral reefs are less healthy if near highly populated areas than near uninhabited areas, we can't say that humans are *causing* the damage observed. Observational studies allow us to observe association between an explanatory variable and a response variable, but we cannot use the observational study to establish a cause-and-effect relationship between the explanatory variable and response variable. In order to establish a cause-and-effect relationship, we need to conduct an **experiment**.

In an experiment, the researcher imposes some **treatment** on the subjects and then observes their responses. In the case of coral reefs, that would mean that we find some uninhabited areas with coral reefs, and then bring people in to live around one or more of those areas, and finally, after a period of time, we compare what happens to the coral reefs. If we observed damage associated with inhabited areas, then we could say that the **factor**, the explanatory variable human habitation, caused the damage to the coral reefs.

The basic ideas of experimental design are among the most important in statistics, and perhaps the most influential. Methods of conducting research in many of the applied sciences, from agricultural research to medicine, have been revolutionized by the use of statistically designed experiments. The distinction between observation and experiment is important. For example, measuring the height, weight, and blood pressure of a doctor's patients during their office visits and similar data-collection procedures are considered observation. While it may be true that observing, say, a patient's weight, may disturb the patient and so change their response (maybe they lose weight before the next visit and their blood pressure goes down), it is not an experiment. Experiments deliberately impose some treatment on the subjects in order to see how they respond.

In the science classroom, many experiments do not have the **randomized, comparative design** that was stressed in the video. Instead, in the controlled environment of a laboratory, many experiments have a simpler design:

Treatment → Observe outcome

For example,

Mix chemicals → Observe explosion

However, for agricultural or ecological experiments in the field or in experiments involving living subjects that vary a lot, this simple design invites **confounding** – mixing of a variety of causes, even those not considered in the experiment. Uncontrolled medical experiments, for example, have led doctors to conclude that worthless treatments worked. The natural optimism of the doctors who had invented the treatment, the **placebo effect** (the strong effect on patients of any treatment given even if the treatment is fake), and perhaps an unrepresentative group of patients combined to make the worthless treatments look good. Later, randomized comparative experiments found that the treatments had no value. Doctors gradually recognized that well-designed experiments were essential.

Agricultural researchers learned this earlier. Variation in the weather alone is enough to force experiments with new crop varieties to be comparative. The yield of a new variety this year may look great, but this was a good growing season. The virtues of the new variety are confounded with the effects of weather unless we do a comparative experiment in which older varieties were also grown. Unpredictable variation in soil makes random assignment of competing crop varieties to small growing plots desirable. In fact, statistical design of experiments first arose in the 1920s to solve the problems encountered by agricultural field trials.

Principles of experimental design

The basic principles of statistical design of experiments are:

1. **Randomization:** Use of impersonal chance to assign subjects to treatments to remove bias and other sources of extraneous variation. Randomization produces groups of subjects that should be similar in all respects before the treatments are applied. It allows us to equalize across all treatments the effects from unknown or uncontrollable sources of variation.
2. **Replication:** Repeat the experiment on many subjects to reduce chance variation in the results.
3. **Local Control:** Control all factors except the ones under investigation. Some examples of local control include: assigning equal numbers of subjects to each treatment, applying treatments uniformly and under standardized conditions, sorting subjects into homogeneous groups, and comparing two or more treatments.

In a **randomized comparative design** subjects are randomly assigned to treatment groups in order to study the effect the treatment has on the response. Figure 15.1 is an outline of the basic design for comparing two treatments.

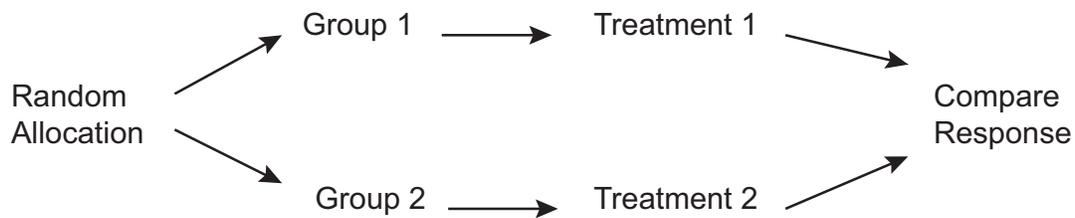


Figure 15.1. Outline of basic design for comparing two treatments.

Use this pictorial description rather than attempting a full description in words of an experimental design. Make the general design specific by adding the size of the groups, substituting the actual treatments compared (in place of Treatment #) and the response recorded. Notice that “treatment” is used for whatever is imposed on the subjects in an experiment, not only medical treatments.

The logic behind a randomized comparative design is straightforward. Randomization produces groups of subjects that should be similar in all respects before the treatments are applied. A comparative design ensures that influences other than the experimental treatments operate equally on all groups. Therefore, differences in the response variable must be due to the effects of the treatments – that is, the treatments are not only associated with the observed differences in the response, but must actually cause them.

To do the random assignment of subjects to groups, use a table of random digits or a random number generator (which is built into graphing calculators, as well as into spreadsheet and statistical software).

Randomizing treatments using a table of random digits

You can find a table of random digits in most introductory statistics textbooks as well as on the Internet. A random digits table has the following properties:

- Each entry in the table is equally likely to be any of the 10 digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.
- The entries are independent of each other. That is, knowledge of one part of the table gives no information about any other part.

A random digits table is one long string of random digits. The numbers are generally arranged in groups of five and the rows numbered for convenience. Two digits from the table are equally likely to be any of the 100 possibilities 00 to 99, and so on. To choose at random, assign the subjects numerical labels and let the table choose from these labels at random.

For example, to choose at random 10 out of 20 subjects to get Treatment 1 (aspirin), label the subjects 01 to 20 and enter the random digits table on any line (vary the entry point each time you use the table), say line 110, which is given below.

38448 48789 18338 24697 39364 42006 76688 08708

Read two-digit groups because we assigned two-digit labels. Just skip over all groups not used as labels. So, skip 38, 44, 84, 87, and 89. The first subject chosen has label number 18. Continue this process until 10 subjects are chosen – next select 20, 06, 08, and so forth. Assign the 10 selected subjects to the aspirin treatment. The remaining subjects get the placebo.

Randomizing treatments using a random number generator

Graphing calculators and spreadsheet and statistical computing software have random number generators. Rand (TI graphing calculators and Excel) generates a number from the uniform distribution on the interval from 0 to 1. Suppose you have a list of 10 subjects and you want to choose five of them to assign to Treatment 1 (aspirin).

Step 1: Use a random number generator to assign a random number to each person. Excel's Rand() was used to generate the set of random numbers in Table 15.1.

Subjects	Rand
Joe	0.305127
Sally	0.130861
Kelly	0.335956
Bruce	0.525466
Marsha	0.288252
Caitlin	0.036762
George	0.084562
Jian	0.763097
Cheryl	0.753380
Ying	0.775420

Table 15.1. Assigning a random number to subjects.

Step 2. Sort the names so that the first name is the name associated with the smallest random number and the last name is associated with the largest random number. Table 15.2 shows the results.

Subjects	Rand
Caitlin	0.036762
George	0.084562
Sally	0.130861
Marsha	0.288252
Joe	0.305127
Kelly	0.335956
Bruce	0.525466
Cheryl	0.753380
Jian	0.763097
Ying	0.775420

Table 15.2. Names sorted by random number.

Step 3. Select the first five names from the sorted list to assign to the aspirin treatment.

So, in this case, Caitlin, George, Sally, Marsha, and Joe would be assigned to the aspirin treatment and the remaining subjects would be assigned to the placebo group.

Who knows who is getting which treatment?

Particularly in the context of medical studies, it is important to know whether or not the participants know which treatment they are getting and whether or not those recording the responses know. In **double-blind experiments** neither the participants nor those conducting the experiment know which participants were assigned to which treatments. In a **single-blind** experiment the participants do not know which treatment they are receiving but those conducting the experiment do know.

KEY TERMS

In an **observational study** researchers observe subjects and measure variables of interest. However, the researchers do not try to influence the responses. The purpose is to *describe* groups of subjects under different situations. In an **experimental study**, researchers deliberately apply some treatment to the subjects in order to observe their responses. The purpose is to study whether the treatment *causes* a change in the response.

In a **double-blind** experiment neither the subjects nor the individuals measuring the response know which subjects are assigned to which treatment. In a **single-blind** experiment the subjects do not know which treatment they are receiving but the individuals measuring the response do know which subjects were assigned to which treatments.

A **placebo** is something that is identical in appearance to the treatment received by the treatment group but has no effect.

A **control group** is an experimental group that does not receive the treatment under study. The control group could receive a placebo to hide the fact that no treatment is being given. In an **active control group**, the subjects receive what might be considered the existing standard treatment.

The explanatory variables in either an observational study or experiment are called **factors**. A **treatment** is any specific condition applied to the subjects in an experiment. If an experiment has more than one factor, then a treatment is a combination of specific values for each factor.

Two factors (explanatory variables) are **confounded** when their effects on a response variable are intertwined and cannot be distinguished from each other.

UNIT ACTIVITY:

IN THE NEWS

Find an article in the print news (newspaper, online news, magazine) that reports on the results of a medical experiment. Describe what information appears in the article about the design of the experiment. Be prepared to discuss your news item in class.

EXERCISES

1. It has been suggested that women and men differ in their political preferences. Women may be more likely than men to prefer Democratic candidates. A political scientist selects a large sample of registered voters, both men and women, and asks them whether they voted for the Democratic or Republican candidate in the last Congressional election. Is this study an experiment? Why or why not?
2. Before a new variety of frozen muffins is put on the market, it is subjected to extensive taste testing. People are asked to taste the new muffin and a competing brand, and to say which they prefer. (The muffins are not identified in the test.) Is this an experiment? Why or why not?
3. You are testing a new medication for relief of migraine headache pain. You intend to give the drug to migraine sufferers and ask them one hour later to estimate what percent of their pain has been relieved. You have 40 patients available to serve as subjects.
 - a. Outline an appropriate design for the experiment, taking the placebo effect into account.
 - b. The names of the subjects are given in Table 15.3. Either use a random digits table beginning at line 131 to do the randomization required by your design or use a calculator's or computer software's random number generator. List the subjects to whom you will give the drug. Explain how you arrived at this assignment.

Abrams	Daniels	Halsey	Lippman	Rosen
Adamson	Durr	Howard	Martinez	Solomon
Afifi	Edwards	Hwang	McNeill	Thompson
Brown	Fluharty	Iselin	Morse	Travers
Cansico	Garcia	Janle	Ng	Turner
Chen	Gerson	Kaplan	Obramowitz	Ullman
Cranston	Green	Krushchev	Rivera	Williams
Curzakis	Gutierrez	Lattimore	Roberts	Wong

Table 15.3. Names of subjects.

4. Determine which of the experiments below are single-blind, double-blind, or neither. Justify your answer.

a. Dr. Colman has a home remedy that he thinks will help his patients recover from colds. He arranges for a colleague to have two sets of identical looking pills made up and bottled – pill A contains his remedy and pill B is a placebo. He recruits some of his patients to take part in the experiment. They are told that they will be randomly assigned to his remedy or a placebo. Dr. Colman does not know which bottles of pills, A or B, contain his remedy. After six months, Dr. Colman interviews his patients to check on the number of colds they contracted and the duration. He then summarizes his results in a report before learning from his colleague which group of participants got his remedy.

b. Pam decides that she wants to know what type of diet cola tastes best – Diet Coke, Caffeine Free Diet Coke, or Coca Cola Zero. So she invites a group of friends over for a taste test. She sets an unopened bottle of each type of soda on a table with some paper cups. Each of the participants pours some cola from each bottle into paper cups, marking the cups with the type of soda. Then they taste each type of diet soda and give Pam their rating – 1 (tastes awful) to 5 (tastes great).

c. Janet wants to know whether her secret recipe for chocolate cake will taste better if she uses cocoa or baking chocolate. She bakes two cakes that appear identical and asks a group of her friends to take part in a taste test. She labels the cake with cocoa as A and the cake with baking chocolate as B. Then she randomly assigns half her friends to taste cake A and the other half to taste cake B. Her friends are then asked to rate each cake from 1 to 10.

REVIEW QUESTIONS

1. You wish to learn if students in an English course write better essays when they are required to use computer word-processing than when they write and revise their essays by hand. There are 120 students in an English course available as subjects.

- a. Outline the design of an experiment to determine if word-processing results in better essays.
- b. What precautions would you take in doing this experiment that don't appear in your outline of the design?

2. Two second grade teachers, Miss Earls (who has been teaching for 10 years) and Mrs. Morrow (who has been teaching for two years), were really excited by a new curriculum that used animations to teach science. They decided to use their classrooms for an experiment. Since Miss Earls had access to computers in her class, she used the animation lessons. Mrs. Morrow covered similar material with her students using handouts followed by discussions. After students had completed the materials, they were given a test designed by Miss Earls. There were 21 students in Miss Earls' class and 29 students in Mrs. Morrow's class. Miss Earls' class scored, on average, 15 points higher on the test. The two teachers decided that the animation science series was an excellent series. Based on this study, Miss Earls talked her school into purchasing this curriculum.

Critique the study conducted by Miss Earls and Mrs. Morrow. Do you think Miss Earls' school made a good decision in purchasing the animation science curriculum based on this study?

3. A study reported in the *Annals of Internal Medicine* (January 2010) followed 10,892 middle-aged adults over a nine year period. At the start of the study none of the subjects had diabetes. Roughly 45% of the subjects were smokers. The study found that compared to those who never smoked, subjects who quit smoking had an increased risk of diabetes.

- a. Is the study described above an observational study or an experiment? Explain.
- b. Based on this study, should you conclude that quitting smoking causes diabetes? Justify your answer.

4. The research question for an undergraduate research project was whether hearing-impaired consumers were treated differently by store clerks than non-hearing-impaired consumers.

There were 20 consumers, 10 of whom were hearing impaired. The consumers were sent in pairs into stores. The hearing-impaired pairs used sign language to communicate with each other and the non-hearing impaired pairs entered stores speaking English to each other. The subjects consisted of 77 sales clerks in 27 stores (from the 175 stores) in a large shopping mall. The response variable was the time from when the pair entered the store and made eye contact with the sales clerk until the clerk approached and offered assistance.

Describe how you would design the rest of the experiment.