

1.3 Intro to Experimental Design with work

1. Stratified: random samples from each stratum are included. Cluster: clusters are randomly selected and then all members of each cluster are included.

3. The advice is wrong. A sampling error talks about the difference in results based on the sample not the entire population. She just missed the serving size.

5. No, while it is random :, some students can be younger than 18 or older than 20 that were not included.

7. a) stratified b) No, each pooled sample would have 100 ticket holders for both men and womens basketball. 125 for men and 75 for women's games type samples are not possible.

9. (ANSWERS VARY) Samples can be biased. a) convenience sampling b) they could be coming from the same class c) they might have chosen to sit there d) most likely all boys

11. (ANSWERS VARY) Use two digit labels.

13. (ANSWERS VARY) Use four digit numbers. **1.2 HW**

15. a) Yes; outcome of die roll can repeat; 2. b) No; process is random.

19. a) SRS b) Cluster c) Convenience d) Systematic e) Stratified

Sep 11-9:13 AM

1.3 Introduction to Experimental Design

Essential Questions:

- What factors can ruin an experiment?
- What is the difference between an observational study and an experiment?

Focus Points:

- Discuss what it means to take a census.
- Describe simulations, observational studies, and experiments.
- Identify control groups, placebo effects, completely randomized experiments, and randomized block experiments.

Sep 4-10:18 AM

Planning a Statistical Study

1. Identify the individuals or objects of interest.
2. Specify the variables as well as the protocols for taking measurements or making observations.
3. Determine if you will use an entire population or a sample.
4. Pick the sampling method.
5. In data collection plan, address issues of ethics, subjects, confidentiality, and privacy. If you are collecting data at a business, store, college, or other institution, be sure to be courteous and to obtain permission as necessary.
6. Collect the data.
7. Use appropriate descriptive statistics methods and make decisions using appropriate inferential statistics methods.
8. Note any concerns you might have about your data collection methods and list any recommendations for future studies.

Sep 4-10:22 AM



Sep 5-3:14 PM

Reminder...

A **census** measures or observes the ENTIRE population.

A **sample** measures or observes part of the population used.

Sep 4-10:38 AM

In an **observational study**, observations and measurements of individuals are conducted in a way that *doesn't change* the response or the variable being measured.

In an **experiment**, a **treatment** is deliberately imposed on the individuals in order to observe a possible change in the response or variable being measured.

Sep 4-10:39 AM

1.3 Intro to Experimental Design with work

Experiment VS Study

In 1778, Captain James Cook landed in what we now call the Hawaiian Islands. He gave the islanders a present, goats. These goats multiplied over the years turning into several thousand. They eat anything, including a famous silver sword plant. Each year this plant had fewer and fewer plants. Biologists suspected that the goats were partially responsible for the decline in the number of plants and conducted a statistical study that verified their theory. In remote areas, biologists set up two plots of land similar in soil conditions, climate, and plant count. One plot was fenced to keep goats out, while the other was not. Regularly these plots were counted for number of plants.

a) Is this an experiment or observational study?

Experiment because the fence was the "treatment" and the non-fenced plot is the control group.



Sep 4-10:41 AM

The **placebo effect** occurs when a subject receives no treatment but **incorrectly** believes he or she is in fact receiving treatment and **responds favorably**.

A **completely randomized experiment** is one in which a random process is used to assign each individual to one of the treatments.

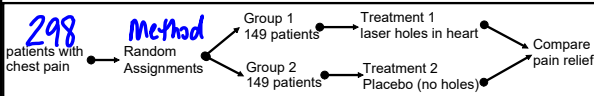
Sep 4-10:53 AM

Can chest pain be relieved by drilling holes in the heart?

Surgeons have been using a laser procedure to drill holes in the heart. Many patients report a lasting and dramatic decrease in chest pain. Is the relief due to the procedure, or is it a placebo effect?

A research project provided some information about the issue by using a completely randomized experiment. The laser treatment was applied through a less invasive process. A group of 298 volunteers with severe, untreatable chest pain were randomly assigned to get the laser or not. The patients were sedated but awake. They could hear the whole procedure. Each patient thought he/she was receiving the treatment.

The experimental design looked like this:



Sep 4-10:55 AM

Since it is hard to control all variables there are other methods:

A **block** is a group of individuals sharing some common features that might affect the treatment. ie: male vs female

In a **randomized block experiment**, individuals are first sorted into blocks, and then a random process is used to assign each individual in the block to one of the treatments. ie: split males and females and split again with two treatments groups.

There is a **control group**. This group receives a dummy treatment, enabling the researchers to control for the placebo effect. In general, a control group is used to account for the influence of other known or unknown variables that might be an underlying cause of a change in response in the experimental group. Such variables are called **lurking** or **confounding variables**.

Randomization is used to assign individuals to the two treatment groups. This helps prevent bias in selecting members for each group.

Replication of the experiment on many patients reduces the possibility that the differences in pain relief for the two groups occurred by chance alone.

Another method is called **double-blind**. In this type of experiment no one knows who is receiving the treatment. Individuals and observers are unaware.

Sep 5-12:43 PM

Example 1: Collecting Data

Which technique for gathering data (sampling, experiment, simulation, or census) do you think might be the most appropriate for the following studies?

- a) study of the effect of stopping the cooling process of a nuclear reactor. **simulation**
- b) study of the amount of time college students taking a full course load spend watching television. **sampling**
- c) study of the effect on bone mass of a calcium supplement given to young girls. **experiment**
- d) study of the credit hour load of *each* student enrolled at your college at the end of the drop/add period this semester. **census**

sampling experiment simulation census

Sep 5-12:52 PM

Example 1: Collecting Data

Which technique for gathering data (sampling, experiment, simulation, or census) do you think might be the most appropriate for the following studies?

- a) study of the effect of stopping the cooling process of a nuclear reactor. **Simulation, since you probably do not want to risk a nuclear meltdown.**
- b) study of the amount of time college students taking a full course load spend watching television. **Sampling and using an observational study would work well. Notice that obtaining the information from a student will probably not change the amount of time the student spends watching television.**
- c) study of the effect on bone mass of a calcium supplement given to young girls. **Experiment. Half were randomly selected and given a placebo. The other half were given calcium supplements.**
- d) study of the credit hour load of *each* student enrolled at your college at the end of the drop/add period this semester. **Census. The registrar can obtain records for every student.**

Sep 5-12:52 PM

1.3 Intro to Experimental Design with work

SURVEYS

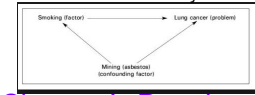
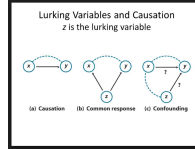
After you decide the type of data collection (sampling, experiment, census, or simulation) then the next step is usually **surveying** by asking individuals questions.

SOME POTENTIAL PITFALLS OF A SURVEY

- **nonresponse**: Individuals either cannot be contacted or refuse to participate. Nonresponse can result in significant **undercoverage** of a population.
- **truthfulness of response**: respondents may lie intentionally or inadvertently.
- **faulty recall**: respondents may not accurately remember when or whether an event took place.
- **hidden bias**: the question may be worded in such a way to elicit a specific response (mustache) The order of questions might lead to biased responses.
- **vague wording**: words such as "often," "seldom," and "occasionally" mean different things to different people.
- **interviewer influence**: factors such as tone of voice, body language, dress, gender, authority, and ethnicity of the interviewer might influence responses.
- **voluntary response**: individuals with strong feelings about a subject are more likely than other to respond. ie: bad driver numbers on semi's

Sep 5-1:00 PM

A **lurking variable** is one for which no data have been collected but that **nevertheless has influence** on other variables in the study.



Simpson's Paradox

Two variables are **confounded** when the effects of one cannot be distinguished from the effects of the other. **Confounding variables** may be part of the study, or they may be outside **lurking variables**.

Explanatory variable (x) explains the outcome on another variable causing **causation**. The **response variable (y)** is the variable changing because of another variable.

Sep 5-1:23 PM

Example 2: Cautions About Data

Talk to your elbow partner or table, and comment on the usefulness of the data collected as described. Mention at least one thing per scenario. Be ready to share as a class.

- A uniformed police officer interviews a group of 20 college freshmen. She asks each one his or her name and then if he or she has used an illegal drug in the last month.
- Jessica saw some data that show that cities with more low-income housing have more homeless people. Does building low-income housing cause homelessness?
- A survey about food in the student cafeteria was conducted by having forms available for customers to pick up at the cash register. A drop box for completed forms was available outside the cafeteria.
- Extensive studies on coronary problems were conducted using men over age 50 as the subjects.

Sep 5-1:29 PM

Example 2: Cautions About Data

These are your books comments....

- A uniformed police officer interviews a group of 20 college freshmen. She asks each one his or her name and then if he or she has used an illegal drug in the last month. **Respondents may not answer truthfully. Some may refuse to participate.**
- Jessica saw some data that show that cities with more low-income housing have more homeless people. Does building low-income housing cause homelessness? **There may be some other lurking variable, such as size of the city. Larger cities may have more low-income housing and more homeless.**
- A survey about food in the student cafeteria was conducted by having forms available for customers to pick up at the cash register. A drop box for completed forms was available outside the cafeteria. **The voluntary response likely produced more negative comments.**
- Extensive studies on coronary problems were conducted using men over age 50 as the subjects. **Conclusions for men over age 50 may or may not generalize to other age and gender groups. These results may be useful for women or younger people, but studies specifically involving these groups may need to be performed.**

Sep 5-1:29 PM

Choosing Data Collection Techniques

Surveys: best for a wide range of variables. best for the use of many questions. Careful of bias issues that can arise.

Observational Studies: fairly convenient with many variables. protocols for measurements or recording observations are very specific.

Experiments: most stringent and restrictive data-gathering. most time-consuming, expensive, and difficult to execute. number of variables are more limited. needs careful design.

Sep 5-1:39 PM

Key Features of Data Collection Plans

- The population and the sampling frame
- the variable(s)
- observational or experimental
- control group, use of placebos, double-blind treatment...?
- sampling technique to be used, block design?
- the method used to collect the data for the variables: survey, method of measurement and accuracy level, count, etc.

Sep 5-1:43 PM

1.3 Intro to Experimental Design with work

HW: pg. 29: 1 - 11 odd only

Sep 5-1:45 PM