Section 1.3

Experimental Design

Section 1.3 Objectives

- Discuss how to design a statistical study
- Discuss data collection techniques
- Discuss how to design an experiment
- Discuss sampling techniques

Designing a Statistical Study

1. Identify the variable(s) of interest (the focus) and the population of the study.

2. Develop a detailed plan for collecting data. If you use a sample, make sure the sample is representative of the population.

3. Collect the data.

4. Describe the data using descriptive statistics techniques.

5. Interpret the data and make decisions about the population using inferential statistics.

6. Identify any possible errors.

3

Data Collection

Observational study

- A researcher observes and measures characteristics of interest of part of a population.

- Researchers observed and recorded the mouthing behavior on nonfood objects of children up to three years old. *(Source: Pediatric Magazine)*

Data Collection

Experiment

- A treatment is applied to part of a population and responses are observed.

- An experiment was performed in which diabetics took cinnamon extract daily while a control group took none. After 40 days, the diabetics who had the cinnamon reduced their risk of heart disease while the control group experienced no change. *(Source: Diabetes Care)*

Data Collection

Simulation

- Uses a mathematical or physical model to reproduce the conditions of a situation or process.
- Often involves the use of computers.

- Automobile manufacturers use simulations with dummies to study the effects of crashes on humans.

Data Collection

Survey

- An investigation of one or more characteristics of a population.
- Commonly done by interview, mail, or telephone.

- A survey is conducted on a sample of female physicians to determine whether the primary reason for their career choice is financial stability.

Example: Methods of Data Collection

Consider the following statistical studies. Which method of data collection would you use to collect data for each study?

1. A study of the effect of changing flight patterns on the number of airplane accidents.

   Solution:
   Simulation (It is impractical to create this situation)

Example: Methods of Data Collection

2. A study of the effect of eating oatmeal on lowering blood pressure.

   Solution:
   Experiment (Measure the effect of a treatment – eating oatmeal)
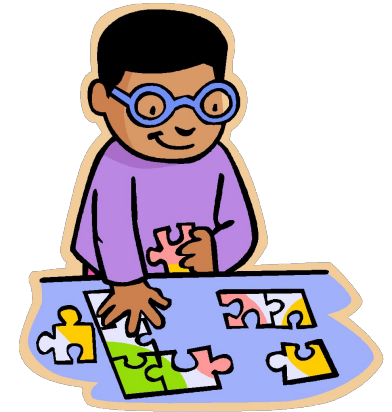
Example: Methods of Data Collection

3. A study of how fourth grade students solve a puzzle.

    Solution:
    Observational study (observe and measure certain characteristics of part of a population)

Example: Methods of Data Collection

4. A study of U.S. residents' approval rating of the U.S. president.

Solution:

Survey (Ask "Do you approve of the way the president is handling his job?")

Key Elements of Experimental Design

- Control
- Randomization
- Replication

Key Elements of Experimental Design: Control

- **Control** for effects other than the one being measured.
- **Confounding variables**
  - Occurs when an experimenter cannot tell the difference between the effects of different factors on a variable.
  - A coffee shop owner remodels her shop at the same time a nearby mall has its grand opening. If business at the coffee shop increases, it cannot be determined whether it is because of the remodeling or the new mall.

Key Elements of Experimental Design: Control

- **Placebo effect**
  - A subject reacts favorably to a placebo when in fact he or she has been given no medical treatment at all.
  - **Blinding** is a technique where the subject does not know whether he or she is receiving a treatment or a placebo.
  - **Double-blind** experiment neither the subject nor the experimenter knows if the subject is receiving a treatment or a placebo.
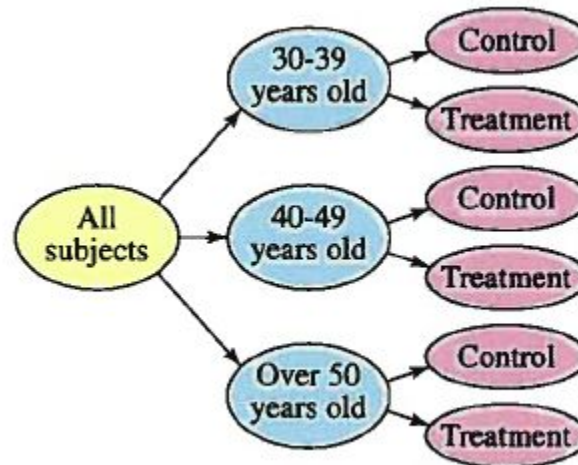
Key Elements of Experimental Design: Randomization

- **Randomization** is a process of randomly assigning subjects to different treatment groups.

- **Completely randomized design**
  - Subjects are assigned to different treatment groups through random selection.

- **Randomized block design**
  - Divide subjects with similar characteristics into **blocks**, and then within each block, randomly assign subjects to treatment groups.

Key Elements of Experimental Design: Randomization

Randomized block design

- An experimenter testing the effects of a new weight loss drink may first divide the subjects into age categories. Then within each age group, randomly assign subjects to either the treatment group or control group.

Key Elements of Experimental Design: Randomization

- **Matched Pairs Design**
  - Subjects are paired up according to a similarity. One subject in the pair is randomly selected to receive one treatment while the other subject receives a different treatment.

Key Elements of Experimental Design: Replication

- **Replication** is the repetition of an experiment using a large group of subjects.

- To test a vaccine against a strain of influenza, 10,000 people are given the vaccine and another 10,000 people are given a placebo. Because of the sample size, the effectiveness of the vaccine would most likely be observed.

Example: Experimental Design

A company wants to test the effectiveness of a new gum developed to help people quit smoking. Identify a potential problem with the given experimental design and suggest a way to improve it.

The company identifies one thousand adults who are heavy smokers. The subjects are divided into blocks according to gender. After two months, the female group has a significant number of subjects who have quit smoking.

Solution: Experimental Design

Problem:

The groups are not similar. The new gum may have a greater effect on women than men, or vice versa.
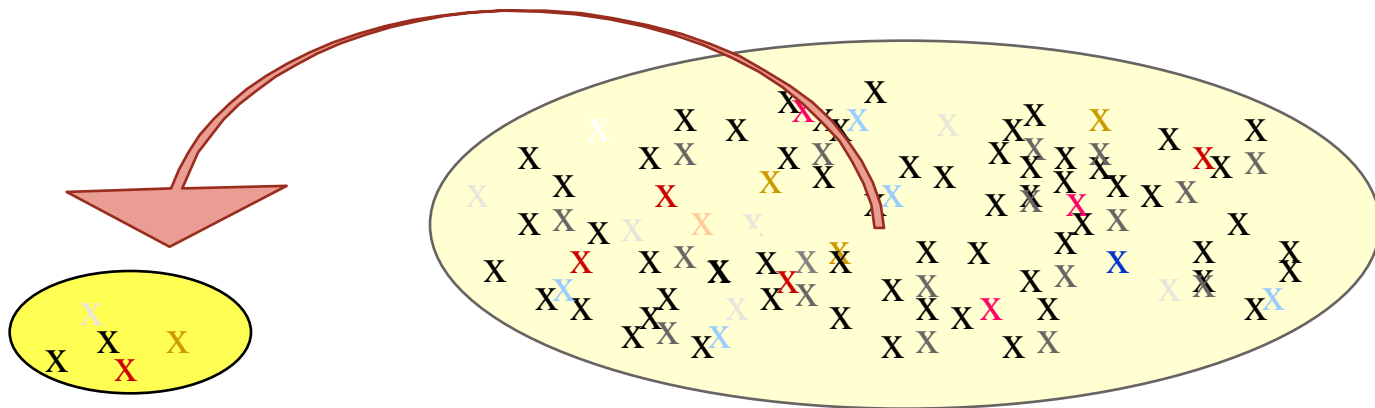
Correction:

The subjects can be divided into blocks according to gender, but then within each block, they must be randomly assigned to be in the treatment group or the control group

Sampling Techniques

Simple Random Sample

Every possible sample of the same size has the same chance of being selected.

Simple Random Sample

- Random numbers can be generated by a random number table, a software program or a calculator.

- Assign a number to each member of the population.

- Members of the population that correspond to these numbers become members of the sample.

Example: Simple Random Sample

There are 731 students currently enrolled in statistics at your school. You wish to form a sample of eight students to answer some survey questions. Select the students who will belong to the simple random sample.

- Assign numbers 1 to 731 to each student taking statistics.
- On the table of random numbers, choose a starting place at random (suppose you start in the third row, second column.)

Solution: Simple Random Sample

Table 1— Random Numbers

92630	78240	19267	95457	53497	23894	37708	79862	76471	66418
79445	78735	71549	44843	26104	67318	00701	34986	66751	99723
	71966	27386	50004	05358	94031	29281	18544	52429	06080
31524	49587	76612	39789	13537	48086	59483	60680	84675	53014
06348	76938	90379	51392	55887	71015	09209	79157	24440	30244
28703	51709	94456	48396	73780	06436	86641	69239	57662	80181
68108	89266	94730	95761	75023	48464	65544	96583	18911	16391
99938	00704	93621	663	393		8534	517	91616	32

- Read the digits in groups of three
- Ignore numbers greater than 731

719|66   2|738|6   50|004| 053|58   9|403|1   29|281| 185|44

The students assigned numbers 719, 662, 650, 4, 53, 589, 403, and 129 would make up the sample.

24

Other Sampling Techniques

Stratified Sample

☐ Divide a population into groups (strata) and select a random sample from each group.

- To collect a stratified sample of the number of people who live in West Ridge County households, you could divide the households into socioeconomic levels and then randomly select households from each level.


Group 1: Low income

Group 2: Middle income

Group 3: High income

Other Sampling Techniques

Cluster Sample

☐ Divide the population into groups (clusters) and select all of the members in one or more, but not all, of the clusters.

- In the West Ridge County example you could divide the households into clusters according to zip codes, then select all the households in one or more, but not all, zip codes.

Zip Code Zones in West Ridge County

Zone 1

Zone 2

Zone 3

Zone 4
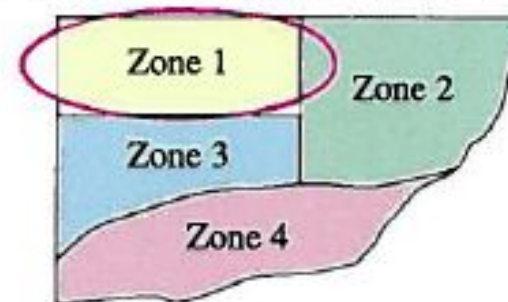
Other Sampling Techniques

Systematic Sample

- Choose a starting value at random. Then choose every k^{th} member of the population.

- In the West Ridge County example you could assign a different number to each household, randomly choose a starting number, then select every 100th household.

Example: Identifying Sampling Techniques

You are doing a study to determine the opinion of students at your school regarding stem cell research. Identify the sampling technique used.

1. You divide the student population with respect to majors and randomly select and question some students in each major.

    Solution:
    Stratified sampling (the students are divided into strata (majors) and a sample is selected from each major)

Example: Identifying Sampling Techniques

2. You assign each student a number and generate random numbers. You then question each student whose number is randomly selected.

Solution:
Simple random sample (each sample of the same size has an equal chance of being selected and each student has an equal chance of being selected.)

Section 1.3 Summary

- Discussed how to design a statistical study
- Discussed data collection techniques
- Discussed how to design an experiment
- Discussed sampling techniques