Understanding Experiments; What Are We Summarizing?

Lecture 11
Sections 3.5, 4.1 - 4.2

Robb T. Koether
Hampden-Sydney College

Mon, Feb 1, 2010
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Suppose a drug is given to 100 patients suffering from a particular disease.
Suppose a drug is given to 100 patients suffering from a particular disease.

After 2 weeks, 90% of the patients have recovered.
Suppose a drug is given to 100 patients suffering from a particular disease.

After 2 weeks, 90\% of the patients have recovered.

The researchers conclude that the drug was effective.
Suppose a drug is given to 100 patients suffering from a particular disease.

After 2 weeks, 90% of the patients have recovered.

The researchers conclude that the drug was effective.

What is wrong with this?
Outline

1. Experiment Design
   - Control Groups
     - Randomized Design
     - Blinded Experiments
2. Homework Review
3. Types of Variable
4. Qualitative Variables
5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables
6. Assignment
Treatment and Control Groups

Definition (Treatment group)
The treatment group is the group that receives the treatment.

Definition (Control group)
The control group is similar to the treatment group in all respects except that it does not receive the treatment.
Outline

1. **Experiment Design**
   - Control Groups
   - **Randomized Design**
   - Blinded Experiments

2. **Homework Review**

3. **Types of Variable**

4. **Qualitative Variables**

5. **Quantitative Variables**
   - Caution
   - Continuous and Discrete Variables

6. **Assignment**
Randomized Design

- Why would it be wrong to allow the individuals themselves to choose whether to be in the treatment group or the control group?
- Why would it be wrong for the researchers to decide, subject by subject, who goes into which group?
A **randomized design** is a design in which the subjects are randomly assigned to either the treatment group or the control group.
Randomized Design

- Suppose that there are 100 subjects.
- Number them 1 - 100.
- Then use a random number generator to obtain 50 (distinct) random numbers from 1 - 100.
- Those 50 subjects would be assigned to the treatment group.
- The rest would be assigned to the control group.
Possible Bias

- Are the subjects in the treatment group aware of the purpose of the experiment?
- Are the subjects in the control group aware that they are not receiving the drug?
- Will it make a difference?
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables
   - Caution
   - Continuous and Discrete Variables

5. Quantitative Variables

6. Assignment
Placebos

Definition (Response bias)
A sampling method exhibits **response bias** if the subjects give what they perceive to be the desired response rather than the true response.

Definition (Placebo)
A **placebo** is a treatment, usually a pill, that is known to have no effect.

Definition (Single-blind experiment)
A **single-blind experiment** is an experiment in which the subjects do not know who is in the treatment group and who is in the control group.
Placebos

- Everybody in the treatment group is administered the drug.
- Everybody in the control group gets the placebo.
- No subject knows which group he is in.
- The researchers look for differences in the groups’ recovery rates.
Definition (Experimenter bias)
A sampling method exhibits experimenter bias if the observer records the desired values rather than the true observed values.

Definition (Double-blind experiment)
A double-blind experiment is an experiment in which neither the subjects nor the observers know who is in the treatment group and who is in the control group.
Outline

1 Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2 Homework Review

3 Types of Variable

4 Qualitative Variables

5 Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6 Assignment
Exercise 3.31, p. 184

A nutrition experimenter intends to compare the weight gain of male rats fed Diet A with that of male rats fed Diet B. Twenty rats are available for this experiment. She would like to randomly allocate 10 rats to Diet A and the remaining 10 rats to Diet B. She decides to label the 20 rats as follows: 01, 02, 03, . . . , 20.

(a) Use the experimenter’s labels and your calculator (with a seed value of 72) to identify the labels of the first 10 rats selected that would be assigned to Diet A.

The first 10 distinct random numbers are 14, 13, 5, 16, 9, 3, 4, 2, 7, and 10.
Exercise 3.31, p. 184

A nutrition experimenter intends to compare the weight gain of male rats fed Diet A with that of male rats fed Diet B. Twenty rats are available for this experiment. She would like to randomly allocate 10 rats to Diet A and the remaining 10 rats to Diet B. She decides to label the 20 rats as follows: 01, 02, 03, . . . , 20.

(a) Use the experimenter’s labels and your calculator (with a seed value of 72) to identify the labels of the first 10 rats selected that would be assigned to Diet A.
Exercise 3.31, p. 184

A nutrition experimenter intends to compare the weight gain of male rats fed Diet A with that of male rats fed Diet B. Twenty rats are available for this experiment. She would like to randomly allocate 10 rats to Diet A and the remaining 10 rats to Diet B. She decides to label the 20 rats as follows: 01, 02, 03, . . . , 20.

(a) Use the experimenter’s labels and your calculator (with a seed value of 72) to identify the labels of the first 10 rats selected that would be assigned to Diet A.

The first 10 distinct random numbers are 14, 13, 5, 16, 9, 3, 4, 2, 7, and 10.
Exercise 3.31, p. 184

(b) The experimenter hopes Diet A is the better diet. To avoid possible bias in her results because of her favoritism, she blinds herself to her experiment. Explain how she would do this.

She would make sure that the person who made the measurements on the rats did not know which rats were given Diet A and which were given Diet B.
Exercise 3.31, p. 184

(b) The experimenter hopes Diet A is the better diet. To avoid possible bias in her results because of her favoritism, she blinds herself to her experiment. Explain how she would do this.

She would make sure that the person who made the measurements on the rats did not know which rats were given Diet A and which were given Diet B.
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Types of Variable

- Statisticians like to quantify everything.

- A sample of political affiliations?
- A sample of body weights?
- A sample of steak preferences (rare, medium, etc.)?
- A sample of family sizes?
- A sample of temperatures throughout the day?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.

How might we summarize:
- A sample of political affiliations?
- A sample of body weights?
- A sample of steak preferences (rare, medium, etc.)?
- A sample of family sizes?
- A sample of temperatures throughout the day?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.
- How might we summarize
  - A sample of political affiliations?
  - A sample of body weights?
  - A sample of steak preferences (rare, medium, etc.)?
  - A sample of family sizes?
  - A sample of temperatures throughout the day?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.
- How might we summarize
  - A sample of political affiliations?
  - A sample of body weights?
  - A sample of steak preferences (rare, medium, etc.)?
  - A sample of family sizes?
  - A sample of temperatures throughout the day?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.
- How might we summarize
  - A sample of political affiliations?
  - A sample of body weights?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.
- How might we summarize
  - A sample of political affiliations?
  - A sample of body weights?
  - A sample of steak preferences (rare, medium, etc.)?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.
- How might we summarize
  - A sample of political affiliations?
  - A sample of body weights?
  - A sample of steak preferences (rare, medium, etc.)?
  - A sample of family sizes?
Types of Variable

- Statisticians like to quantify everything.
- Given a set of data, they want to summarize it with a single number.
- How might we summarize
  - A sample of political affiliations?
  - A sample of body weights?
  - A sample of steak preferences (rare, medium, etc.)?
  - A sample of family sizes?
  - A sample of temperatures throughout the day?
Outline

1. **Experiment Design**
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. **Homework Review**

3. **Types of Variable**

4. **Qualitative Variables**

5. **Quantitative Variables**
   - Caution
   - Continuous and Discrete Variables

6. **Assignment**
Qualitative Variables

Definition (Qualitative variable)
A qualitative variable is a variable whose values are nonnumerical.

- The values of a qualitative variable may or may not have a natural order.
  - Political affiliation.
  - Steak preference.
Summarizing Qualitative Variables

Typically, we use percentages or proportions to summarize qualitative variables.

- 40% of the subjects are Democrats.
- 50% of the people prefer their steak medium.
1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Quantitative Variables

Definition (Quantitative variable)

A **quantitative variable** is a variable whose values are numerical.

- The values of a quantitative variable always have a natural order.
  - A person’s weight.
  - Number of children.
  - Temperature.
Summarizing Quantitative Variables

- Typically, we use averages to summarize quantitative variables.
  - The people in the sample weigh an average of 156.2 lbs.
  - The people in the sample have an average of 2.3 children.
  - The average temperature for the day was $-2.2^\circ \text{C}$ ($28^\circ \text{F}$).
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Some qualitative variables may appear to be quantitative when they are really qualitative.

Rate your child’s intelligence:
- 1  Way below average
- 2  Below average
- 3  Average
- 4  Above average
- 5  Way above average
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Continuous and Discrete Variables

- Quantitative variables fall into two categories.

**Definition (Continuous variable)**

A *continuous variable* is a variable whose set of *possible* values forms a complete interval of real numbers.

**Definition (Discrete variable)**

A *discrete variable* is a variable whose set of *possible* values forms a set of isolated points on the number line.
Continuous Variables

- Typically continuous variables are *measured* quantities.
  - Length
  - Time
  - Area
  - Weight
Discrete Variables

- Typically discrete variables are things that are *counted*.
  - Family size = number of children in the family.
  - Heart rate = number of beats per minute.

- A verbal description usually contains the phrase “the number of.”

- Caution: Would weight (“number of pounds”) be discrete?
Discrete vs. Continuous

- Some data may be mistakenly thought to be discrete.
  - Time = “number of minutes.”
  - Weight = “number of pounds.”
- What if the minutes and pounds were rounded off to the nearest minute or pound.
- What about heart rate, measured in beats per minute?
Outline

1. Experiment Design
   - Control Groups
   - Randomized Design
   - Blinded Experiments

2. Homework Review

3. Types of Variable

4. Qualitative Variables
   - Caution

5. Quantitative Variables
   - Caution
   - Continuous and Discrete Variables

6. Assignment
Assignment

Homework

- Page 182, exercises 25, 26, 28 - 31.
- Chapter 3 review, p. 196, exercises 39 - 43, 45, 47, 49 - 51, 59, 60, 67, 68, 70.
- Read Sections 4.1 - 4.2, pages 212 - 219.
- Let's Do It! 4.1.
- Page 219, exercises 1 - 5.
3.26 (a) treatments are assigned to the subjects.

3.28 Randomly assign 4 to Group I. The remaining 4 are in Group II. Train Group I by the old method and train Group II by the new method. Group I is \{Rover, Spot, Klaus, Charlie\} and Group I is \{Jasmin, Chelsea, Maggie, Bo\}.

3.30 (a) They are already labeled (?)
(b) Take a simple random sample of 10. Assign them to the first group, which gets the aspirin. The remainder will get the placebo.
Page 182, Problems 26, 28, 30

3.30 (c) First group is \{Steve, James, Eric, Ralph, Kyle, Stan, Ed, Herb, Ryan, Pablo\}. The second group is \{Lee, Mark, Bill, Matt, Tim, Robb, Phil, Mike, Doug, Henry\}. 
Page 196, Problems 40, 42, 50, 60, 68, 70

3.40 Treatment 1: 2, 7
Treatment 2: 8, 4
Treatment 3: 3, 6
Treatment 4: 1, 5.

3.42 (a) 18.
(b) 90.

3.50 (a) explanatory variable.
(b) response variable.
(c) a retrospective observational study.
### 3.60

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>An experiment.</td>
</tr>
<tr>
<td>(b)</td>
<td>The teaching method.</td>
</tr>
<tr>
<td>(c)</td>
<td>Yes. The teachers’ personalities.</td>
</tr>
<tr>
<td>(d)</td>
<td>No mention of randomization.</td>
</tr>
<tr>
<td>(e)</td>
<td>( H_0 ): The two teaching methods are equally effective. ( H_1 ): The two teaching methods are not equally effective.</td>
</tr>
<tr>
<td>(f)</td>
<td>It is greater than 0.10.</td>
</tr>
<tr>
<td>(g)</td>
<td>Possibly. If the ( p )-value is 0.12, for example, then the results would not be significant at the 10%, but they would be significant at the 15% level.</td>
</tr>
</tbody>
</table>
3.68  (a) Experiment. The explanatory variables were not manipulated.

(b) Gender is explanatory and car purchase price is response.

(c) A statistic. It’s value is based on the sample, not the population.

(d) confounding.

(e) 0.12.

(f) response.
### Page 196, Problems 40, 42, 50, 60, 68, 70

**3.70**

(a) The individuals were told which movie to watch.

(b) The explanatory variable was the movie that the person watched. The levels were “The Bridges of Madison County,” “The Godfather: Part II,” and the rainforest documentary. The response variables were the person’s hormone levels before, after, and 45 minutes after watching the movie.

(c) Randomly assign individuals to the different movies.
3.70 (d) The hormone level of each individual before the movie was matched with the hormone levels of the same individual after the movie.

(e) Was a randomized design used?
Page 219, Problems 2, 4

4.2 (a) Quantitative discrete.
    (b) Quantitative continuous.
    (c) Quantitative discrete.
    (d) Quantitative continuous.

4.4 No. There could be confounding factors, making the unemployment rate under the combination of proposals unpredictable.