Controlling Bias; Types of Variables

Lecture 11
Sections 3.5.2, 4.1 - 4.2

Robb T. Koether

Hampden-Sydney College

Mon, Feb 6, 2012

Outline

Review Quiz

Example (Review Quiz)

1. An observational study is one in which (select as many as are correct)
   (a) The response variable is observed.
   (b) The explanatory variable is observed.
   (c) The response variable is manipulated.
   (d) The explanatory variable is manipulated.

Review Quiz

Example (Review Quiz)

2. An experimental study is one in which (select as many as are correct)
   (a) The response variable is observed.
   (b) The explanatory variable is observed.
   (c) The response variable is manipulated.
   (d) The explanatory variable is manipulated.

Review Quiz Answers

Example (Review Quiz Answers)

1. (a), (b)
2. (a), (b), (d)

Outline

Experiment Design

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

Experiment Design

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

Experiment Design

- 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.

Experiment Design

- 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

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

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?

Randomized Design

Definition (Randomized design)

A randomized design is a design in which the subjects are randomly assigned to either the treatment group or the control group.

Randomized Design

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

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.

Double-Blind Experiments

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

Types of Variable

- Statisticians like to quantify everything.

Types of Variable

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

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

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?

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

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.
- A proportion is a percentage that is expressed a decimal.

Outline

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° C (28° F).

Outline

Caution

- Some qualitative variables may appear to be quantitative when they are really qualitative.
- The president is doing a fine job.
  1. Strongly agree
  2. Agree
  3. Neutral/no opinion
  4. Disagree
  5. Strongly disagree

Outline

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 people in the family.
- A verbal description usually contains the phrase “the number of.”
- Caution: Would weight (“number of pounds”) be discrete?

Outline

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.