Chapter 4: Designing Studies

Section 4.3
Using Studies Wisely

The Practice of Statistics, 4th edition – For AP*
STARNES, YATES, MOORE
Chapter 4
Designing Studies

4.1 Samples and Surveys
4.2 Experiments
4.3 Using Studies Wisely
Section 4.3
Using Studies Wisely

Learning Objectives

After this section, you should be able to…

✓ DESCRIBE the challenges of establishing causation
✓ DEFINE the scope of inference
✓ DESCRIBE data ethics in designing studies
Scope of Inference

What type of inference can be made from a particular study? The answer depends on the design of the study.

Well-designed experiments randomly assign individuals to treatment groups. However, most experiments don’t select experimental units at random from the larger population. That limits such experiments to inference about cause and effect.

Observational studies don’t randomly assign individuals to groups, which rules out inference about cause and effect. Observational studies that use random sampling can make inferences about the population.

<table>
<thead>
<tr>
<th>Were individuals randomly assigned to groups?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were individuals randomly selected?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference about the population: YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference about cause and effect: YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference about the population: NO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference about cause and effect: YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference about the population: NO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference about cause and effect: NO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Alterate Example – Silence is Golden

Many students insist that they study better when listening to music. A teacher doubts this claim and suspects that listening to music actually hurts academic performance. What can we conclude for each design?

Select a random sample of students from your school to participate in a study. Randomly assign half of the students to listen to music while studying for the entire semester and have the remaining half abstain from listen to music while studying.

With random selection, the results of the study can be applied to the entire population – in this case, all the students at this school. With random assignment, we can conclude that there is a cause-and-effect relationship between listening to music while studying and GPA for all the students at the school.
The Challenges of Establishing Causation

A well-designed experiment tells us that changes in the explanatory variable cause changes in the response variable.

Lack of realism can limit our ability to apply the conclusions of an experiment to the settings of greatest interest.

In some cases it isn’t practical or ethical to do an experiment. Consider these questions:
- Does texting while driving increase the risk of having an accident?
- Does going to church regularly help people live longer?
- Does smoking cause lung cancer?

It is sometimes possible to build a strong case for causation in the absence of experiments by considering data from observational studies.
The Challenges of Establishing Causation

When we can’t do an experiment, we can use the following criteria for establishing causation.

- The association is strong.
- The association is consistent.
- Larger values of the explanatory variable are associated with stronger responses.
- The alleged cause precedes the effect in time.
- The alleged cause is plausible.

Discuss how each of these criteria apply to the observational studies of the relationship between smoking and lung cancer.
Data Ethics

Complex issues of data ethics arise when we collect data from people. Here are some basic standards of data ethics that must be obeyed by all studies that gather data from human subjects, both observational studies and experiments.

**Basic Data Ethics**

- All planned studies must be reviewed in advance by an institutional review board charged with protecting the safety and well-being of the subjects.
- All individuals who are subjects in a study must give their informed consent before data are collected.
- All individual data must be kept confidential. Only statistical summaries for groups of subjects may be made public.
Section 4.3
Using Studies Wisely

Summary

In this section, we learned that...

✓ **Inference about the population** requires that the individuals taking part in a study be randomly selected from the larger population. A well-designed experiment that randomly assigns treatments to experimental units allows **inference about cause-and-effect**.

✓ **Lack of realism** in an experiment can prevent us from generalizing its results.

✓ In the absence of an experiment, good evidence of causation requires a strong association that appears consistently in many studies, a clear explanation for the alleged causal link, and careful examination of possible lurking variables.

✓ **Studies involving humans must be screened in advance by an institutional review board**. All participants must give their **informed consent**, and any information about the individuals must be kept **confidential**.
Looking Ahead…

In the next Chapter…

We’ll learn how to apply the mathematics of chance as a basis for inference.

We’ll learn about

✓ Randomness and Probability
✓ Simulations
✓ Probability Rules
✓ Conditional Probability and Independence