

Chapter 4.3 Lecture Notes & Examples
Using Studies Wisely

Learning Targets

- Determine the scope of inference for a statistical study.

Section 4.3 - Using Studies Wisely (pp. 261-265)

1. Scope of Inference. It is very important to understand what conclusions can be drawn from statistical studies. This is called **scope of inference**.

Case 1: If we randomly sample individuals from a population

We can make inferences (associations) about the population based upon the sample.

Example: The US Census Bureau carries out a monthly Current Population Survey of about 60,000 households. Their goal is to use data from these randomly selected households to estimate the percent of unemployed individuals in the population.

What can be inferred from this study?

We can infer the % of unemployed individuals in the population based upon the same

Case 2: If we randomly assign subjects in an experiment

We can make inferences about cause & effect; (we could even apply cause & effect for the population the subjects came from if they were randomly chosen)

Example: Scientists performed an experiment that randomly assigned 21 volunteer subjects to one of two treatments: sleep deprivation for one night or unrestricted sleep. The experimenters hope to show that sleep deprivation causes a decrease in performance two days later.

What can be inferred from the study?

We can infer cause and effect for this group of volunteers. We are limited on inferring about population due to voluntary sample.

Random selection of individuals from the population -----> allows inference from the population.

Random assignment of individuals to groups -----> allows inference about cause and effect.

((((Both random selection and random assignment introduce *chance* variation into a statistical study. When performing inference, statisticians use the laws of probability to describe this chance variation.))))

		Were Individuals randomly assigned to groups?	
		Yes	No
Were individuals randomly selected?	Yes	Inference about population: <i>yes</i>	Inference about population: <i>yes</i>
		Inference about cause/effect: <i>yes</i>	Inference about cause/effect: <i>no</i>
	No	Inference about population: <i>no</i>	Inference about population: <i>no</i>
		Inference about cause/effect: <i>yes (only about individuals randomly selected)</i>	Inference about cause/effect: <i>no</i>

Team Work: 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. Here are four possible study designs to address this question at your school. In each case the response variable is the students' GPA at the end of the semester. For each design, suppose that the mean GPA for students who listen to music while studying was significantly lower than the mean GPA of students who did not listen to music while studying. What can you conclude from each design?

1. Get all the students in your AP Statistics class to participate in a study. Ask them whether or not they study with music on and divide them into two groups based on their answer to this question.

No random selection → results cannot apply to population, only to AP Stats class
No random assignment → can't infer cause & effect

Conclusion: AP students listen to music ⇒ lower grade we do not know why

2. Select a random sample of students from your school to participate in a study. Then divide them into groups as in Design 1.

Random sample → results apply to population (entire school)
No random assignment - can't infer cause & effect

Conclusion: Students at this school who listen to music have lower GPA's than those who don't, we don't know why.

3. Get all the students in your AP Statistics class 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 abstain from listening to music while studying.

No Random Selection - results cannot apply to population, only AP Stats class students

Random assignment → we can infer cause and effect, but only for AP Stats class

Conclusion: Listening to music causes Lower GPAs for AP Stats students in this study

4. 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 listening to music while studying.

Random sample - results apply to population (entire school)

Random assignment - can infer cause & effect

Conclusion: Listening to music causes lower GPA for ^{all} students at this school.

2. The Challenges of Establishing Causation - A well-designed experiment tells us that changes in the explanatory variable cause changes in the response variable. More precisely, it tells us that this happened for specific individuals in the specific environment of this specific experiment.

The serious threat is that the treatments, subjects, or the environment of the experiment is not *realistic*. Lack of realism can limit our ability to apply the conclusions of an experiment to the settings of greatest interest.

Examples

- Determining if texting while driving causes accidents?
- Does smoking cause lung cancer?
- Does not wearing seatbelts cause more fatalities than wearing seatbelts?

What are the criteria for establishing causation when we cannot do an experiment?

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