Introduction to Statistics

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

August 19, 2019

Notes

Topics

Notes

Objectives

Objectives:

- Identify the explanatory and response variables in a study.
- Distinguish between observational studies and experiments.
- State which type of study can establish cause-and-effect.
- State the three principles of experimental design.

Notes

Collecting Data: Experiments and Observational Studies (1.4)

Explanatory and Response Variables

- Many studies are carried out to examine how two variables are *related* to each other. For example:
  - Is a person's risk of a heart attack related to his or her weight?
  - Does a person's income depend on his or her gender?
  - Are crime rates related to the state of the economy?

Notes

- In such studies, the variables usually play different roles: The ***explanatory variable*** is the one that *explains* differences or *causes* changes in the ***response variable***.

Notes

Exercise

For each of the following studies, decide which is the **explanatory variable** and which is the **response**:

a) A study examines the relationship between the weight (in lbs) of a car and the car's gas mileage (in mpg).

b) A study is designed to determine if aspirin usage reduces blood pressure.

c) A study is designed to determine if eating food rich in antioxidants (such as fruits and vegetables) reduces the risk of colon cancer?

d) A study examines how the number of hours spent studying for an exam is related to the score on the exam.

Notes

Observational Studies versus Experiments

- Studies that examine how two variables are *related* can be either of two types:
  - Observational studies
  - Experiments

Notes

- An ***observational study*** is a study in which the investigator merely *observes* individuals in order to draw conclusions about how the two variables are related.

  No attempt is made to cause changes in the response variable by manipulating the explanatory variable. Instead its relation to the explanatory variable is just observed

Notes

- An ***experiment***, on the other hand, is a study that **involves a deliberate attempt to cause changes in the response variable** by imposing *treatments* on individuals.

  The goal of an experiment is to establish a **cause-and-effect** relationship between the treatments and the response.

Notes

- In experiments,the ***treatments*** are different *levels* of the explanatory variable, which is sometimes called the ***factor***.

Notes

Exercise

Decide whether each study below is an **observational study** or an **experiment**.

If it's an *observational study*, determine what the **explanatory** and **response variables** are.

If it's an *experiment*, determine what the **explanatory variable** (**factor**) is, what the **treatments** are, and what the **response variable** is.

a) A study in the British Journal of Medicine found that ”owls” – people who go to sleep after 11 p.m. and rise after 8 a.m. – tend to have higher incomes than ”larks” - people who go to sleep before 11 p.m. and rise before 8 a.m.

Notes

b) In a study of whether vitamin C helps prevent colds, a group of 100 subjects was divided randomly into two groups of 50 each, with one group receiving daily vitamin C and the other, acting as *controls*, receiving only a daily *placebo*.

Notes

c) In a study to determine whether college students learn better when listening to music, participants are randomly divided into 4 groups:

Group 1: Classical music
Group 2: Rock music
Group 3: Jazz
Group 4: No music (just silence)

The subjects are asked to read a chapter from a history textbook during the music or silence, and then given a test on the material.

Notes

d) A group of smokers and a group of nonsmokers who visited a particular clinic were asked to come in for a physical exam every five years for the rest of their lives so that the researchers could monitor and compare their heart rates.

Notes

e) "The patch" is a device that dispenses nicotine into the blood when applied to the skin.

In a study at the Mayo Clinic, 240 smokers aged 20 to 65 were divided randomly into two groups. The first group used "the patch" for eight weeks and the second used a "placebo patch" that resembled the real patch, but dispensed no nicotine.

The subjects in the study were blinded as to whether they were receiving "the patch" or a placebo (i.e. they were not told which treatment they were receiving). After the eight weeks, each subject was asked if he or she was able to quit smoking.

Notes

- **An observational study, by itself, *cannot* establish cause-and-effect relationships** due to the possible presence of *"lurking"* variables.
- A ***"lurking"*** variable is a third variable whose effect on the response can't be distinguished from the effect (if any) of the explanatory variable because the "lurking" variable is related to *both* the explanatory variable *and* the response.

Notes

Exercise

Suppose, for each of the world's nations, we observe the number of TV sets per capita and the average life expectancy.

We'd find that nations with more TV sets have higher life expectancy.

a) Is this study an experiment or an observational study?

b) Could we conclude from the study that owning more TVs *causes* a longer life?

Notes

c) List a possible "lurking" variable that provides a better explanation for the relationship between TVs and life expectancy. ***Hint***: Try to identify another characteristic that distinguishes nations that have more TVs per capita from those that have fewer.

Notes

Exercise

An observational study showed that people who eat food rich in antioxidants (such as fruits and vegetables) have lower rates of colon cancer than those who don't eat these foods.

a) Can we conclude that eating such foods *reduces* the risk of colon cancer?

b) List a few possible "lurking" variables that provide equally plausible explanation for the lower rates of colon cancer than cause-and-effect. ***Hint***: Try to identify other ways in which people who eat lots of fruits and vegetables might differ from those who don't.

Notes

- **In order to establish a cause-and-effect relationship between two variables, we need to carry out an experiment**, as in the next example.

Notes

Exercise

A clinical trial (experiment on humans) to study the effect of dietary antioxidants (vitamins A, C, and E) on colon cancer rates *randomly* divided 864 subjects into four treatment groups. Each group was given a different amount of antioxidants in their diet:

Group 1: Daily beta carotene (vitamin A)
Group 2: Daily vitamins C and E
Group 3: All three vitamins daily
Group 4: No vitamin supplements.

After 4 years, researchers were surprised to find no significant difference in colon cancer among these groups.

Notes

- Note:
  1. In the experiment in the last example, because individuals were **randomly assigned** to treatment groups, the groups were **similar** to each other with respect to variables (such as health consciousness, amount of exercise, etc.) that might differ from one group to the next in an observational study (i.e. variables that might be ”lurking”).
  2. In this way, the only real difference between the groups was which treatment (vitamin combination) they received.
  3. Thus if we observe significant differences in the groups’ colon cancer rates, those differences *must* be due to (i.e. have been **caused** by) differences in the effects of the treatments (vitamins), as opposed to being due to a ”lurking” variable.

Notes

- Randomly assigning individuals to treatment groups is called ***randomization***. It helps assure that the groups are similar to each other prior to receiving the treatments.
- A properly designed experiment should include some form of ***control*** so that the effect, if any, of the treatment can be compared to the effect of receiving no treatment at all.

  This is usually done by using a ***control group*** that receives a dummy treatment (i.e. a ***placebo***) or no treatment at all.
- The following two examples illustrate what can go wrong if the experiment doesn’t include a control group.

Notes

Exercise

Ulcer pain is the result of too much stomach acid. ”Gastric freezing” is a clever treatment for ulcers in which the patient swallows tubes through which a cold fluid is pumped to cool the stomach and, perhaps, reduce stomach acid.

In a poorly designed experiment, a group of patients was treated with gastric freezing, and they reported that their pain subsided.

The results were reported in the *Journal of the American Medical Association*, and gastric freezing gained widespread usage.

Notes

But the patients' response may have been due to the *placebo effect*. A *placebo* is a dummy treatment. Many patients respond favorably to any treatment, even a placebo, due to trust in the doctor or expectations of a cure. The response to a placebo is called the ***placebo effect***.

In this study, the placebo effect was "lurking" in the background and couldn't be distinguished from the effect (if any) of the treatment.

Notes

A later experiment randomly divided ulcer patients into two groups. One group was treated with gastric freezing, and the other with a placebo treatment in which the fluid pumped through the tubes was at body temperature.

This new study found that indeed 34% of the 82 patients in the treatment group reported experiencing pain relief. But so did 38% of the 78 patients in the control group. Gastric freezing was shown to be no better than a placebo, and its use was soon abandoned.

Notes

Exercise

In a poorly designed experiment to test the effect of fertilizer on corn yield, it was found that this year, when fertilizer was used, the yield was much higher than in years past, when fertilizer wasn't used.

But if this year's weather conditions were much more favorable than in past years, then the effect of weather is "lurking" in the background, and cant' be distinguished from the effect of fertilizer.

Notes

A better experimental design would involve comparing the yield on a plot where fertilizer was used with that on a plot where fertilizer wasn't used so that the favorable weather acts equally on both plots. Here the plot without fertilizer acts as a control much like the placebo group did in the ulcer example.

Notes

- **Three Principles of Experimental Design**: The previous examples help illustrate the importance of the ***Three Principles of Experimental Design***:

1. ***Randomization*** - Random assignment of individuals to treatment groups so that the groups are similar to each other prior to receiving the treatments.
2. ***Control*** - Including in the experiment a control group that receives a placebo or no treatment at all so that the effect of the treatment can be compared to the effect of no treatment.
3. ***Replication*** - Using enough individuals in the experiment so that the results are reliable.

Notes

Notes

Notes

Notes