Experiments and Observational Studies

Ryan Miller

Study Design

- In the candy weighing activity we saw that how a sample is collected is critically important in getting accurate information about a population
- Today we will see that even when our data are a *representative sample*, or an *entire population*, accurately answering a research question can still be difficult
- We'll explore these challenges using a few detailed case studies

Case Study - Racial Bias in Florida Courts

A 1981 study (Link) analyzed murders that took place during a felony in Florida between 1972 and 1977. The study recorded whether or not the offender was sentenced to the death penalty and their race:

	Death Penalty	Murder during a Felony
White Offender	46	198
Black Offender	38	180

Question: Did the court's sentencing have a racial bias?

Death penalty rates were 23.2% (46/198) for white offenders and 21.1% (38/180) for black offenders

Association and Causation

Before providing a definitive answer to that question, let's take a step back and think about the features of these data and how they relate to our research question

The previous slide relates two variables: an explantory variable - "the race of the offender" and an outcome variable - "death penalty sentencing"

- Two variables said to be **associated** if certain values of one variable are related with certain values of the other variable across cases
- Two variables said to be **causally associated** if changing the value of one variable (for a given case) directly influences the value of the other variable
- Which is a better description of the relationship shown on the last slide? Can we be sure?

Revisiting the Florida Racial Bias Data

The researchers in the study were careful to collect more variables than just the two shown in previous table. Below is a new table that includes information on the race of the victim:

	White Victim		Black Victim	
	Death	Murders	Death	Murders
White Offender	46	190	0	8
Black Offender	37	78	1	102

Our analysis can now **condition** upon the race of the victim:

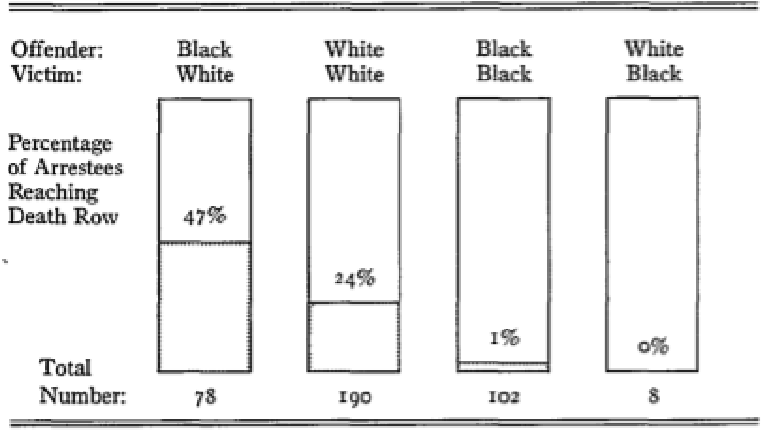
- In cases involving a white victim, the death penalty rate for white offenders was 24.2% (46/190) and 47.4% (37/78) for black offenders
- In cases involving a black victim, the rate was 0% for white offenders and 1% for black offenders

What just happened?

Visualization

FIGURE 2

RATIO OF OFFENDERS ON FLORIDA DEATH ROW TO THE NUMBER OF FLORIDA ARRESTS FOR MURDER DURING A FELONY, BY RACE OF OFFENDER AND RACE OF VICTIM



Confounding Variables

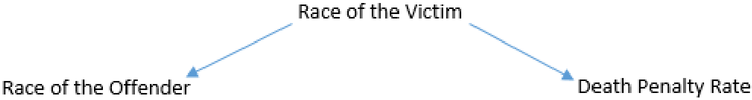
Race of the victim is *strongly associated* with both:

1.) The explanatory variable - the race of the offender

- 65% of white victims had white offenders, only 7% of black victims had white offenders

2.) The response variable - the rate of death penalty sentencing

- 28.4% of cases with a white victim resulted in a death sentence, only one death sentence ($< 1\%$) occured for black victims



We accounted for these associations by *conditioning*, specifically we used a technique called **stratification** to split up the cases by the confounding variable: Stratification Link

Confounding Variables

In this example, the race of the victim is a **confounding variable**, a third variable that is associated with both the explanatory variable and the response variable. Confounding variables will obscure the relationship between explantory and response variables.

How can we establish causal relationships in the face of confounding?

- Careful study design!
- Statistical techniques? (*Causal Inference* is currently a hot research area in statistics)

Randomized Experiments

If we could somehow force the confounding variable (race of the victim) to be unrelated with the explanatory variable our problems would be solved

- We did this using stratification, but what if there are more confounders?
- A way to prevent *all* confounding is to *randomly assign* values of the explanatory variable
- In our example this would, on average, equalize the proportions of white/black victims for both races - breaking the link between the explanatory variable and the confounder
- **Randomized experiments**, where the values of the explanatory variable are randomly assigned, are not always possible
- Often the only feasible approach to collecting data are **observational studies**, which are inherently troubled by confounding

Discussion - Planning an Experiment

Suppose we want to know: "Is arthroscopic surgery is effective in treating arthritis of the knee?" Theorize an *observational study* and a *randomized experiment* you could conduct to answer this question. Be sure to address the following during your discussion:

1. How costly will it be for the researchers to collect data with each design?
2. Are there any feasibility problems or ethical issues with each design?
3. What, if any, confounding variables might be problematic with each design?

Sham Knee Surgery

In the 1990s a study was conducted in 10 men with arthritic knees that were scheduled for surgery. They were all treated identically expect for one key distinction: only half of them actually got surgery! Once each subject was in the operating room and anesthetized, the surgeon looked at a randomly generated code indicating whether he should do the full surgery or just make three small incisions in the knee and stitch up the patient to leave a scar. All patients received the same post-operative care, rehabilitation, and were later evaluated by staff who didn't know whether they had actually received the surgery or not. The result? Both the sham knee surgery and the real knee surgery showed indistinguishable levels of improvement

Source: https://www.nytimes.com/2000/01/09/magazine/the-placebo-prescription.html

Control Groups, Placebos, and Blinding

The Sham Knee Surgery example illustrates several important aspects of a well-designed experiment that we've yet to discuss:

- **Control Group** - Some patients were randomly assigned not to receive the knee surgery, providing a comparison group that is, on average, balanced with surgery group in all baseline characteristics
- **Placebo** - Patients in the control group received a fake surgery
- **Blinding** - Using a placebo is not helpful if patients know which group they're in. Similarly, the staff interacting with the patients might treat them differently if they knew the patient's group
  - **Single-blind** - the participants don't know the treatment assignments
  - **Double-blind** - the participants *and* everyone interacting with the participants don't know the treatment assignments

The Realities of Randomized Experiments

- Randomized, placebo-controlled, double-blind experiments are considered the *gold standard* in research, but they aren't always possible
- When determining how to collect data you must assess the practical and ethical considerations
  - It is functionally impossible to implement randomization in the Florida racial bias study
  - It is ethically inappropriate to randomize harmful explanatory variables like exposure to hazardous chemicals
- In these scenarios, observation studies at least help us determine associations
- How do you think smoking was "proven" harmful?

Discussion Example

- University of Iowa researcher was conducting an experiment on lab monkeys
- Because these monkeys are expensive, his experiment consisted of 2 groups of 4
- Having taken a statistics course, the researcher randomly assigned treatment/control groups
- After conducting the experiment and getting suprising results, the research recognizes that the 4 monkeys in the control group were also the oldest 4 monkeys
- The researcher knew that the age of the monkey had an important relationship with the outcome, but he expected randomization to handle that

Should he report his results? What could he have done differently?

Minneapolis Police Case Study

Even if you can implement randomization, it might not produce the intended results:

- Police departments have long been uncertain about how to best respond to cases of domestic abuse
- Minneapolis Police conducted a study comparing three different response strategies:
    - Arrest
    - Advice
    - Seperate

- Officers were randomly assigned a strategy to use on each case, but they were allowed discretion to change the strategy if necessary - Precautions were taken to ensure the officers were as faithful as possible to each assigned strategy - The outcome of interest was whether or not violence reoccured

Minneapolis Police Case Study

The columns of the table indicate the strategy assigned, the rows indicate the strategy actually used:

	Arrest	Advice	Seperate
Arrest	91	18	26
Advice	0	84	5
Seperate	1	6	82

Overall there was 82% adherence to the randomly assigned strategy, but do you see any problems?

Minneapolis Police Case Study

- A common pattern was for officers to "upgrade" to the arrest strategy
- The advice and seperate groups were likely losing some of their highest risk members to the arrest group
- This seemingly well-designed experiment ended up needing to be bailed out by complex statistical approaches that jointly modeled new violence (the outcome variable) and adherence to the randomized strategy

Minneapolis Police Article Link

Intention-to-Treat Principle

- The Food and Drug Administration (FDA) mandates an **intent-to-treat principle** as the the primary design and analysis strategy for clinical trials
- This means that all subjects who are randomized be included in the final analysis, even if they, cross-over, do not adhere to any protocol, or drop out of the study
- It also means that clinical trials estimate the effect of the *treatment assignment* rather than the treatment itself

Intent-to-treat Article Link

Conclusion

Right now you should...

1. Recognize that not every association implies causation
2. Be able to identify and describe potential confounding variables
3. Know the strengths and weaknesses of randomized experiments vs observational studies
4. Feel comfortable designing and implementing simple randomized experiments

These notes cover Section 1.3 of the textbook, I encourage you to read through the section and its examples