Experiments and Observational Studies

Ryan Miller

Randomized Experiments

- Lab 2 provided an example where a **randomized experiment**, where the values of the explanatory variable are randomly assigned, could not easily be conducted
- Often, the only feasible approach to collecting data are **observational studies**, which are inherently troubled by confounding variables
- Even still, randomized experiments are the *ideal way* to answer a research question, so we should study some of their nuances

Discussion - Planning an Experiment

Suppose we want to know: "Is arthroscopic surgery is effective in treating arthritis of the knee?" Outline an *observational study* and a *randomized experiment* that 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, observational studies at least help us determine associations
- How do you think smoking was "proven" harmful?

Discussion - Can Randomization Fail?

- University of Iowa researcher was conducting an experiment on lab monkeys
- Monkeys are expensive, so 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 seeing surprising results, the researcher 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?

Discussion - Can Randomization Fail?

- Randomization is not guaranteed to negate confounding variables unless the sample size is relatively large
- At smaller sample sizes, alternative approaches, such as a **matched pairs** randomization design, are advantageous
  - In a matched pairs experiment, cases are paired based upon a **blocking variable**
  - Within each pair, one case is assigned to the treatment group, the other case to the control group
- In the previous example, monkeys could have been matched based upon the variable "age", then treatment could be randomized within pairs

Minneapolis Police Case Study

Even with a large sample, randomization 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 given 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 reoccurred.

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 Article Link

Minneapolis Police Case Study

- A common pattern was to "upgrade" to the arrest strategy
- The advice and separate groups likely lost their highest risk members to the arrest group
- This seemingly well-designed experiment ended up needing to be bailed out by complex statistical approaches used to jointly model reoccurance of violence (the outcome variable) and adherence to the randomized strategy

The Intention-to-Treat Principle

- The Food and Drug Administration (FDA) mandates an **intent-to-treat principle** (ITT) as 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

Intention-to-Treat Example

Below are the results of the MN police case study *as randomized*:

	Recurrence	No Recurrence
Arrest	10	82
Advice	24	84
Seperate	26	87

Below are the results *by treatment used*:

	Recurrence	No Recurrence
Arrest	18	117
Advice	16	73
Seperate	26	63

With your group, compare the difference in recurrence proportions of the Advice and Separate treatments using an ITT analysis. Compare this with the same analysis done using the treatments used.

Intention-to-Treat Example - Solution

Using ITT, $p_{re|advice} = 24/106 = 0.22$ and
$p_{re|separate} = 26/113 = 0.23$

- We can conclude that telling an officer to use the "Advice" strategy leads to essentially the same outcomes as telling the officer to use the "Separate" strategy

Using as treated, $p_{re|advice} = 16/89 = 0.18$ and
$p_{re|separate} = 26/89 = 0.29$

- We have trouble concluding anything; while there is a clear difference, we don't know if it is due to the strategies themselves, or a disproportionate switching of high/low risk cases into different strategies

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

If you want more information:

- Read Ch 1.3
- Randomized Experiments vs. Random Sampling