Study Design and Multivariate Relationships

Ryan Miller

Outline

1. Confounding Variables
2. Experimental vs. observational studies
3. Examples and common pitfalls

Multivariate relationships

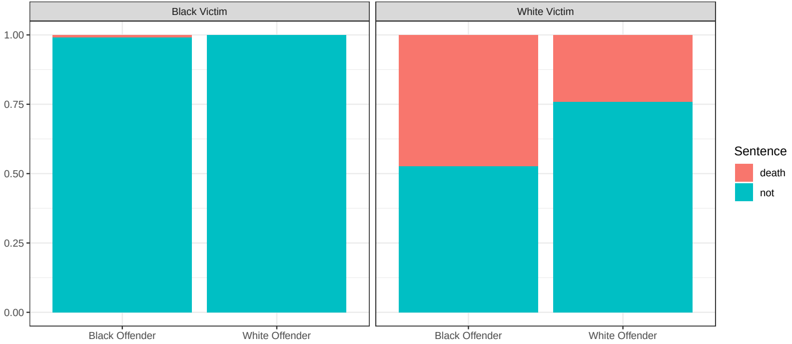
In Lab #2, we analyzed data describing death penalty sentencing in Florida during the 1970s:

	death	not
black	38	142
white	46	152

- White offenders received the death penalty in 23.3% of cases, while black offenders received death penalty in only 21.1% of cases
  - What problem arises when concluding that sentencing was biased against white offenders?

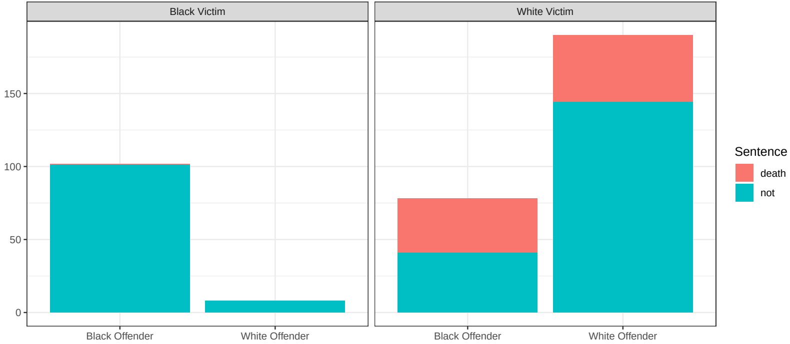
Multivariate relationships

Overall, white offenders received the death penalty slightly more often, but this ignored the influence of the victim's race:



Confounding Variables

Because offenders *disproportionately* committed crimes against victims of their own race, the overall death penalty rates were skewed in a way that obscured the racially biased sentencing:



Balance

- We can view the problems caused by confounding variables as an issue of **imbalanced groups**
  - Offenders were more likely to victimize their own race, and crimes against whites tended to be punished more severely
  - The groups white offenders and black offenders were systematically different in an important way (victims race)

Balance

- We can view the problems caused by confounding variables as an issue of **imbalanced groups**
  - Offenders were more likely to victimize their own race, and crimes against whites tended to be punished more severely
  - The groups white offenders and black offenders were systematically different in an important way (victims race)
- We can try to correct for imbalances using techniques like *stratification*
  - However, a better way to eliminate *all* confounding variables is to use an entirely different type of study design

Two Types of Studies

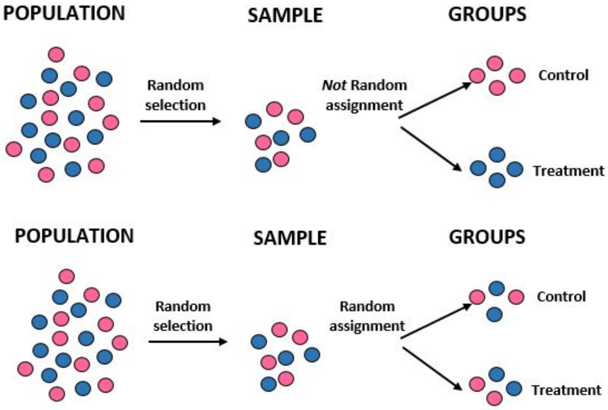
There are two broad types of studies:

- **Observational studies**: the explanatory and response variables are *observed* by the researchers (separate samples)
- **Experimental studies**: the explanatory variable is *assigned* by the researchers (the researchers split up a single sample)

The Florida Death Penalty data provide an example of an observational study (since the researchers merely observed the race of the offender, they did not assign it to cases)

Random Assignment

If the explanatory variable is *randomly assigned*, both groups will be *balanced* in terms of *all* possible confounding variables.



Barriers

- Random assignment isn't always feasible, some variables are unethical or too costly to randomly assign
  - For example, we couldn't assign cases to consume toxic chemicals or expose themselves to harm
  - We also cannot randomly assign explanatory variables that pre-date the study like genetics, etc.

Barriers

- Random assignment isn't always feasible, some variables are unethical or too costly to randomly assign
  - For example, we couldn't assign cases to consume toxic chemicals or expose themselves to harm
  - We also cannot randomly assign explanatory variables that pre-date the study like genetics, etc.
- Despite their flaws, observational studies are still very valuable
  - But they will always fall short of *randomized experiments*

Example #1 - Experiment vs. Observational?

Suppose we want to know: "Is arthroscopic surgery is effective in treating arthritis of the knee?" Describe both 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?

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 the 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

Example #2 - Can Randomization Fail?

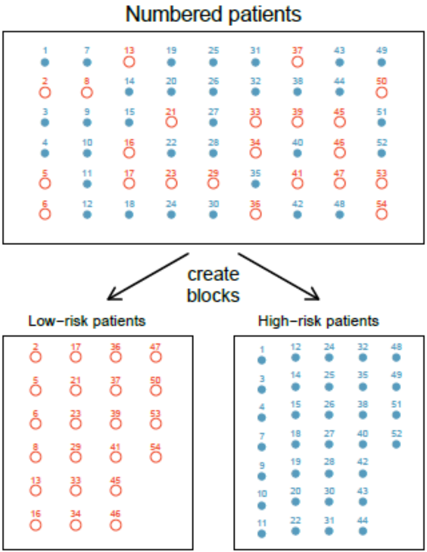
- A University of Iowa researcher was conducting an experiment on lab monkeys
- Lab monkeys are expensive, so his experiment had $n = 8$
- Having taken a statistics course, he randomly assigned treatment/control groups
- After conducting the experiment and seeing surprising results, the researcher recognizes 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 on the outcome variable, but he expected randomization to handle that

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

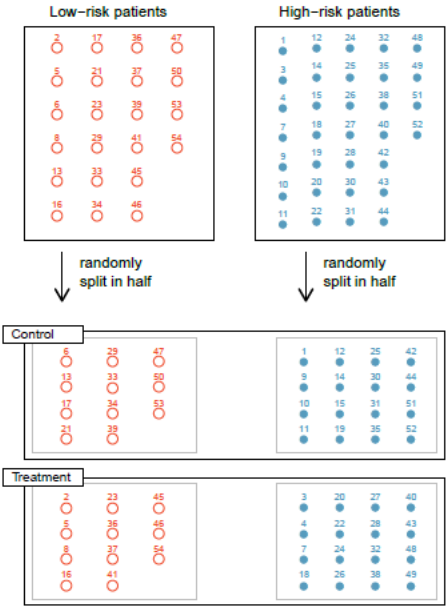
Blocking

- Randomization is not guaranteed to properly balance the treatment and control groups unless the sample size is relatively large
- At smaller sample sizes, strategies such as **blocking** can be used
  - In this design, cases are first split using a **blocking variable**, then random assignment is done within each block
  - This ensures the blocking variable is balanced in each group

Blocking


Numbered patients
create blocks
Low–risk patients
High–risk patients

Blocking



Low–risk patients
2 17 36 47
5 21 37 50
6 23 39 53
8 29 41 54
13 33 45
16 34 46
High–risk patients
1 12 24 32 48
3 14 25 35 49
4 15 26 38 51
7 18 27 40 52
9 19 28 42
10 20 30 43
11 22 31 44
randomly
split in half
randomly
split in half
Control
6 29 47
13 33 50
17 34 53
21 39
1 12 25 42
9 14 30 44
10 15 31 51
11 19 35 52
Treatment
2 23 45
5 36 46
8 37 54
16 41
3 20 27 40
4 22 28 43
7 24 32 48
18 26 38 49

Example #3 - Limitations of Randomization

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

Example #3 - Limitations of Randomization

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

- The overarching goal of a statistician is to *rule out* as many possible explanations for an observed association as possible

Conclusion

- The overarching goal of a statistician is to *rule out* as many possible explanations for an observed association as possible
- So far we've considered the following design-related explanations, as well as methods for addressing for them
  - Sampling bias - Simple random sampling
  - Confounding variables - Random assignment of the explanatory variable, or stratification
  - Other biases - Using placebo, double-blinding, proper measurement, etc.