Study Design (part 1)

Ryan Miller

Introduction

- So far we've learned that the procedure used to obtain a sample is critically important, but this is not the only factor worthy of considering in the design of a study
- Suppose researchers develop a new COVID-19 treatment, how would you design a study to determine if it is effective or not?

Introduction

- So far we've learned that the procedure used to obtain a sample is critically important, but this is not the only factor worthy of considering in the design of a study
- Suppose researchers develop a new COVID-19 treatment, how would you design a study to determine if it is effective or not?
- The only meaningful designs will compare the new treatment with something else
  - Therefore, we'll either need to use two samples, or emply a method for splitting a single sample into two

Two Types of Studies

This leads us to distinguish between two types of studies:

- **Observational studies**: the explanatory and response variables are *observed* by the researchers
- **Experimental studies**: the explanatory variable is *assigned* by the researchers

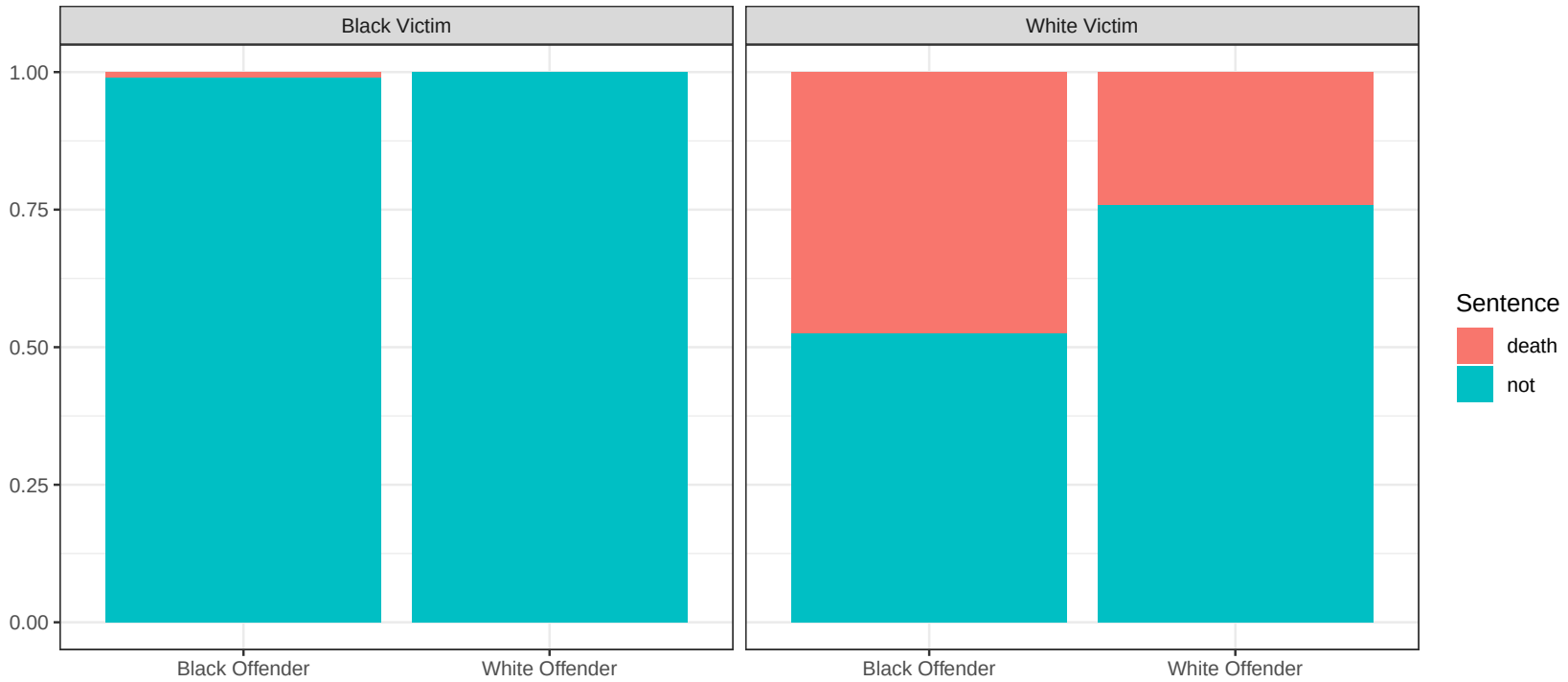
Observational Studies

- We've already seen an example observational study in the Florida Death Penalty Sentencing case study
- The researchers recorded the race of the offender, as well as whether the offender was sentenced to the death penalty or not
  - Did the offender's race appear to be associated with their sentence?

	death	not
black	38	142
white	46	152

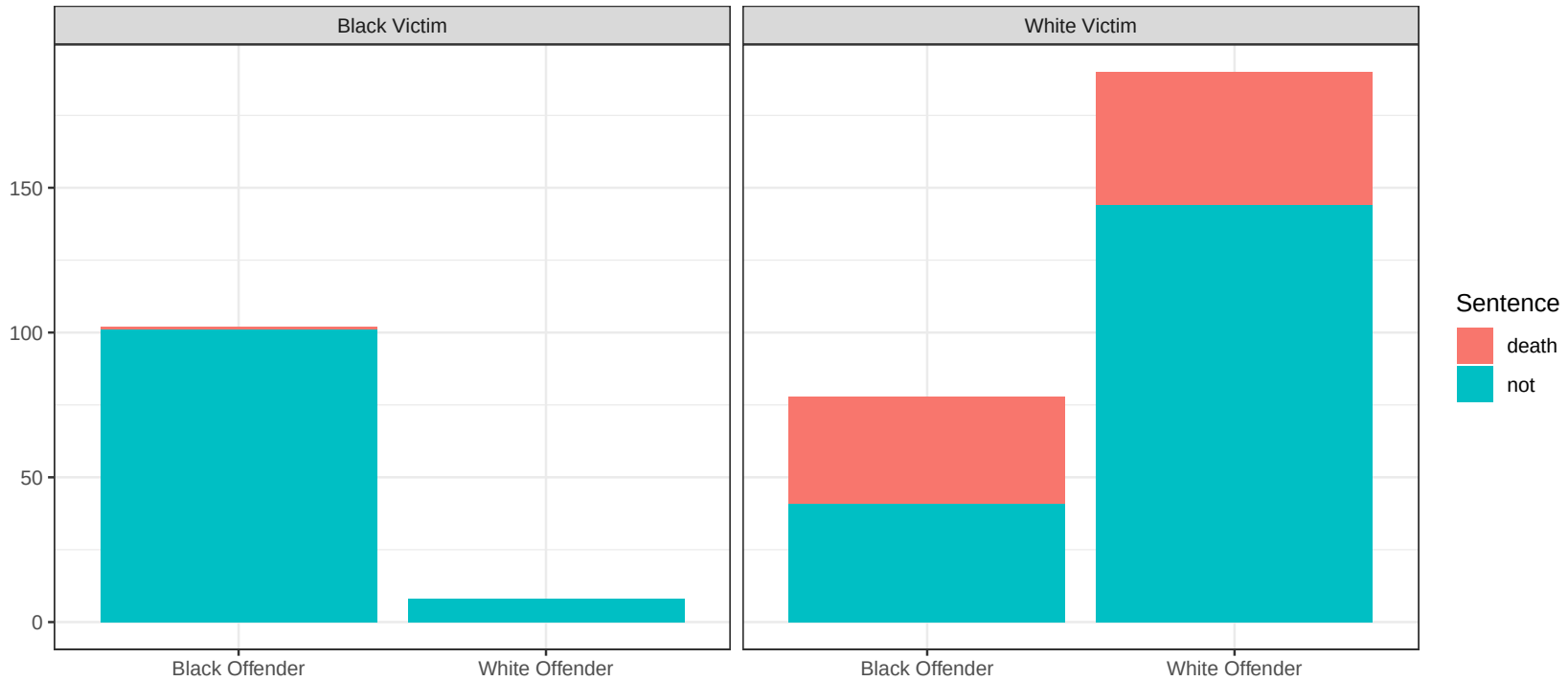
Confounding Variables

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:



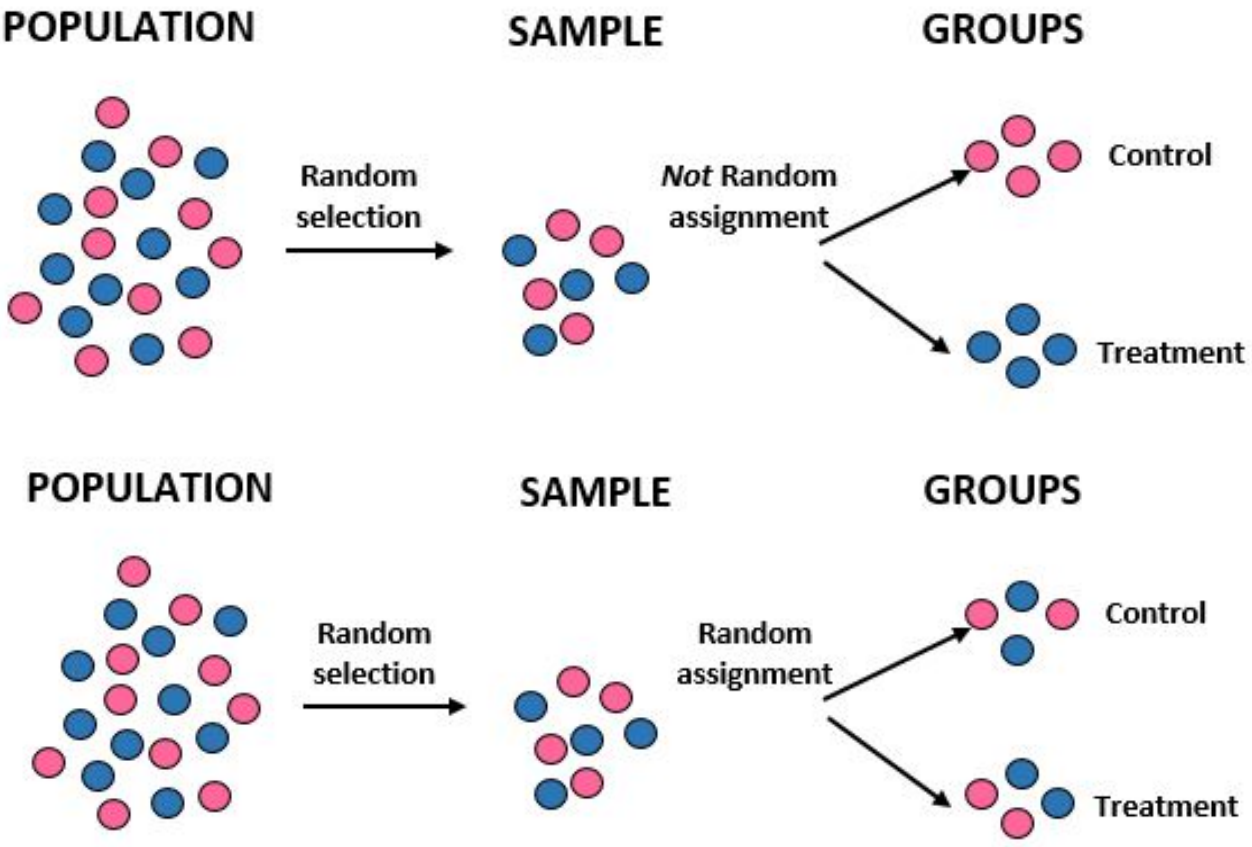
Balance

- 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)
- Going back to the COVID-19 example, if study participants are allowed to choose whether they receive the vaccine we might expect the vaccine group to be older, sicker, working riskier jobs, etc.
  - However, these factors would all occur in equal proportions in the vaccinated and control groups if we **randomly assigned** which participants received the vaccine

Random Assignment



Barriers

- Obviously random assignment isn't always feasible, some explanatory variables are too unethical or 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 universally pre-date the study like genetics, etc.

Barriers

- Obviously random assignment isn't always feasible, some explanatory variables are too unethical or 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 universally pre-date the study like genetics, etc.
- Despite their flaws, observational studies are very valuable
  - But they will always fall short of *randomized experiments*