Causal Infernece: Lab 1

Spring 2026

Xiangyu Song Chanhyuk Park

Washington University in St. Louis

Introduction

In this lab, we explore fundamental concepts in causal inference: distinguishing between the **Average Treatment Effect (ATE)** and the **Average Treatment Effect on the Treated (ATT)**, and dissecting potential sources of bias in our estimators (particularly **selection bias** and **heterogeneous treatment effects**).

We begin with a **job-training program** example featuring bankers and doctors in a futuristic scenario, then pivot to a more abstract medical example with potential outcomes. Throughout, you'll see why **random assignment** is so crucial for obtaining unbiased estimates of treatment effects.

Setup

```
{r setup, include=FALSE} knitr::opts_chunk$set(echo = TRUE)
knitr::opts_knit$set(results.folding=NULL)
```

```
{r packages} library(tidyverse)
```

ATE vs. ATT with Bankers & Doctors

In our futuristic scenario, **bankers** have largely lost their jobs to AI-enhanced ATMs, while **doctors** still have robust employment. We imagine a job-training program that aims to retrain bankers as programmers. However, since doctors aren't eligible for this program, including them in the overall population can lead to misunderstandings about the **treatment effect** if we simply do a naive difference-in-means. This section highlights:

1. **Why the naive ATE might be misleading** if our treated group (bankers) differs substantially from our control group (a mix of doctors and bankers).

2. **How constructing a control group that resembles the treated group** (bankers only) leads to an estimate closer to the **ATT**—often the real-world quantity of interest.

---

1. Constructing the Population

"'{r constructing-population} # Set a seed for reproducibility set.seed(123)

Create vectors of doctor salaries ~ U[160, 260] and banker salaries ~ U[43, 64]

doctor <- runif(1000, 160, 260) bankers <- runif(1000, 43, 64)

Create a tibble of workers

workers <- tibble(job = c(rep('doctor', 1000), rep('banker', 1000)), salary = c(doctor, bankers), id = 1:2000)

```
Here, we simulate a population with distinctly different salary distributions for doctors vs

---

## 2. Creating Treatment and Control Groups

```{r treatment-control-groups}
set.seed(123)

# Draw a sample of 500 bankers to be treated
treated_rows <- sample(1001:2000, size = 500)

# The remaining rows
drop_treats <- workers[-treated_rows,]

# Draw a control sample from the full population (including doctors)
control <- sample_n(drop_treats, size = 500)

# Treatment group
treated <- workers[treated_rows,]
```
```

Notice that only bankers can receive the treatment, reflecting real-life situations where certain programs are only available to certain subgroups. But our control group is drawn from everyone, including higher-earning doctors. Keep an eye on how this affects our estimated treatment effect.

---

3. Applying the Treatment Effect

"'{r applying-treatment-effect} # Treatment effect: # - If salary < 60, uniform increase between 30-50 # - If salary >= 60, uniform increase between 0-15 set.seed(123) treat_eff <- vector() for(i in 1:nrow(treated)){ treat_eff <- append(treat_eff, unlist(ifelse(treated[i,'salary'] < 60, treated[i,'salary'] + runif(1, 30, 50), treated[i,'salary'] + runif(1, 0, 15)))) }

Control group salaries remain unchanged

control_sals <- control$salary

Combine treatment and control into one experimental dataset

exp_group <- bind_rows(treated, control) exp_group$new_sal <- -c(treat_eff, control_sals)exp_grouptreated <- c(rep(1, 500), rep(0, 500))

```
Almost all bankers get the "good" treatment effect (a large salary boost), whereas doctors-i

---

## 4. Exploring the ATE vs. the ATT
(Hint: Check the summary statistics for the treatment and control groups, check the simple d
```{r exploration-ate-att}
# Your Code Here
```

Decomposing the Difference in Means

We now switch gears to a smaller-scale **medical scenario** that lets us peek directly into **potential outcomes** (both treatment and control outcomes for each individual). Of course, in real life, we never observe both potential outcomes for the same person; we see only one, making causal inference tricky. This section breaks down how **selection bias** and **heterogeneous treatment effects** each can corrupt naive difference-in-means estimates.
```

---

1. Setting Up Potential Outcomes for 10 Cancer Patients

“‘{r potential-outcomes-small} po_patients <- tibble(patients = c(1:10), y_i1 = c(7 , 5, 5, 7, 4, 10, 1, 5, 3, 9), y_i0 = c(1, 6, 1, 8, 2, 1, 10, 6, 7, 8), po_diff = y_i1 - y_i0)

po_patients

True ATE from potential outcomes

ate <- mean(po_patients$po_diff) ate

```
Here, `y_i1` is how many years a patient might live if assigned to an experimental surgery (

---

## 2. Sorting Everyone Into Their Best Outcome (Generating Bias)

```{r best-outcomes}
po_patients2 <- po_patients %>%
  mutate(
    D = if_else(po_diff >= 0, 1, 0),
    Y = y_i1 * D + y_i0 * (1 - D)
  )

po_patients2

# Compute the actual ATT
att <- po_patients2 %>%
  filter(D == 1) %>%
  mutate(tt = Y - y_i0) %>%
  pull(tt) %>%
  mean()

att

# Compute the actual ATC
atc <- po_patients2 %>%
  filter(D == 0) %>%
  mutate(tt = y_i1 - Y) %>%
  pull(tt) %>%
  mean()
```
```

```
atc
```

By giving everyone the treatment only if it's beneficial to *them*, we're introducing strong correlation between potential outcomes and treatment assignment—exactly how **selection bias** arises when sicker (or healthier) patients are more (or less) likely to receive treatment in a real-world setting.

---

3. Estimating Using the Simple Difference in Means

"'{r simple-diff-means} po_patients2 %>% group_by(D) %>% summarise(diff = mean(Y))

diff_means <- po_patients2 %>% group_by(D) %>% summarise(diff = mean(Y)) %>% pull(diff) %>% diff()

diff_means ate

```
---

## 4. Breaking Down the Bias Components

(Hint: decompose $\mathbb{E}[Y_{i1} \mid D_i = 1] - \mathbb{E}[Y_{i0} \mid D_i = 0]$.)

```{r bias-decomposition}
# Your Code Here
```

Random Assignment as a Remedy

Randomization aims to break the link between who is treated and which potential outcomes they have, mitigating selection bias and enabling a simpler, more reliable difference-in-means estimation of the treatment effect.

1. Random Assignment with 10 Patients (Multiple Reps)

"'{r random-assignment-small} set.seed(123) dim_est <- vector()

for(i in 1:1000){ dim_est <- append(dim_est, po_patients %>% mutate(D = rbinom(10, 1, .5), Y = y_i1 * D + y_i0 * (1 - D)) %>% group_by(D) %>% summarise(dim = mean(Y)) %>% pull(dim) %>% diff()) }

Average difference-in-means

mean(dim_est)
```

Distribution of the estimates

ggplot(tibble(d = dim_est), aes(x = d)) + geom_histogram(binwidth = 0.5, fill="skyblue", color="white") + geom_vline(xintercept = ate, linetype = 'dashed') + labs(x = 'Difference in Means Estimate', y = 'Count')

95% quantiles

quantile(dim_est, c(.025, .975))

```
With only 10 patients, random chance can still produce wide variance in the difference-in-me

---

## 2. Larger Sample Random Assignment Example (1000 Patients)

$$
Y_{i1} \sim N(0.6, 0.5^2) \quad \text{(Treatment)} \\
Y_{i0} \sim N(0, 0.5^2) \quad \text{(Control)}
$$
```{r random-assignment-large}
# Your Code Here
```

Key Takeaways

1. **ATE vs. ATT**: Including individuals who are not actually eligible or never choose treatment (e.g., doctors in our banker example) can produce a misleading **ATE**. Often we care more about the **ATT**—the treatment effect for the subgroup that *would* receive treatment.

2. **Potential Outcomes**: Although we used artificial data to reveal both outcomes per individual, in reality, each person's counterfactual remains unknown. Causal inference methods are designed to cope with this partial-information challenge.

3. **Bias**:
   - **Selection bias** arises when the likelihood of being treated depends on potential outcomes.

   - **Heterogeneous treatment effects** mean that an "average" effect may mask large differences among subgroups.
```

4. **Random Assignment**: Underpins credible causal inference because it decouples treatment assignment from individual characteristics, enabling unbiased estimates of the ATE—particularly with sufficiently large samples.