

Causal Treatment Effects

Randomized Control Trial

The Difference-In-Means Estimator

Assumptions

Unbiasedness

Numerical Example

Randomized Controlled Trial and Average Treatment Effect Code ▾

Ruoqing Zhu

Last Updated: November 18, 2024

Average Treatment Effect

In many social and medical studies, we are interested in the effect of a treatment or intervention on an outcome. For example, we may want to know

- What is the effect of a new drug on the survival rate of patients?
- What is the effect of a new teaching method on students' test scores?
- What is the effect of college education on the income of individuals?

In a naive analysis, some researchers would calculate the effect in the following way, say, in the college degree problem:

- Collect data of incomes and education status by randomly sample individuals from a society
- Split the data into two groups: one with a college degree and one without
- Compute the averaged income from each group and compare them

Although this analysis is simple, it is flawed, or in statistical terms, biased. The reason is that the two groups may not be comparable. For example, individuals with a higher talent or motivation may be more likely to obtain a college degree, and these individuals may also have higher income potential. Therefore, the difference in income between the two groups may not be due to the college degree itself, but rather due to the difference in income potential. The bias in our proposed estimation procedure is caused by ignoring the confounding effect of this income potential. This is also known as the confounding bias.

There are, however, situations where the naive analysis is valid. For example, if the treatment is randomly assigned to individuals. then the two groups are comparable, and the naive analysis is valid. However, this random assignment is not always possible or ethical. For example, it is not ethical to randomly assign individuals to smoking and non-smoking groups to study the effect of smoking on lung cancer. In this lecture and the following several lectures, we will discuss concepts and methods associated with the estimation of treatment effects. Moreover, we will introduce methods that can be used to suggest a better treatment for individuals based on their characteristics, which is known as personalized medicine.

Causal Treatment Effects

To estimate the best treatment for an individual, one fundamental concept is the treatment effect. Let's define some notation. Suppose we have a binary treatment label $A \in \{0, 1\}$, for which 1 indicates a type of treatment and 0 indicates another treatment or simply the placebo control. We also have a response variable Y , which is the outcome of interest. For an individual i , this could potentially give us two outcomes: $Y_i(0)$ and $Y_i(1)$, and the difference between the two is the **causal treatment effect**¹:

$$\Delta_i = Y_i(1) - Y_i(0)$$

The fundamental difficulty in estimating the treatment effect is that we can only observe one of the two outcomes. For example, if you received offers from two universities, UIUC and Harvard, and you can only choose one of them. If you choose UIUC and observe the outcome of your salary after graduation, then you would never observe the outcome of going to Harvard. The two are essentially in two parallel universes. This dilemma between the **actual** (or realized) and **counterfactual** outcomes is known as the **fundamental problem of causal inference** by Holland (1986). However, some statistical framework would allow us to estimate the treatment effect, or its related quantities. The most commonly accepted view is the **potential outcome** framework, which originated in Neyman (1923) who considered a randomized experiment, and Rubin (1974) who considered observational (non-randomized) studies. In this section, let us explore this framework using the randomized control trial and discuss some of its crucial assumptions.

Randomized Control Trial

Let's consider the **average treatment effect** (ATE), which is the treatment effect averaged over the entire population:

$$\tau = E[\Delta_i] = E[Y_i(1)] - E[Y_i(0)]$$

To entertain the idea of potential outcomes, let's suppose that we could actually walk both the parallel universe. Then we would observe both $Y_i(1)$ and $Y_i(0)$. A nature estimator would be

$$\frac{1}{n} \sum_{i=1}^n \Delta_i = \frac{1}{n} \sum_{i=1}^n (Y_i(1) - Y_i(0)) \text{ (}\#eq : sate\text{)}$$

This quantity is called a **sample average treatment effect** (SATE), a hypothetical unbiased estimator for τ . We can easily see that it is an unbiased estimator of the ATE. However, this estimator is not possible in reality.

In reality, we could randomly assign the treatment to a group of individuals and compare the outcomes between the two groups. This is known as the **randomized control trial** (RCT). In this case, we only observe one of the two potential outcomes for each individual:

$$\begin{aligned} Y_i &= A_i Y_i(1) + (1 - A_i) Y_i(0) \\ &= \begin{cases} Y_i(1) & \text{if } A_i = 1 \\ Y_i(0) & \text{if } A_i = 0 \end{cases} \end{aligned}$$

where A_i is the treatment assignment for individual i .

The Difference-In-Means Estimator

A naive (but pretty good) idea is to estimate the ATE using the mean differences of the two groups, which is called the **difference-in-means estimator**:

$$\begin{aligned}\hat{\tau} &= \frac{1}{n_1} \sum_{A_i=1} Y_i - \frac{1}{n_0} \sum_{A_i=0} Y_i \\ &= \frac{1}{n_1} \sum_{i=1}^n A_i Y_i - \frac{1}{n_0} \sum_{i=1}^n (1 - A_i) Y_i\end{aligned}$$

where n_1 and n_0 are the sample sizes of the two groups. The advantage of using RCT, as we will see later, is that this estimator is **unbiased even if there are other covariates** that affect the outcome and we will discuss the utilization of other covariates later.

Assumptions

The unbiasedness of this estimator can be established by setting a connection with the SATE estimator @ref(eq:sate). Besides assuming an i.i.d. (independent and identically distributed) set of samples, we need two important assumptions:

Independence: $A_i \perp \{Y_i(0), Y_i(1)\}$

This is a pretty natural assumption, which suggests that the assignment of treatment has nothing to do with all the possible potential outcomes. What situation would violate this assumption? For example, a patient may choose what he/she believes to be the better treatment. In this case, the treatment assignment is not independent of the potential outcomes.

SUTVA (Stable Unit Treatment Values Assumption): $Y_i = Y_i(A_i)$

This assumption is a bit subtle. It has two components.

- **No interference**: the outcome of individual i is not affected by other individuals.
- **Consistency**: There are no hidden forms of treatment, i.e., the outcome of individual i is the same as the outcome of individual i under the same treatment assignment.

The first part is relatively easy to understand. For example, if you are a patient in a clinical trial, your outcome should not be affected by the treatment of other patients. However, if all patients share and are competing on the same medical resources, then the outcome of one patient may be affected by the treatment of others. The second part is somewhat philosophical and it can be closely related to how we define the treatment. One could simply say that this is automatically satisfied. But if the patient does not take the treatment as prescribed, then the outcome may be different. In that case, there is an entire area of research on **noncompliance** and **intention-to-treat** (Gupta 2011) analysis.

Unbiasedness

As we discussed previously, we could consider the difference-in-means estimator. In fact, under the assumptions we gave previously, this estimator is unbiased for the SATE, and hence, the ATE. To see this, let's consider the conditional expectation of the estimator given the potential outcomes and the treatment assignment ²:

$$\begin{aligned}
& E \left[\frac{1}{n_1} \sum_{i=1}^n A_i Y_i \mid \{Y_i(0), Y_i(1)\}_{i=1}^n, n_1 \right] \\
&= E \left[\frac{1}{n_1} A_i Y_i(1) \mid \{Y_i(0), Y_i(1)\}_{i=1}^n, n_1 \right] \quad \text{by SUTVA} \\
&= \frac{1}{n_1} \sum_{i=1}^n Y_i(1) E \left[A_i \mid \{Y_i(0), Y_i(1)\}_{i=1}^n, n_1 \right] \quad Y_i(1)'s \text{ are constant} \\
&= \frac{1}{n_1} \sum_{i=1}^n Y_i(1) \frac{n_1}{n} \quad \text{by Independence} \\
&= \frac{1}{n} \sum_{i=1}^n Y_i(1)
\end{aligned}$$

Here, we made an assumption that the treatment assignment is random, i.e., $E[A_i] = \Pr(A_i = 1) = n_1/n$. The same argument can be applied to the second term. Therefore, the difference-in-means estimator is unbiased for SATE, and so is for ATE.

Numerical Example

Let's use some simulation study to illustrate the properties of the difference-in-means estimator. Suppose we have $n = 200$ patients, and we randomly assign $n_1 = 100$ patients to the treatment group and $n_0 = 100$ patients to the control group. Suppose our outcomes are generated from a linear model

$$E(Y|X = x) = 0.5 \times x + A \times x^2. \text{ (\#eq : model)}$$

In this case, the potential outcomes for a subject with covariate value x are $0.5 \times x + x^2$ for treatment 1 and $0.5 \times x$ for treatment 0. If X follows a standard normal distribution, the expected treatment effect is $\tau = E(X^2) = 1$. In a regression problem, we typically want to observe X and model the relationship between X and Y and then infer the treatment effect. However, in RCT, this is not necessary. Let's generate the observed outcomes and estimate the ATE using the difference-in-means estimator without using X . We repeat this process for 200 times and point estimations and their confidence intervals.

Hide

```

# setting parameters
n <- 200
n1 <- 100
n0 <- 100
tau <- 1 # true ATE
nsim = 500

set.seed(1)

tauhat <- rep(NA, nsim)
tauhatsd <- rep(NA, nsim)

for (i in 1:nsim) {
  # generate potential outcomes
  X <- rnorm(n)
  Y1 <- rnorm(n, mean = 0.5*X+X^2)
  Y0 <- rnorm(n, mean = 0.5*X)

  # treatment label
  A = sample(c(rep(1, n1), rep(0, n0)))

  # observed outcomes
  Y <- A * Y1 + (1 - A) * Y0

  tauhat[i] <- mean(Y[A == 1]) - mean(Y[A == 0])
  tauhatsd[i] <- sqrt(var(Y[A == 1]) / n1 + var(Y[A == 0]) / n0)
}

# make two plots on the same row
par(mfrow = c(1, 2))

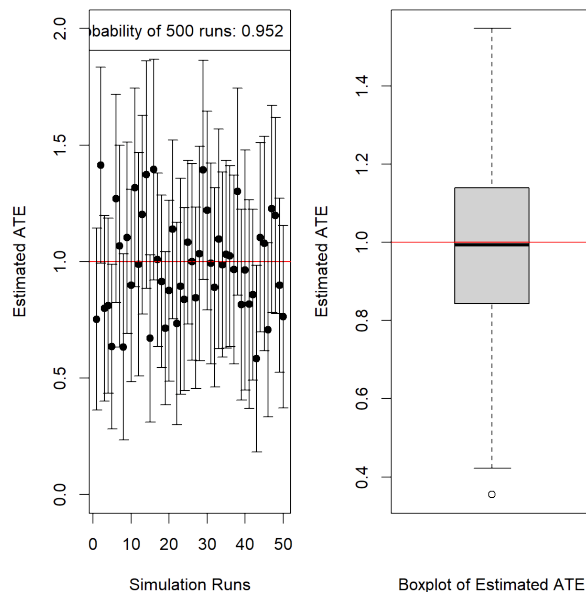
# set margin of figure
par(mar = c(4, 4, 1, 1))
plot(tauhat[1:50], pch = 19, ylim = c(0, 2),
      xlab = "Simulation Runs", ylab = "Estimated ATE")
abline(h = tau, col = "red")

# adding confidence intervals
for (i in 1:50) {
  ci_lower <- tauhat[i] - 1.96 * tauhatsd[i]
  ci_upper <- tauhat[i] + 1.96 * tauhatsd[i]
  arrows(x0 = i, y0 = ci_lower, x1 = i, y1 = ci_upper, angle = 90, code = 3, leng
th = 0.05)
}

coverage = sum((tauhat - 1.96 * tauhatsd < tau) & (tauhat + 1.96 * tauhatsd > ta
u)) / nsim
legend("topright", paste("Coverage probability of", nsim, "runs:", coverage))

boxplot(tauhat, xlab = "Boxplot of Estimated ATE", ylab = "Estimated ATE")
abline(h = tau, col = "red")

```



The result roughly demonstrates that the point estimation is unbiased and the confidence interval has the correct coverage probability. We should note that in this estimation, we **did not use the covariate information**. The unbiasedness comes purely from the randomization of the treatment assignment.

Gupta, Sandeep K. 2011. "Intention-to-Treat Concept: A Review." *Perspectives in Clinical Research* 2 (3): 109.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60.

Neyman, Jerzy. 1923. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles." *Ann. Agricultural Sciences*, 1–51.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688.

1. In this notation, we usually consider each Δ_i as a random variable. They may or may not have the same mean value across different individuals (e.g., if they depend on additional covariates), but that will not change the main result of this section. In our later lectures, we could consider using additional covariates X to improve our estimation if they are available. ↩
2. We condition on the potential outcomes because it avoids making assumptions on how they are generated. And this is also the reason why we need to assume SUTVA and independence. ↩