Randomised Controlled Trials

Methods and Applications of Machine Learning

Elisa Tosetti

2023-2024

Outline

- Causal effects
- Introduction to the potential outcome notation
- Introduction to randomised controlled trials
- Average treatment effects vs conditional average treatment effects
- The propensity score

References: *Angrist and Pischke* book

Causal relashionship

- A **causal relationship** refers to the cause-effect connection between two variables: changes in one variable may cause changes in the other
- **Treatment variable**: the variable where the change originates
- **Outcome variable**: the variable that may change in response to a change in the treatment variable
- In mathematical notation we represent the treatment variable as D and the outcome variable as Y and represent the causal relationship between them as

$$D \rightarrow Y$$

- We initially focus on treatment variables that are binary

Introduction to the potential outcome notation

Imagine we want to know the causal effect of hospitalisation on the health of people

Research question: Within the population of elderly individuals, does visiting the emergency room for primary care improve health?

Two questions from the National Health Interview Survey (NHIS) in the US:

- During the past 12 months, was the respondent a patient in a hospital?
- Would you say your health in general is excellent (1), very good (2), good (3), fair (4), poor (5)?

$$\textit{Hospitalisation} \rightarrow \textit{Health}$$

Treatment variable and outcome variable

$$D_i = \begin{cases} 1 \text{ if individual i receives the treatment} \\ 0 \text{ otherwise} \end{cases}$$

Based on whether the individual receives the treatment, we speak of two different conditions:

- **Treatment** is the condition with the treatment ($D_i = 1$) (e.g., hospitalised)
- **Control** is the condition without the treatment ($D_i = 0$) (e.g., not hospitalised)

Outcome variable for individual i (e.g., health score) is y_i

Introduction to the potential outcomes framework

- When estimating the causal effect of D on Y, we attempt to quantify the change in the outcome variable Y, ΔY, that is caused by a change in the treatment variable D
- To measure this change in the outcome Y, ideally we would compare two potential outcomes, the outcome when the treatment is present and the outcome when the treatment is not present
- This leads us to the framework known as the **Neyman-Rubin potential outcomes** framework

Introduction to the potential outcomes framework

For each individual there are two potential outcomes:

$$\text{Potential outcome} = \begin{cases} y_{1i} \text{ if } D_i = 1 \\ y_{0i} \text{ if } D_i = 0 \end{cases}$$

If, for each individual i, we could observe both potential outcomes, then the difference between these two gives us the **causal effect** of D_i (e.g., hospital treatment) on y_i (e.g., health of individuals):

$$\tau_i = y_{1i} - y_{0i}$$

This is known as the **individual causal effect** of a treatment on an outcome

Introduction to the potential outcomes framework

i	health(hospitalised=1)	health(hospitalised=0)	Δhealth
1	5	1	3
2	3	2	1
3	1	1	0
4	4	5	-1
5	5	1	4
6	5	4	1

Unfortunately, this kind of analysis is not possible as we *never* observe both potential outcomes for the same individual

Introduction to the potential outcomes framework

This leads us to the fundamental problem of causal inference:

- We can never simultaneously observe y_{1i} and y_{0i}: **for each individual** both outcomes are possible but they are **never observed at the same time** (a unit either receives the treatment or not, never both)
- To get around the fundamental problem of causal inference we must find good approximations for the counterfactual outcomes
- Most of applied econometrics focuses on addressing this problem

Note: we can think of causal inference as a prediction problem. How could we predict the counterfactual given that we never observe it?

Average treatment effects

- If a unit receives the treatment the observed outcome is y_{1i} and **counterfactual outcome** is y_{0i}
- We must find good approximations for the counterfactual outcomes
- To accomplish this, we move away from individual-level effects, and focus on the *average* causal effect across a group of individuals
- Average causal effect of D on Y: average change in Y caused by a change in X for a group of individuals

$$\tau = E(y_{1i}) - E(y_{0i})$$

The switching equation

We can write the observed outcome y_i for each individual i in terms of the potential outcomes using the so-called **switching equation**:

$$y_i = \begin{cases} y_{1i} \text{ if } D_i = 1 \\ y_{0i} \text{ if } D_i = 0 \end{cases}$$

$$= y_{0i} + (y_{1i} - y_{0i})D_i$$

where $y_{1i} - y_{0i}$ is the causal effect of D_i for an individual

Average treatment effects

Consider $E(y_i|D_i = 1) - E(y_i|D_i = 0)$. From the switching equation we can rewrite it as follows:

$$\underbrace{E(y_i|D_i = 1) - E(y_i|D_i = 0)}_{\text{observed difference in average outcome}}$$

$$= \underbrace{E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 1)}_{\text{(1): average treatment effect on the treated (ATT)}} +$$

$$+ \underbrace{E(y_{0i}|D_i = 1) - E(y_{0i}|D_i = 0)}_{\text{(2): selection bias}}$$

Average treatment effects

- (1): is the **ATT**: captures the average difference between the outcome of the treated, $E(y_{1i}|D_i = 1)$, and what would have happened to them had they not been treated $E(y_{0i}|D_i = 1)$:

$$E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 1) = E(y_{1i} - y_{0i}|D_i = 1) = E(\tau_i|D_i)$$

- (2): is the **selection bias**: it is the difference in average y_{0i} between those who were and were not treated. The extent to which the "control group" provides a bad counterfactual for the treated individuals

In our example

(1): **ATT**: average difference between the health of the hospitalised and what would have happened to them had they not been hospitalised. This is the average causal effect of hospitalisation on those who were hospitalised

(2): **selection bias**: Average y_{0i} between those who were and were not hospitalised. Because the sick are more likely than the healthy to seek treatment, those who were hospitalised are likely to have worse y_{0i} creating selection bias in this example

The selection bias is given by how the treated and control group differ before the treatment, in case neither of them has received the treatment

The selection bias may be so large (in absolute value) that it completely masks a positive treatment effect

Selection bias

Angrist and Pischke (MHE, p. 12):

"The goal of most empirical economic research is to overcome selection bias, and therefore to say something about the causal effect of a variable like D"

How do experiments - the gold standard of empirical economic (and scientific) research - accomplish this goal and overcome selection bias?

If $E(y_{0i}|D_i = 1) = E(y_{0i}|D_i = 0)$ then, association is causation

To say $E(y_{0i}|D_i = 1) = E(y_{0i}|D_i = 0)$ is to say that treatment and control group are comparable before the treatment. Or, when the treated had not been treated, if we could observe its y_{0i}, its outcome would be the same as the untreated

Randomised experiments: the golden standard

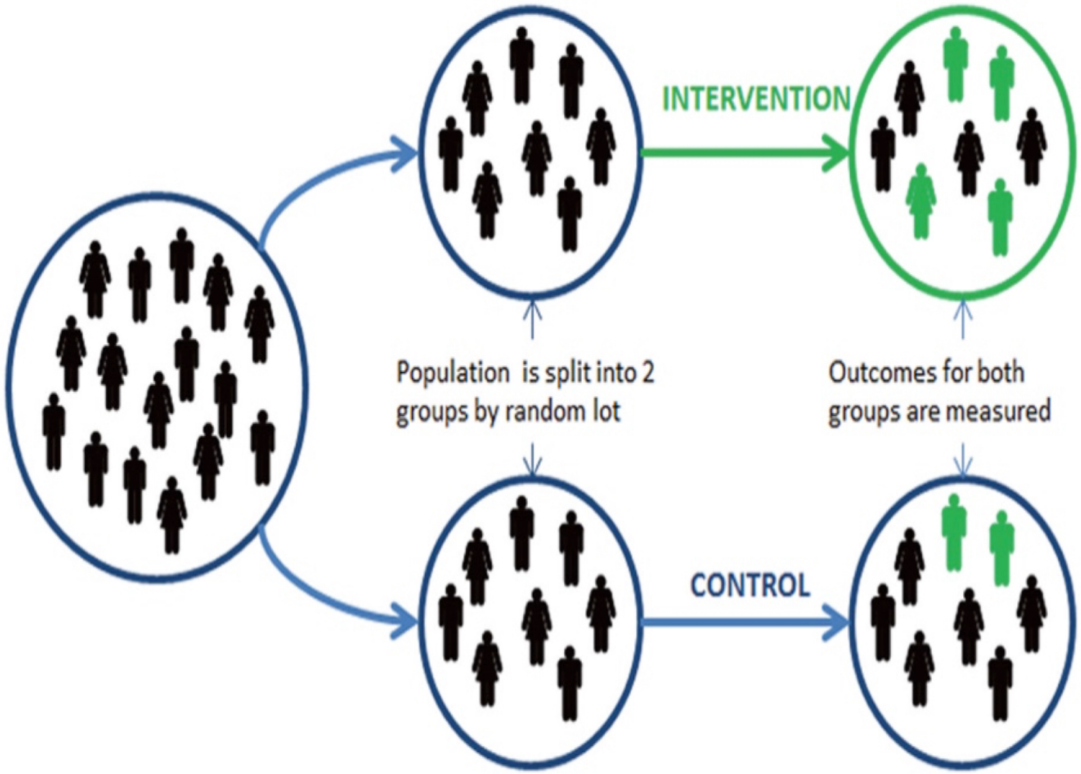
In a randomised experiment (also known as **randomised controlled trial**), researchers decide who receives the treatment based on a *random process*

In our health experiment: hospitalisation on a randomly chosen patient

When treatment assignment is randomised, the only thing that distinguishes the treatment group from the control group, besides the reception of the treatment, is chance:

$\Rightarrow$ This means that the two groups are comparable to each other, on average, in all respects other than whether or not they received the treatment

Random treatment assigment



Random treatment assigment

Random treatment assignment makes the treatment and control groups on average identical to each other in all observed and unobserved pre-treatment characteristics

Pre-treatment characteristics are the characteristics of the individuals in a study before the treatment is administered

For example, the average age of individuals going to hospital and of individuals *not* going to hospital

Randomised experiments

- Random assignment of D_i solves the selection problem because random assignment makes D_i independent of which outcome we observe
- Assume $\{y_{1i}, y_{0i}\} \perp D_i$. Then:

$$E(y_i|D_i = 1) - E(y_i|D_i = 0) = E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 0)$$
$$= E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 1)$$

where the independence of y_{0i} and D_i allows us to swap $E(y_{0i}|D_i = 0)$ for $E(y_{0i}|D_i = 1)$

Randomised experiments

Given random assignment, this simplifies further to

$$E(y_{1i}|D_i = 1) - E(y_{0i}|D_i = 1) = E(y_{1i} - y_{0i}|D_i = 1)$$

$$= E(y_{1i} - y_{0i})$$

The effect of randomly-assigned hospitalisation on the hospitalised is the same as the effect of hospitalisation on a randomly chosen patient. The main thing, however, is that random assignment of D_i eliminates selection bias

Randomly assigned treatment

Random assignment of treatment gives us the **average treatment effect** (ATE):

$$E(y_{1i}) - E(y_{0i})$$

With random assignment the observed means difference between the two groups is an unbiased estimator of the average treatment effect

In other words, there is no selection bias

Thus, through randomisation, although we never observe $y_{1i} - y_{0i}$, we can consistently estimate it using the ATE $E(y_{1i} - E(y_{0i}))$

Difference-in-means estimator

The **difference-in-means estimator** of the average treatment effect is the sample average of outcomes in treatment minus the sample average of outcomes in control:

$$\hat{\tau} = \frac{1}{n_1} \sum_{i:D_i=1} y_i - \frac{1}{n_0} \sum_{i:D_i=0} y_i$$

By using random treatment assignment, we can assume that the treatment and control groups were comparable before the administration of the treatment

As a result, we can rely on the difference-in-means estimator to provide a valid estimate of the average treatment effect

Difference-in-means estimator

We can also estimate the ATE of a binary treatment via a linear regression of observed outcomes. From expression in slide 11:

$$y_i = \underbrace{\beta_0}_{\| \; E(y_{0i})} + \underbrace{\tau}_{\| \; y_{1i}-y_{0i}} D_i + \underbrace{u_i}_{\| \; y_{0i}-E(y_{10})}$$

Evaluating the conditional expectation of this equation with treatment status switched off and on gives:

$$E(y_i|D_i = 1) = \alpha + \tau + E(u_i|D_i = 1)$$

$$E(y_i|D_i = 0) = \alpha + E(u_i|D_i = 0)$$

so that:

$$E(y_i|D_i = 1) - E(y_i|D_i = 0) = \tau + E(u_i|D_i = 1) - E(u_i|D_i = 0)$$

where $E(u_i|D_i = 1) - E(u_i|D_i = 0)$ is the selection bias

Difference-in-means estimator in experiments

Interestingly, selection bias amounts to the correlation between the error term, u_i, and the regressor, D_i

Selection bias should remind you a lot of **omitted-variable bias**: there is something in the error term that is affecting y_i and is also correlated with D_i. In other words, our treatment D_i is **endogenous**

In RCTs, since D_i is randomly assigned, the selection term disappears, and the simple OLS estimator of τ in the above model is a consistent estimate of the causal effect

Why study experiments?

- RCTs provide a conceptual benchmark for assessing observational studies
- RCTs are often considered the **Gold Standard** of empirical research: random assignment allows identification of causal effects
- By design D_i independent of any pre-treatment influences on y_i allowing **causal estimation** of the effect of D_i on y_i
- RCTs give researchers the ability to let questions determine the data to be obtained, instead of the data determining the questions that can be asked
- Thinking about experiments helps us to understand quasi-experiments, or *natural experiments* in which there some variation is "as if" randomly assigned.

Why not

Unfortunately, we are not always able to conduct an experiment. Three types of obstacles:

- **Ethical**: It would not be ethical to randomize certain treatments, such as a potentially lethal drug
- **Logistical**: Some treatments, such as race, cannot be easilty manipulated
- **Financial**: Experiments are often expensive

Threats to Validity of RCTs

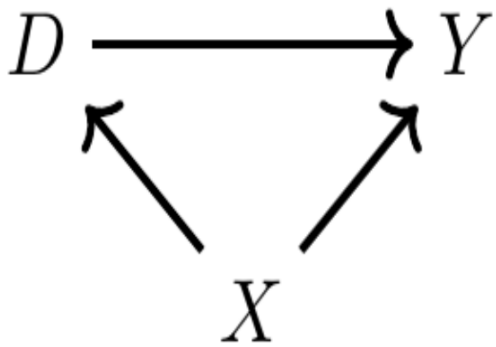
- Failure to randomize
- Non compliance
- Attrition
- Small sample sizes

Confounding variables

When condition $\{y_{1i}, y_{0i}\} \perp D_i$ is violated, we say that we are in an **observational setting**

A confounding variable, also known as **confounder**, is a variable that affects both (i) The likelihood of receiving the treatment (ii) The outcome Y

A confounder obscures the causal relationship between D and Y



In the presence of confounding variables, treatment and control groups are not comparable, correlation does not necessarily imply causation and the difference-in-means estimator does not provide a valid estimate of the average treatment effect

Propensity score models

In social and health research, RCTs are not always practical, ethical, or even desirable

Under such conditions, researchers often use quasi-experimental designs, which - in most instances - are vulnerable to omitted variable bias

Propensity score models help to remove the bias though:

- Reweighting
- Stratification and matching

Some of the methods that we will study in this course will use propensity scores

Propensity scores

Given a treatment D and a set of covariates X, the **propensity score** is:

$$e(x) = P(D_i = 1|X = x)$$

In essence, it is the probability that an individual with observed covariate x is in the treatment group

In randomised control trials, we exactly know $e(x)$ (they are constant)

In observational studies, we do not know $e(x)$, although we can estimate it from data

The variability of the propensity score gives a measure of how far we are from a randomised trial

Propensity scores

One key assumption on propensity scores is that there exists a constant k such that:

$$0 < e(x) < k \text{ for all x}$$

This assumption is known as **overlap**

it means that for all types of individuals in our population (i.e., all values of observable characteristics) we can find some portion of individuals in treatment and some in control

Intuitively, this is necessary because we'd like to be comparing treatment and control at each level of the covariates and then aggregate those results

Propensity scores

One intuitive use of the propensity scores is use them as weights to create a balanced sample of treated and control observations

It is possible to show that

$$E(y_{1i} - y_{0i}) = E\left[\frac{D_i y_i}{e(x_i)} - \frac{(1 - D_i)y_i}{1 - e(x_i)}\right]$$

This implies that the following inverse-propensity weighted estimator is unbiased for the average treatment effect:

$$\frac{1}{n}\sum\left[\frac{D_i y_i}{\hat{e}(x_i)} - \frac{(1 - D_i)y_i}{1 - \hat{e}(x_i)}\right]$$

see Angrist and Pischke p.60