Causal inference and counterfactuals

Silje Synnøve Lyder Hermansen

2020-12-01

Distinctions

1st part of the course

The goal is to provide a parsimonious but correct description of the observed data (y).

- model choice based on *outcomes*: a discrete y
- interpretation and model assessment: focus on prediction

2nd part of the course

The goal is to test a causal claim (x) by leveraging the data strategically.

- model choice based on a *treatment*: a discrete x
- interpretation and model assessment: focus on research design/case selection using counterfactuals and marginal effects

Causal claim

Causal claims abound in political science.

- Does X cause Y?
- If so, how big is the effect?

$\Rightarrow$ *This part of the course surveys strategies to credibly make such claims*

Causal empiricist critic

Causal empiricist critic

Causal empiricist critic of traditional regression approaches (Samii, 2016)

Pseudo-generalizability: external reliability

Trad. approach: Regression study on data representing the universe of cases and therefore generalizes

Critique: this neglects the identifying variation

- assumes *positivity*; if not, your results will be *model dependent*
- regression weights mean giving influence to observations where the covariates poorly explain the treatment

$\Rightarrow$ *Large-N studies do not necessary draw variation (information) from a large N; sometimes even the wrong N*

Pseudo-facts: internal reliability

Trad approach: assumes that

- covariates (selected on theory) account for confounding factors
- the functional form is correct

Critique:

- adressing mis-spesification with (formal) theory merely shows that it is possible to come up with a theory for any correlation
- lack of reflection of causal relation leads to control away the effect of teoretical interest (controlling for "post-treatment")

Pillars of causal empiricism

Focus on research design

What is the identifying variation? How are treatment values determined?

- Randomness rules out unobservable confounders of the treatment: from design (experiment) natural experiment, discontinuity...
- Substantial knowledge

$\Rightarrow$ *The problem is that true randomness is a fairly rare event*

Specific causal facts: "go local"

Identify the population you want to generalize to, and be narrow

- focus on one sub-population
- focus on one causal factor (not the other xs)

$\Rightarrow$ *Test narrow theoretical claims, leave theory development to others. This is a research project, not a paper.*

The counterfactual approach

The counterfactual approach

Morgan and Winship (2015, mainly ch 2)

Why a "counterfactual" framework?

The causal effect is the difference between an observed and a hypothetical outcome ($Y^1 - Y^0$)

Table 2.1 The Fundamental Problem of Causal Inference

Group	Y^1	Y^0
Treatment group ($D = 1$)	Observable as Y	Counterfactual
Control group ($D = 0$)	Counterfactual	Observable as Y

Some notations

Uppercase for population-level variables, lowercase for individuals

- an intervention: treatment (D=1) or control (D=0)
- two outcome states: if treatment was administered (Y^1) or not (Y^0)

$\Rightarrow$ collapses into $Y = DY^1 - (1 - D)Y^0$

We have thus four states:

- $Y^1|D=1$: treatment was administered and treatment-outcome observed
- $Y^0|D=1$: treatment was administered and control-outcome not observed
- $Y^1|D=0$: treatment was not administered and treatment-outcome not observed
- $Y^0|D=0$: treatment was not administered and control-outcome observed

Mentimeter: Where are the counterfactuals?

Why a "counterfactual" framework?

We would like to make an individual-level causal claim, but we cannot observe to versions of the same individual

Table 2.1 The Fundamental Problem of Causal Inference

Group	Y^1	Y^0
Treatment group ($D=1$)	Observable as Y	Counterfactual
Control group ($D=0$)	Counterfactual	Observable as Y

$\Rightarrow$ *We'll look for real-life examples that approximate the counterfactual*

Claims at the group-level

We study larger groups and make claims at the group-level

individual-level causal effects $\rightarrow$ average causal effect

- Average treatment effect (ATE): $E[\delta] = E[Y^1] - E[Y^0]$

$\Rightarrow$ *How well does this describe the effect?*

Base-line assumptions

Requirement 1 for the counterfactual framework

We need to have clearly identifiable causal states. This implies

- discrete x: binary ("treatment"/"control") or categorical ("alternative states")
- assumptions about causal mechanisms in x: categories can be more or less broad, founded in theory etc.
- assumptions about the counterfactual: how realistic is it to observe the control state in our current world?

Requirement 2 for the counterfactual framework

Stable unit treatment value assumption (STUVA) The effect of the treatment for one individual should not be dependent on the assignment of another individual to that treatment.

- influence from contact among individuals once assigned
- influence due to the prevalence of the treatment

$\Rightarrow$ *be modest about your claims*

Requirement 3

Treatment status has to be independent of the potential outcomes (but not vice-verca)

$(Y^1, Y^0) \perp\!\!\!\perp D$

$\Rightarrow$ *treatment assignment is ignorable*

Conditional average effects

Example: ATE $= Y^1 - Y^0$

Table 2.3 An Example of Inconsistency and Bias of the Naive Estimator When the ATE Is the Causal Effect of Interest

Group	$E[Y^1 \mid D]$	$E[Y^0 \mid D]$
Treatment group ($D=1$)	10	6
Control group ($D=0$)	8	5

- A group passes a test (Y), and here is the score
- 30% of them has a bachelor's degree ($D = 1$)

Mentimeter: What is the ATE?

Example: ATE $= Y^1 - Y^0$

- Average Treatment effect among the Treated (ATT): 4
- Average Treatment effect among the Controlled (ATC): 3
- Average Treatment effect (ATE): 3.3

Example: ATE $= Y^1 - Y^0$

There are two sources of bias in the naïve estimation

- Base-line bias in expected outcome among the two groups: $(E[Y^0|D=1] - E[Y^0|D=0])$
- Differential treatment bias: Difference in treatment effect across groups.

A counterfactual framework

Two assumptions for approximating the counterfactual

Table 2.1 The Fundamental Problem of Causal Inference

Group	Y^1	Y^0
Treatment group ($D = 1$)	Observable as Y	Counterfactual
Control group ($D = 0$)	Counterfactual	Observable as Y

$\Rightarrow$ *Both or either can be true; this determines which groups we can generalize to.*

Two assumptions

If both assumptions are true, then $ATE = ATT = ATC$

1. $E[Y^1|D=1] = E[Y^1|D=0]$
2. $E[Y^0|D=1] = E[Y^0|D=0]$

$\Rightarrow E[\delta] = E[Y^1] - E[Y^0]$

We can compare the average between the treated and the untreated and generalize to both groups.

Assumption 1 is true

We can still make claims about the control group (ATC)

1. $E[Y^1|D=1] = E[Y^1|D=0]$ Same response to treatment for all
2. $E[Y^0|D=1] \neq E[Y^0|D=0]$ Base-line bias in control group

$\Rightarrow E[Y^1|D=0] - E[Y^0|D=0]$

Ex: Put normally catholic school student in public school.

Assumption 2 is true

we can estimate the ATT

1. $E[Y^1|D=1] \neq E[Y^1|D=0]$ Different response to treatment
2. $E[Y^0|D=1] = E[Y^0|D=0]$ Base-line bias in control group

$\Rightarrow E[Y^1|D=1] - E[Y^0|D=1]$

Ex: Effect of catholic school on normally catholic school students.