Causality, potential outcomes and randomization

Matti Sarvimäki

Principles of Empirical Analysis
Lecture 4

Course outline and learning objectives

- Data and measurement
  1. introduction, data
  2. descriptive statistics
  3. more descriptive statistics
- Experimental methods
  1. **today: causality and research designs**
  2. statistical significance
  3. statistical power
  4. noncompliance
- Quasi-experimental methods
  1. observational data and quasi-experiments
  2. difference-in-difference (DiD)
  3. regression discontinuity design (RDD)

Course outline and learning objectives

- Data and measurement
  1. introduction, data
  2. descriptive statistics
  3. more descriptive statistics
- Experimental methods
  1. **today: causality and research designs**
  2. statistical significance
  3. statistical power
  4. noncompliance
- Quasi-experimental methods
  1. observational data and quasi-experiments
  2. difference-in-difference (DiD)
  3. regression discontinuity design (RDD)

- Todays' learning objectives:
  - Good understanding of what is
    1. causality
    2. counterfactual
    3. potential outcomes
    4. treatment effect
    5. selection bias
  - Good understanding of why randomization eliminates selection bias
  - Basic understanding of the ethics and limitations of RCTs

Course outline and learning objectives

- Data and measurement
  1. introduction, data
  2. descriptive statistics
  3. more descriptive statistics
- Experimental methods
  1. **today: causality and research designs**
  2. statistical significance
  3. statistical power
  4. noncompliance
- Quasi-experimental methods
  1. observational data and quasi-experiments
  2. difference-in-difference (DiD)
  3. regression discontinuity design (RDD)

- Todays' learning objectives:
  - Good understanding of what is
    1. causality
    2. counterfactual
    3. potential outcomes
    4. treatment effect
    5. selection bias
  - Good understanding of why randomization eliminates selection bias
  - Basic understanding of the ethics and limitations of RCTs
- Also: the first feedback survey is out

Causal questions

- Thus far, we have focused on **descriptive** questions (lectures 1–3)
  - aim: measure the actual state of the world
  - "what is joint distribution of X and Y?"

Causal questions

- Thus far, we have focused on **descriptive** questions (lectures 1–3)
  - aim: measure the actual state of the world
  - "what is joint distribution of X and Y?"
- We would often need to evaluate the **impact** of X on Y, e.g.
  - education on earnings
  - marketing campaing on sales
  - carbon tax on emissions
  - R&D subsidy on innovation
  - fiscal stimulus on unemployment

Causal questions

- Thus far, we have focused on **descriptive** questions (lectures 1–3)
  - aim: measure the actual state of the world
  - "what is joint distribution of X and Y?"
- We would often need to evaluate the **impact** of X on Y, e.g.
  - education on earnings
  - marketing campaing on sales
  - carbon tax on emissions
  - R&D subsidy on innovation
  - fiscal stimulus on unemployment
- These are **causal** questions
  - aim: compare *counterfactual* states of the world
  - "how would Y change if we changed X?"
    - we typically refer to Y as "outcome" and to X as "treatment"

Randomized experiments

- The next four lectures will focus on answering causal questions using research designs based on **randomization**
  - the simplest context for learning relevant statistical concepts

Randomized experiments

- The next four lectures will focus on answering causal questions using research designs based on **randomization**
  - the simplest context for learning relevant statistical concepts
- The prime example is **randomized controlled trials** (RCT)
  - RCTs have become an important part of economits' toolkit
  - you might end up running them for living
  - you will definitely end up interpretting results from other people's RCTs

Randomized experiments

- The next four lectures will focus on answering causal questions using research designs based on **randomization**
  - the simplest context for learning relevant statistical concepts
- The prime example is **randomized controlled trials** (RCT)
  - RCTs have become an important part of economits' toolkit
  - you might end up running them for living
  - you will definitely end up interpretting results from other people's RCTs
- Even when we can't run an experiment, it is often helpful to ask: what would be the **ideal experiment** for answering this question?
  - helpful benchmark for "naturally occurring" or "quasi" experiments
    - we'll discuss an example of a "natural experiment" involving actual randomization already on Wednesday's class
    - you'll see other types of quasi-experimental approaches in lectures 8–11

In-class discussion: Impact of a new integration program

- Imagine that you have been asked to assist the government to evaluate the following proposal by a private investor:
  - the investor has designed a new type of integration program for newly arrived immigrants (which seems reasonably good)
  - the government pays for performance, i.e., payment is a function of how well the participants perform in the labor market after participating

In-class discussion: Impact of a new integration program

- Imagine that you have been asked to assist the government to evaluate the following proposal by a private investor:
    - the investor has designed a new type of integration program for newly arrived immigrants (which seems reasonably good)
    - the government pays for performance, i.e., payment is a function of how well the participants perform in the labor market after participating
- What would be your advice?
    - no need to get to the "right" answer! the point is to collect some thoughts and then we'll start thinking through them systematically

In-class discussion: Impact of a new integration program

- Imagine that you have been asked to assist the government to evaluate the following proposal by a private investor:
  - the investor has designed a new type of integration program for newly arrived immigrants (which seems reasonably good)
  - the government pays for performance, i.e., payment is a function of how well the participants perform in the labor market after participating
- What would be your advice?
  - no need to get to the "right" answer! the point is to collect some thoughts and then we'll start thinking through them systematically
- My take: helpful to break this into two parts
  - what is the question one needs to answer?
  - how to answer it (ideal experiment)?

Elements of causal questions

1. **Treatment**
   - impact *of* [...]
2. **Counterfactual**
   - impact *in comparison to* [...]
3. **Outcome**
   - impact *on* [...]
4. **Population**
   - impact *for* [...]

Elements of causal questions

1. **Treatment**
   - impact *of* [...]
2. **Counterfactual**
   - impact *in comparison to* [...]
3. **Outcome**
   - impact *on* [...]
4. **Population**
   - impact *for* [...]

- What is a well-defined question for our case study?

Elements of causal questions

1. **Treatment**
   - impact *of* [...]
2. **Counterfactual**
   - impact *in comparison to* [...]
3. **Outcome**
   - impact *on* [...]
4. **Population**
   - impact *for* [...]

- What is a well-defined question for our case study?
  - my take: what is the impact of the **new program** in comparison to **business-as-usual programs** on **participants'** **cumulative unemployment benefits during their first three years in Finland**?

Elements of causal questions

1. **Treatment**
   - impact *of* [...]
2. **Counterfactual**
   - impact *in comparison to* [...]
3. **Outcome**
   - impact *on* [...]
4. **Population**
   - impact *for* [...]

- What is a well-defined question for our case study?
  - my take: what is the impact of the **new program** in comparison to **business-as-usual programs** on **participants'** **cumulative unemployment benefits during their first three years in Finland**?
  - this is just one example of a well-defined question, there are also many others (even in the context of this specific example)
- Next: formal definitions using the potential outcomes framework

Potential Outcomes

- We focus on binary (0/1) treatments and denote **treatment status** of individual i as

$$D_i = \begin{cases} 1 & \text{if she receives the treament} \\ 0 & \text{if she doesn't} \end{cases}$$

Potential Outcomes

- We focus on binary (0/1) treatments and denote **treatment status** of individual i as

$$D_i = \begin{cases} 1 & \text{if she receives the treament} \\ 0 & \text{if she doesn't} \end{cases}$$

- We denote **outcomes** by y and define

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

Potential Outcomes

- We focus on binary (0/1) treatments and denote **treatment status** of individual i as

$$D_i = \begin{cases} 1 & \text{if she receives the treament} \\ 0 & \text{if she doesn't} \end{cases}$$

- We denote **outcomes** by y and define

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

in words: y_{1i} is the outcome of individual i in the state of the world where she is treated and y_{0i} is her outcome in the state of the world where she was *not* treated

Only one potential outcome can occur

Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;

...

I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.

Robert Frost (1915): The Road Not Taken

Robert Lee Frost (1874–1963) was an American poet, who frequently wrote about settings from rural life, using them to examine complex social and philosophical themes. Source: Wikipedia

Treatment effect

- The treatment effect for individual i is:

$$y_{1i} - y_{0i}$$

in words: difference in the potential outcomes with and without the treatment

Treatment effect

- The treatment effect for individual i is:

$$y_{1i} - y_{0i}$$

in words: difference in the potential outcomes with and without the treatment

- The fundamental challenge of causal inference is that we cannot observe both y_{1i} and y_{0i} for the same individual. Instead, we observe

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

Treatment effect

- We can never identify the treatment effect for an individual person, but sometimes we can estimate average treatment effects:

$$\text{Average treatment effect (ATE)} \quad = \quad \mathbb{E}\left[y_{i1} - y_{0i}\right]$$

Treatment effect

- We can never identify the treatment effect for an individual person, but sometimes we can estimate average treatment effects:

$$\begin{aligned} \text{Average treatment effect (ATE)} &= \mathbb{E}\left[y_{i1} - y_{0i}\right] \\ \text{ATE for the treated (ATT)} &= \mathbb{E}\left[y_{i1} - y_{0i} \middle| D_i = 1\right] \end{aligned}$$

where $\mathbb{E}\left[a|b\right]$ is the expectation of a conditional on b

Treatment effect

- We can never identify the treatment effect for an individual person, but sometimes we can estimate average treatment effects:

$$\begin{aligned} \text{Average treatment effect (ATE)} &= \mathbb{E}\left[y_{i1} - y_{0i}\right] \\ \text{ATE for the treated (ATT)} &= \mathbb{E}\left[y_{i1} - y_{0i} \middle| D_i = 1\right] \end{aligned}$$

where $\mathbb{E}\left[a|b\right]$ is the expectation of a conditional on b

- Why ATE *and* ATT?
  - treatment effect may be different for those getting the treatment than it would be for those not getting it (e.g. specific integration policy)
  - internal validity: do we learn the true effect for the treated population?
  - external validity: can we extrapolate to other populations?

Research designs and control groups

- We use a *comparison* or *control* group to approximate what would have happened to the treated in the absence of the treatment
  - that is, we *estimate* the counterfactual $\mathbb{E}[y_{0i}|D_i = 1]$

Research designs and control groups

- We use a *comparison* or *control* group to approximate what would have happened to the treated in the absence of the treatment
  - that is, we *estimate* the counterfactual $\mathbb{E}[y_{0i}|D_i = 1]$
- In economics parlance, this approach is know as "**design-based**" or "**reduced form**" or "**experimental**" approach
  - the alternative is the "structural" approach, where we use quantitative economic models to simulate counterfactual states of the world

Research designs and control groups

- We use a *comparison* or *control* group to approximate what would have happened to the treated in the absence of the treatment
  - that is, we *estimate* the counterfactual $\mathbb{E}[y_{0i}|D_i = 1]$
- In economics parlance, this approach is know as "**design-based**" or "**reduced form**" or "**experimental**" approach
  - the alternative is the "structural" approach, where we use quantitative economic models to simulate counterfactual states of the world
- Invalid control group leads to **selection bias**
  - whether the control group provides a good counterfactual or not is the key question of all design-based causal inference

Selection Bias

- As the amount of data increases, the sample averages approach the population average (expectations)

$$\underbrace{Avg[y_i|D=1]}_{\text{treatment group}} - \underbrace{Avg[y_i|D=0]}_{\text{control group}} \rightarrow \mathbb{E}[y_i|D=1] - \mathbb{E}[y_i|D=0]$$

Selection Bias

- As the amount of data increases, the sample averages approach the population average (expectations)

$$\underbrace{Avg[y_i|D=1]}_{\text{treatment group}} - \underbrace{Avg[y_i|D=0]}_{\text{control group}} \rightarrow \mathbb{E}[y_i|D=1] - \mathbb{E}[y_i|D=0]$$

$$= \mathbb{E}[y_{1i}|D=1] - \mathbb{E}[y_{0i}|D=0]$$

Selection Bias

- As the amount of data increases, the sample averages approach the population average (expectations)

$$\underbrace{Avg[y_i|D=1]}_{\text{treatment group}} - \underbrace{Avg[y_i|D=0]}_{\text{control group}} \rightarrow \mathbb{E}[y_i|D=1] - \mathbb{E}[y_i|D=0]$$

$$= \mathbb{E}[y_{1i}|D=1] - \mathbb{E}[y_{0i}|D=0]$$

- Where the second row emphasizes that we observe y_{0i} only for the control group, while our objective is to estimate ATT, i.e.

$$\mathbb{E}\left[y_{i1} - y_{0i}|D_i = 1\right] = \mathbb{E}[y_{1i}|D=1] - \underbrace{\mathbb{E}[y_{0i}|D=1]}_{\text{never observed}}$$

Selection Bias

- As the amount of data increases, the sample averages approach the population average (expectations)

$$\underbrace{Avg[y_i|D=1]}_{\text{treatment group}} - \underbrace{Avg[y_i|D=0]}_{\text{control group}} \quad \rightarrow \quad \mathbb{E}[y_i|D=1] - \mathbb{E}[y_i|D=0]$$

$$= \quad \mathbb{E}[y_{1i}|D=1] - \mathbb{E}[y_{0i}|D=0]$$

- Where the second row emphasizes that we observe y_{0i} only for the control group, while our objective is to estimate ATT, i.e.

$$\mathbb{E}\left[y_{i1} - y_{0i}|D_i = 1\right] \quad = \quad \mathbb{E}[y_{1i}|D=1] - \underbrace{\mathbb{E}[y_{0i}|D=1]}_{\text{never observed}}$$

- Selection bias arises when a control group leads to an incorrect estimate of the counterfactual, i.e. $\mathbb{E}[y_{0i}|D=0] \neq \mathbb{E}[y_{0i}|D=1]$

Selection Bias

- A particularly informative way to illustrate selection bias is:

$$\mathbb{E}[y_i|D=1] - \mathbb{E}[y_i|D=0] \quad = \quad \mathbb{E}[y_{1i}|D=1] - \mathbb{E}[y_{0i}|D=0]$$

Selection Bias

- A particularly informative way to illustrate selection bias is:

$$\begin{aligned}
\mathbb{E}[y_i|D=1]-\mathbb{E}[y_i|D=0] &= \mathbb{E}[y_{1i}|D=1]-\mathbb{E}[y_{0i}|D=0] \\
&= \underbrace{\mathbb{E}[y_{1i}|D=1]-\mathbb{E}[y_{0i}|D=1]}_{\text{ATT}}+ \\
&\quad \underbrace{\mathbb{E}[y_{0i}|D=1]-\mathbb{E}[y_{0i}|D=0]}_{\text{Selection bias}}
\end{aligned}$$

Selection Bias

- A particularly informative way to illustrate selection bias is:

$$\begin{aligned}\mathbb{E}[y_i|D=1]-\mathbb{E}[y_i|D=0] &= \mathbb{E}[y_{1i}|D=1]-\mathbb{E}[y_{0i}|D=0] \\ &= \underbrace{\mathbb{E}[y_{1i}|D=1]-\mathbb{E}[y_{0i}|D=1]}_{\text{ATT}}+ \\ &\quad \underbrace{\mathbb{E}[y_{0i}|D=1]-\mathbb{E}[y_{0i}|D=0]}_{\text{Selection bias}}\end{aligned}$$

where the first step is from the previous slide and the second step is taken by simply adding and substracting $\mathbb{E}[y_{0i}|D=1]$

  - i.e. $\mathbb{E}[y_{0i}|D=1]-\mathbb{E}[y_{0i}|D=1]=0$, so including it does not change the result, but allows us to rewrite the equation as ATT+SB

Selection Bias

- A particularly informative way to illustrate selection bias is:

$$\begin{aligned}\mathbb{E}[y_i|D=1]-\mathbb{E}[y_i|D=0] &= \mathbb{E}[y_{1i}|D=1]-\mathbb{E}[y_{0i}|D=0]\\ &= \underbrace{\mathbb{E}[y_{1i}|D=1]-\mathbb{E}[y_{0i}|D=1]}_{\text{ATT}}+\\ &\quad \underbrace{\mathbb{E}[y_{0i}|D=1]-\mathbb{E}[y_{0i}|D=0]}_{\text{Selection bias}}\end{aligned}$$

  where the first step is from the previous slide and the second step is taken by simply adding and substracting $\mathbb{E}[y_{0i}|D=1]$
  - i.e. $\mathbb{E}[y_{0i}|D=1]-\mathbb{E}[y_{0i}|D=1]=0$, so including it does not change the result, but allows us to rewrite the equation as ATT+SB
  - in words: differences in the average outcomes between treatment and control groups include the treatment effect *and* the selection bias (the difference between the two groups if neither had been treated)

In-class discussion: Selection bias and integration policies

- Let's return to the case of new integration program and speculate about the likely selection bias in two alternative control groups:
  1. all immigrants not participating in the program
  2. unemployed immigrants entering the employment services at the same time, but participate in other types of programs

In-class discussion: Selection bias and integration policies

- Let's return to the case of new integration program and speculate about the likely selection bias in two alternative control groups:
  1. all immigrants not participating in the program
  2. unemployed immigrants entering the employment services at the same time, but participate in other types of programs
- Let's assume that the new program consists of
  - 60 days intensive language training
  - followed by 6 months of guaranteed real low-skilled job

  while the business-as-usual model includes
  - 1yr standard language training
  - "graduation" into standard unemployment

Randomization eliminates selection bias

- Random assignment into treatment/control ensures that the control groups is *comparable* to the treatment group

Randomization eliminates selection bias

- Random assignment into treatment/control ensures that the control groups is *comparable* to the treatment group
- Formally: their potential outcomes are in expectation the same, i.e.

$$\mathbb{E}[y_{1i}|D = 1] = \mathbb{E}[y_{1i}|D = 0]$$
$$\mathbb{E}[y_{0i}|D = 1] = \mathbb{E}[y_{0i}|D = 0]$$

Randomization eliminates selection bias

- Random assignment into treatment/control ensures that the control groups is *comparable* to the treatment group
- Formally: their potential outcomes are in expectation the same, i.e.

$$\mathbb{E}[y_{1i}|D=1] = \mathbb{E}[y_{1i}|D=0]$$
$$\mathbb{E}[y_{0i}|D=1] = \mathbb{E}[y_{0i}|D=0]$$

- Thus $\mathbb{E}[y_{0i}|D=1] - \mathbb{E}[y_{0i}|D=0] = 0$, i.e. no selection bias
  - in words: the control group tells us what would have happened to the treatment group in the absence of the treatment

The ethics of RCTs

- The key ethical concern of RCTs is the unequal treatment of the treatment and control group
  - sometimes a question of life and death (e.g. early AIDS medication)
- Nevertheless, drug approval requires extensive clinical trials. Why?

The ethics of RCTs

- The key ethical concern of RCTs is the unequal treatment of the treatment and control group
  - sometimes a question of life and death (e.g. early AIDS medication)
- Nevertheless, drug approval requires extensive clinical trials. Why?
  - The 1960's thalidomide tragedy led to stricker requirements that drugs have to be proved to be safe and effective before they are marketed
  - the proof comes from clinical trials (RCTs)

The ethics of RCTs

- No country requires similar proof for social policies
    - inevitable because not all policies can be tested with RCTs
- But why not study the impact of policies suitable for experimental research designs using the most reliable methods?

The ethics of RCTs

- No country requires similar proof for social policies
  - inevitable because not all policies can be tested with RCTs
- But why not study the impact of policies suitable for experimental research designs using the most reliable methods?
  - my interpretation: policy makers often have a gut feeling that RCTs are somehow immoral (without having really thought this through)
- Aalto Economic Institute is part of this debate
  - see e.g our recent reports on social experiments and ex-post evaluations (if you are interested; i.e. this is not a requirement for this course)
  - we've closely worked with the government in designing RCTs
    - e.g., the ongoing two-year preschool experiment

The ethics of RCTs

- RCTs can reveal what works (and for whom)
  - benefits those potentially getting the treatment later
  - Finnish parliament: acceptable to deviate from equal treatment in social experiments if there is sufficiently strong reason (PeVL 51/2016)

The ethics of RCTs

- RCTs can reveal what works (and for whom)
  - benefits those potentially getting the treatment later
  - Finnish parliament: acceptable to deviate from equal treatment in social experiments if there is sufficiently strong reason (PeVL 51/2016)
- Resource constraints often create unequal treatment in any case
  - e.g. fewer slots in a new training program than suitable participants
  - randomization provides all eligible persons equal chance to participate

The ethics of RCTs

- RCTs can reveal what works (and for whom)
  - benefits those potentially getting the treatment later
  - Finnish parliament: acceptable to deviate from equal treatment in social experiments if there is sufficiently strong reason (PeVL 51/2016)
- Resource constraints often create unequal treatment in any case
  - e.g. fewer slots in a new training program than suitable participants
  - randomization provides all eligible persons equal chance to participate
- Typically we do not know whether the treatment is beneficial or not
  - if we knew the effects, we would not have to run the experiment

The ethics of RCTs

- RCTs can reveal what works (and for whom)
  - benefits those potentially getting the treatment later
  - Finnish parliament: acceptable to deviate from equal treatment in social experiments if there is sufficiently strong reason (PeVL 51/2016)
- Resource constraints often create unequal treatment in any case
  - e.g. fewer slots in a new training program than suitable participants
  - randomization provides all eligible persons equal chance to participate
- Typically we do not know whether the treatment is beneficial or not
  - if we knew the effects, we would not have to run the experiment
- Features of ethically sound experiments
  - always: never cause harm knowingly, privacy protection, pre-evaluation of risks and benefits, reliable measurement, approriate test population
  - usually: informed consent (e.g. possibility to opt-out)

The limits of RCTs

- RCTs are a powerful tool for reliably answering some questions
  - but: many questions cannot be be answered with RCTs
  - it would be crazy to focus *only* on question suitable for RCTs

The limits of RCTs

- RCTs are a powerful tool for reliably answering some questions
  - but: many questions cannot be be answered with RCTs
  - it would be crazy to focus *only* on question suitable for RCTs
- RCTs are not helpful and/or possible when
  - treatment affects everyone (e.g. monetary policy)
  - the experiment would be unethical or too impractical/expensive
  - the study population differs (too much) from the relevant population
  - relevant follow-up period is impractically long
- Even when RCTs are feasible, they only guarantee internal validity

Summary

- **Causality**: how one thing *affects* another thing
  - requires comparing counterfactual states of the world to each other ("how would Y change if we changed X?")
  - at most, one of them is observed
- **Control group** in an experimental research design
  - the outcomes of the control group are used to infer what would have happened to the treatment group in the absence of the treatment
- **Selection bias** occurs when the control group is not comparable to the treatment group, i.e. $\mathbb{E}[y_{0i}|D=0] \neq \mathbb{E}[y_{0i}|D=1]$
  - = potential outcomes differ between the treatment and control groups
- **Randomization** eliminates selection bias
  - on expectation, the only difference between the groups is that the treatment group gets the treatment and the control group does not
  - $\rightarrow$ differences in average outcomes must be due to the treatment