# Causal Inference

Falco J. Bargagli-Stoffi

UCLA
falco@ucla.edu

Week 1: *Foundations of Causal Inference*
Spring, 2025

# Outline

# Course description

- Biostat M235 is a class in causal inference
- Course meets on Monday and Tuesday 4 - 5.50 pm
- Office hours: schedule online at `https://doodle.com/bp/falcoj.bargaglistoffi/meeting`
- Webiste: `https://dviviano.github.io/classes/webpage_topicsmetrics.html`

# Intended audience

- This course covers key concepts and state-of-the-art methods for causal inference from randomized experiments and observational studies
- The course is designed for all students either:
  (i) interested in pursuing research in methodological causal inference
  (ii) interested in learning recent advances in causal inference for applied research

# Tentative list of topics

I will cover a subset of the following modules (depending on time).

- 

# Assignments

- 3 problem sets in total due Monday before class on date
  - Problem sets count towards 45% of the grade (15% each)
  - Submit the problem sets on where
- Each assginment has two graded questions
  - A short referee report of one paper between two papers provided
  - A question of your choice between two alternative questions (one conceptual and one computational)
- Solutions will be posted shortly after assignements are due (no late assignments)
- You can collaborate on assignments, but shoud write your own solution

# Final exam

- Final exam consists of choosing a paper of your choice and write an essay about the paper
- The essay can be written (and submitted) together with another student
- The final essay should be about ten pages (and no more than 15 pages with references)
- The essay should choose one paper from the list on the website
- You can pick a different paper in which case please come and talk with me to choose the paper together
- You are welcome to discuss the essay and its topics any time during office hours

# Structure of the essay

- The essay should be structured as follows
  - (i) Intro that motivates the method for applied econ research
  - (ii) A section discussing the setup
  - (iii) A section discussing the methodology and presenting the main analysis – this needs to contain only one of the main argument of the paper
  - (iv) A section presenting the numerical properties of the procedure and an empirical application (this needs to present a new set of simulations using the method in the paper, and an application using some different existing dataset)
- Rrecommended reading: The Little Book of Research Writing by Varanya Chaubey (can be ordered online)

# Evaluation

- 10% in-class participation (attendance sufficient, participation encouraged)
- 45% assignments
- 45% final essay and 15 minutes presentation during the last week

# Other notes

- Slides are self-contained, no textbook is mandatory
- I will point to some chapters of some books or papers as we go through the material.
- Online you can find lots of additional material: I will point to some online lectures
- The class is intended for students interested in applying causal inference in their research
  - If you are interested in doing research in causal inference come and talk to me to get more references on things you find interesting

# Outline

# Part 1: Organization

- Definitions
- Fundamental Assumptions
- Average effects and estimands

Chapter readings:

- Ch. 2 of [**?**], Ch. 1, 3 of [**?**]

# Causality in Ancient Greek Philosophy

*I would rather discover one causal law than be King of Persia.*
*— Democritus*

*We have knowledge of a thing only when we have grasped its cause.*
*— Aristotle, Posterior Analytics*

*We do not have knowledge of a thing until we have grasped its why, that is to say, its cause.*
*— Aristotle, Physics*

# "Mechanical" notions of Causality

- "*Development of Western science is based on two great achievements: the invention of the formal logical system (in Euclidean geometry) by the Greek philosophers, and the discovery of the possibility to find out causal relationships by systematic experiment (during the Renaissance)*" (Einsten, 1953)
- Causality is at the core of science
- In physical sciences is (sometimes) mechanical (experiments suffice)
- In social sciences... tricky: for a given individual, we can only observe one world ("fundamental problem of causal inferenceâ, Holland 1986)
- Even with experiments, we need assumptions to get to causal effects

# Questions on Causation

- Relevant questions about causation:
  - The philosophical meaningfulness of the notion of causation
  - Deducing the causes of a given effect
  - Understanding the details of causal mechanism

- In this class we focus on measuring the effects of causes—a place where statistics, which is concerned with measurement, has most contributions to make

# Causal Inference in Statistics

- Causality is a subject at the frontier of academic debate in statistics and in many areas of medical, social, economic and political sciences

- Research questions motivating most studies in statistics-based sciences are causal in nature:

*Without causal conclusions there are little policy implications*

# Examples of Causal Questions

- What are the effects of job-market training programs on labor market outcomes (e.g., yearly earnings or employment status)?
- What are the effects of financial aids to firms on firms' development?
- How effective is a micro-credit program in reducing poverty in a developing country?
- What are the effects of educational programs on students' academic performance (as measured by, e.g., test scores)?
- What is the effect of incarceration on re-offending?
- What are the effects of changing a health policy?
  - ✓ What is the effect of stricter regulatory policies for air pollution on life-expectation?
- How effective is a drug treatment?
- What are the effects of the mask mandate to slow down the spread of COVID-19 on mortality?

# Association/Correlation versus Causation

- Standard statistical analysis is to infer associations among variables, based on which one may do some prediction
- Causal analysis is one step further: it is about counterfactual prediction:
  i.e., predict what would have happened to the same units/subjects had they been exposed to a different (counterfactual) condition
- In most cases, *association/correlation does not imply causation*

# Causal Inference needs Assumptions

- How to make the leap from association to causation?
- Key: causal assumptions - structural and/or modeling
- Causal inference is about
  1. build a framework and define causal effects under general scenarios
  2. specify assumptions under which one can decare/identify causation from association
  3. assess the sensitivity to the causal assumptions and find ways to mitigate

# Outline

# Frameworks for Causal Inference

- The purpose is to construct a model or a framework that is complex enough to allow us to formalize basic intuitions concerning cause and effect
- Two commonly used frameworks
  - The *potential outcome framework*, also known as the counterfactual framework, or the Neyman-Rubin Causal Model (Neyman, 1923; Rubin, 1974; Imbens and Rubin, 2015; Hernan and Robins, 2020)
  - The *causal diagram* framework (Pearl, 2009)
  - Mathematically the two frameworks are connected (Richardson and Robins, 2013), but each has different established goals, tools and applicable areas (Imbens, 2020)
  - This class focuses on the potential outcome framework, and will occasionally draw directed acyclic graphs (DAGs) for simple illustration

# Potential Outcome Framework

- The potential outcome (PO) framework is also known as the counterfactual framework, or the Rubin Causal Model (*Holland, 1986*)
- Textbooks: Imbens and Rubin (2015, key textbook in this course); Angrist and Pischke (2009, 2015); Hernan and Robins (2020)
- Reviews: Imbens and Wooldridge (2009), Mealli et al., (2011); Mattei et al. (2020); Dominici et al. (2021)
- The PO framework has its roots in the work on randomized experiments and randomization-based inference by Fisher (1925) and Neyman (1923) and has been extended by Rubin (Rubin, 1974, 1977, 1978), and subsequently exploited by others, to more complicated situations, including nonrandomized studies, and alternative modes of inference other than randomization-based

# Potential Outcome Framework

- The potential outcome framework is arguably the most widely used causal framework across many disciplines, e.g., policy, social sciences, economics, public health

- *No causation without manipulation*: "cause" must be (hypothetically) manipulatable, e.g., intervention, treatment

  - ✓ Causality is tied to an action (intervention, treatment, manipulation) applied to a unit

  - ✓ Manipulation need not be performed, but should be theoretically possible

- Gender, race and age are not well defined "causes" under the potential outcome framework

# PO Framework: Basic Concepts/Primitives (I)

- Unit: A physical object at a particular place and point in time
  - ✓ A physical object may be a person, a firm, a region, a state or a thing or collection of objects or persons, such as a family, a classroom, a school, a market
  - ✓ The physical object at two different times is a different unit
- Treatment: An intervention, the effects of which (on some particular measurement of the units) the investigator wishes to assess
  - ✓ The same unit can be subject or exposed to a specific treatment (a particular action, intervention, policy or program) or to alternative treatments, which could be different active treatments, or no treatment at all
- We will mostly focus on settings with two treatments (binary treatments)
  - ✓ "Active/new treatment" versus "control/standard treatment" or, simply, treatment versus control

# PO Framework: Basic Concepts/Primitives (II)

- Potential Outcomes: The values of a unit's measurement of interest after ($a$) application of the treatment and ($b$) non-application of the treatment (i.e., under control)
- A potential outcome for each unit/treatment pair: for each unit there are two associated potential outcomes at a future point in time after treatment
  - ✓ The value of the outcome of interest if the unit was exposed to the active treatment; and
  - ✓ The value of the outcome of interest at the same future point in time if the unit was exposed to the control treatment
- Causal Effect: For each unit, the comparison of the potential outcome under treatment and the potential outcome under control
- Covariate/pre-treatment variable: A background characteristic (measured or unmeasured) of a unit that could not have been affected by treatment assignment

# PO Framework: Basic Concepts/Primitives (III)

- Estimand: "to to be estimated"
- Estimate: approximation of estimand using finite sample
- Estimator: the method/algorithm to construct the estimate

- Identification: "Econometric identification really means just one thing: model parameters or features being uniquely determined from the observable population that generates the data" lewbel2019identification
- Testability: an hypothesis (or assumption) is testable if we can find a necessary condition of such an hypothesis that can be rejected using information from a given sample draw.
- Experiment: a research design under which we know treatment assignment mechanism (typically manipulated by the researcher)
- Observational study: research design under which we do not know the assignment mechanism (typically need assumptions)

## Example: Assessing the Impact of Aspirin on Headache

- Unit: A subject (e.g., you), at a particular point in time, with headache
- Treatment: Taking an aspirin (active treatment) or not taking an aspirin (control treatment)
- Potential Outcomes: The future values of the subject's headache, e.g., an hour later the decision of taking versus not taking the aspirin ($a$) if the subject took the aspirin and ($b$) if the subject did not take the aspirin
- Causal Effect: The comparison–e.g., difference–between the potential future headache status if the subject took the aspirin and the potential future headache status if the subject did not take the aspirin
- Covariates: Individual background characteristics (e.g., gender, age, education levels) + Prior health histories (e.g., intensity of the headache before making the decision to take the aspirin or not)

# Well- and Ill-defined Causal Statements

- My headache went away because I took an aspirin
  - ✓ Well-defined causal statement
- She got a good job last year because she went to college
  - ✓ Poorly defined causal statement: It is not clear what the treatment and its alternatives are
- She has long hair because she is a girl
  - ✓ Ill-defined causal statement: There is no clear action described that would have allowed us to observed the unit exposed to the alternative treatment ("being a boy")

# Notation

- Hypothetical example: one unit
  - ✓ A single person, e.g., You, at a particular point in time, with headache, contemplating whether or not to take an aspirin for your headache
- Binary treatment variable: $W = w$,

  $w = 1$ (active/new treatment) and $w = 0$ (control/standard treatment)

  - ✓ $W$ = Indicator for taking versus not taking an aspirin:
    $w = 1$ (Aspirin) and $w = 0$ (No aspirin)
- Outcome variable: $Y$
  - ✓ $Y$ = Intensity of your headache an hour later the decision of taking versus not taking the aspirin
- Potential outcomes: $Y(0)$ and $Y(1)$
  - ✓ $Y(0)$ = Intensity of your headache an hour later if you did not take the aspirin
  - ✓ $Y(1)$ = Intensity of your headache an hour later if you took the aspirin
- Causal effect: $Y(1)$ versus $Y(0)$
  - ✓ $Y(1) - Y(0)$

# Causal Effect: Examples

- Outcome = Intensity of headache

| | Potential Outcomes | | Causal effect on $Y$ |
|---|---|---|---|
| Unit | $Y(0)$ | $Y(1)$ | $Y(1) - Y(0)$ |
| You | 75 | 25 | $25 - 75 = -50$ |

- Gain score = Variation in the outcome

| | Pre-treatment headache intensity | Gain Scores Potential Outcomes | | Causal effect |
|---|---|---|---|---|
| Unit | $X$ | $Y(0) - X$ | $Y(1) - X$ | $[Y(1) - X] - [Y(0) - X$ |
| You | 85 | $-5$ | $-55$ | $-55 - (-5) = -50$ |

- Improvement due to the aspirin

# The Fundamental Problem of Causal Inference

- The fundamental problem of causal inference: We can observe at
  most one of the potential outcomes for each unit, the other(s) are
  missing/counterfactual
- Realized and possibly observed potential outcome: potential outcome corresponding to the action actually taken

$$Y^{obs} = Y(W) = WY(1) + (1 - W)Y(0) = \begin{cases} Y(1) & \text{if } W = 1 \\ Y(0) & \text{if } W = 0 \end{cases}$$

- Ex post, the other potential outcomes cannot be observed because the corresponding actions that would lead to them being realized were not taken

$$Y^{mis} = Y(1-W) = (1-W)Y(1)+WY(0) = \begin{cases} Y(0) & \text{if } W = 1 \\ Y(1) & \text{if } W = 0 \end{cases}$$

# The Fundamental Problem of Causal Inference: Example

| | Not Observable | | | Known | |
|---|---|---|---|---|---|
| | Potential Outcomes | | Causal effect on $Y$ | Actual Treatment | Observed outcome |
| Unit | $Y(0)$ | $Y(1)$ | $Y(1) - Y(0)$ | $W$ | $Y^{obs}$ |
| You | 75 | 25 | −50 | Aspirin ($W = 1$) | 25 ($Y^{obs} = Y(1)$) |

# Causal Effects: Remarks

- The definition of the causal effect depends on the potential outcomes, but it does not depend on which outcome is actually observed
- The causal effect is the comparison of potential outcomes, for the same unit, at the same moment in time post-treatment
  - ✓ The causal effect is not defined in terms of comparisons of outcomes at different times, as in a before-and-after comparison
  - ✓ In the aspirin example the causal effect is not defined as before and after comparison of your headache before and after deciding to take or not to take the aspirin

$$Y^{obs} - X = 25 - 85 = -60$$

# Causal Inference

- The fundamental problem of causal inference makes it clear that causal inference under the potential outcome framework is essentially a missing data problem
- How do we learn about the missing potential outcomes and therefore the causal effects?
- To identify causal effects from observed data, under any mathematical framework, one must make assumptions (structural
or/and stochastic)

*Separate the Science (object of inference) from what we do to learn about the Science (intervene at a particular point in time)*

# Learning about Causal Effects: Multiple Units

- We must observe multiple units, some exposed to the active treatment, some exposed to the alternative (control) treatment to make causal inferences

  - ✓ Observe the same physical object under different treatment levels at different points in time

  - ✓ Observe different physical objects at approximately the same time

# Multiple Units: The Same Unit at Different Times

- Comparison of multiple units: the same physical object at different times under treatment and under control
- Fundamental mistake: Viewing the same physical object at different times as the same unit
  - ✓ The same physical unit, "yourself at different times" at different times is not the same unit in the PO approach to causality
- Time matters for many reasons
  - ✓ It may be reasonable to assume that time makes little difference for inanimate objects, but this assumption is typically less reasonable with human subjects

# Multiple Units: Many Units at the Same Time

- Two subjects, you and I, have headache, but only one of us take an aspirin
- We may attempt to infer the efficacy of taking aspirin by comparing our subsequent headaches
- It is more obvious here that "you" and "I" at the same point in time are different units
- Warning: By itself, however, the presence of multiple units does not solve the problem of causal inference

# Multiple Units: Many Units at the Same Time

| Treatment status | | | |
|---|---|---|---|
| You (Unit 1) | I (Unit 2) | Potential outcomes for you | Potential outcomes for me |
| 1 | 1 | $Y_1([1,1]) = 0$ | $Y_2([1,1]) = 0$ |
| 0 | 0 | $Y_1([0,0]) = 100$ | $Y_2([0,0]) = 100$ |
| 1 | 0 | $Y_1([1,0]) = 50$ | $Y_2([1,0]) = 100$ |
| 0 | 1 | $Y_1([0,1]) = 75$ | $Y_2([0,1]) = 0$ |

# Multiple Units: Many Units at the Same Time

- We can still only observe at most one of these four potential outcomes for each unit: the one corresponding to whether you and I took, or did not take, an aspirin
- Each level of the treatment now indicates both whether you take an aspirin and whether I do
- Six different comparisons defining causal effects for each of us, depending on which two of the four potential outcomes for each unit are conceptually compared: For each $i = 1, 2$,

$$Y_i([1,1]) - Y_i([0,0]) \quad Y_i([1,1]) - Y_i([1,0]) \quad Y_i([1,1]) - Y_i([0,1$$

$$Y_i([1,0]) - Y_i([0,0]) \quad Y_i([1,0]) - Y_i([0,1]) \quad Y_i([0,1]) - Y_i([0,0$$

# The Stable Unit Treatment Value Assumption (SUTVA)

## SUTVA (*Rubin, 1980*)

The potential outcomes for any unit do not vary with the treatments assigned to other units, and, for each unit, there are no different forms or versions of each treatment level, which lead to different potential outcomes

- SUTVA includes two sub-assumptions:
  1. No interference
  2. No different versions of a treatment, also known as consistency
- Under SUTVA, we have only two potential outcomes for each unit, $Y_i(0)$ and $Y_i(1)$, and for each unit $i$

$$Y_i^{obs} = Y_i(W_i) = W_i Y_i(1) + (1 - W_i) Y_i(0) = \begin{cases} Y_i(1) & \text{if } W_i = 1 \\ Y_i(0) & \text{if } W_i = 0 \end{cases}$$

# SUTVA in the Aspirin Example

- The no-interference component of SUTVA implies that whether you take an aspirin has no effect on the status of my headache
- Under SUTVA, for you (unit 1) and for me (unit 2), we have

$$Y_1(1) = Y_1([1,1]) = Y_1([1,0]) \quad \text{and} \quad Y_1(0) = Y_1([0,1]) = Y_1([0,0$$

$$Y_2(1) = Y_2([1,1]) = Y_2([0,1]) \quad \text{and} \quad Y_2(0) = Y_2([1,0]) = Y_2([0,0$$

and, for $i = 1, 2$,

$$Y_i^{obs} = \begin{cases} Y_i(1) & \text{if } W_i = 1 \text{ (aspirin)} \\ Y_i(0) & \text{if } W_i = 0 \text{ (no aspirin)} \end{cases}$$

# SUTVA: Remarks

- SUTVA seems trivial, actually very strong assumption
- SUTVA is a type of exclusion restriction assumption
  - ✓ It rules out the existence of a causal effect of a particular treatment relative to an alternative
- In the aspirin example. . .
  - ✓ Your taking or not taking aspirin has not any effect on my headache; and
  - ✓ There is no the possibility that the aspirin tablets available to me are of different strengths
- SUTVA is not directly informed by observations: It relies on previously acquired knowledge of the subject matter knowledge

# Basic Setup under SUTVA

- Data: a random sample of $N$ units from a target population
  - ✓ Target Population: a well-defined population of individuals
    whose outcomes are going to be compared
- A vector of $p$ covariates $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})$
- A treatment with two levels, $w = 0, 1$
- For each unit $i$, there exsist two potential outcomes $(Y_i(0), Y_i(1))$: the outcomes under the two values of the treatment, at most one of which is observed
- $\boldsymbol{X} = [X_{i1}, \ldots, X_{ip}]_{i=1}^{N}$, $\boldsymbol{Y}(0) = [Y_i(0)]_{i=1}^{N}$, $\boldsymbol{Y}(1) = [Y_i(1)]_{i=1}^{N}$
- The science: $[\boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)]$

# The Science Under SUTVA

| | Covariates | | | Potential Outcomes | |
|---|---|---|---|---|---|
| | | | | Treatment | Control |
| Units | $X_1$ | $\ldots$ | $X_p$ | $Y(1)$ | $Y(0)$ |
| 1 | $X_{11}$ | $\ldots$ | $X_{1p}$ | $Y_1(1)$ | $Y_1(0)$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $i$ | $X_{i1}$ | $\ldots$ | $X_{ip}$ | $Y_i(1)$ | $Y_i(0)$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $N$ | $X_N$ | $\ldots$ | $X_{Np}$ | $Y_N(1)$ | $Y_N(0)$ |

# Causal Estimands

- Unit-level causal effect: a comparison of treatment and control potential outcomes for a single unit $i$: $Y_i(1)$ versus $Y_i(0)$

  - ✓ Examples: $Y_i(1) - Y_i(0)$; $\log\{Y_i(1)\} - \log\{Y_i(0)\}$; $Y_i(1)/Y_i(0)$

- Summary causal effects Comparison of $Y_i(1)$ vs $Y_i(0)$ for a common set of units

  - ✓ Examples: Summary causal effects for all units or for sub-groups of units defined by, e.g., values of covariates

# The Science and Causal Estimands

| | Covariates | | | Potential Outcomes | | | |
|---|---|---|---|---|---|---|---|
| | | | | Treatment | Control | Unit-level | Summary |
| Units | $X_1$ | ... | $X_p$ | $Y(1)$ | $Y(0)$ | Causal Effects | Causal Effects |
| 1 | $X_{11}$ | ... | $X_{1p}$ | $Y_1(1)$ | $Y_1(0)$ | $Y_1(1)$ vs $Y_1(0)$ | Comparison o |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |
| $i$ | $X_{i1}$ | ... | $X_{ip}$ | $Y_i(1)$ | $Y_i(0)$ | $Y_i(1)$ vs $Y_i(0)$ | $Y_i(1)$ vs $Y_i(0$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | for a commor |
| $N$ | $X_N$ | ... | $X_{Np}$ | $Y_N(1)$ | $Y_N(0)$ | $Y_N(1)$ vs $Y_N(0)$ | set of units |

# Finite-Sample Causal Estimands

- Finite-sample perspective: Focus on causal estimands for the finite set of $N$ units participating in the study
- Finite Sample Average Treatment Effect

$$ATE_{FS} = \frac{1}{N}\sum_{i=1}^{N}[Y_i(1) - Y_i(0)] = \frac{1}{N}\sum_{i=1}^{N}Y_i(1) - \frac{1}{N}\sum_{i=1}^{N}Y_i(0) = \overline{Y}(1) - \overline{Y}$$

- Finite Sample Average Treatment Effect for female

$$ATE_{FS}(f) = \frac{1}{N_f}\sum_{i:X_i=f}[Y_i(1) - Y_i(0)]$$

where $N_f$ is the number of female in the finite population/sample

- Finite Population Average Treatment Effect on the Treated units

$$ATT_{FS} = \frac{1}{N_t}\sum_{i:W_i=1}[Y_i(1) - Y_i(0)]$$

where $N_t$ is the number of units exposed to the active treatment in the finite population/sample

# Finite-Sample Causal Estimands: Example

| | | Potential Outcomes | | Unit-level |
|---|---|---|---|---|
| | Gender | Control | Treatment | Causal Effects |
| Unit | $X_i$ | $Y_i(0)$ | $Y_i(1)$ | |
| 1 | $M$ | 48 | 48 | 0 |
| 2 | $M$ | 54 | 20 | −34 |
| 3 | $M$ | 75 | 25 | 50 |
| 4 | $M$ | 90 | 30 | −60 |
| 5 | $F$ | 60 | 65 | 5 |
| 6 | $F$ | 62 | 18 | −44 |
| 7 | $F$ | 45 | 20 | −25 |
| 8 | $F$ | 66 | 26 | 40 |

$$ATE_{FS} = -\frac{248}{8} = \frac{252}{8} - \frac{500}{8} = 62.50 - 31.50 = -31$$

$$ATE_{FS}(f) = -\frac{104}{4} = \frac{129}{4} - \frac{233}{4} = 32.25 - 58.25 = -26$$

# Finite-Sample Causal Estimands: Example

| | | Potential Outcomes | | Unit-level | |
|---|---|---|---|---|---|
| | Gender | Control | Treatment | Causal Effects | Treatment |
| Unit | $X_i$ | $Y_i(0)$ | $Y_i(1)$ | | $W_i$ |
| 1 | $M$ | 48 | 48 | 0 | 1 |
| 2 | $M$ | 54 | 20 | –34 | 1 |
| 3 | $M$ | 75 | 25 | 50 | 0 |
| 4 | $M$ | 90 | 30 | –60 | 1 |
| 5 | $F$ | 60 | 65 | 5 | 0 |
| 6 | $F$ | 62 | 18 | –44 | 0 |
| 7 | $F$ | 45 | 20 | –25 | 1 |
| 8 | $F$ | 66 | 26 | 40 | 1 |

$$ATT_{FS} = \frac{159}{5} = \frac{144}{5} - \frac{303}{5} = 28.8 - 60.6 = -31.8$$

# Super-Population Causal Estimands

- Super-population perspective: Focus on causal estimands for a large super- population from which the $N$ units are considered as a random sample

- Super Population Average Treatment Effect

$$ATE_{SP} = \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$$

- Super Population Average Treatment Effect for female

$$ATE_{SP}(f) = \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = f]$$

- Super Population Average Treatment Effect on the Treated units

$$ATT_{SP} = \mathbb{E}[Y_i(1) - Y_i(0) \mid W_i = 1]$$

# Observed Data

- For each unit $i$, we observe
  - ✓ A vector of $p$ covariates $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{ip})$
  - ✓ The (binary) treatment status $W_i \in \{0, 1\}$
  - ✓ An outcome $Y_i^{obs}$
- $\boldsymbol{W} = [W_i]_{i=1}^N$, $\boldsymbol{Y}^{obs} = [Y_i^{obs}]_{i=1}^N$
- Potential outcomes and assignments jointly determine the values of the observed and missing outcomes:

$$Y_i^{obs} = Y_i(W_i) = W_i Y_i(1) + (1 - W_i) Y_i(0) = \begin{cases} Y_i(1) & \text{if } W_i = \\ Y_i(0) & \text{if } W_i = \end{cases}$$

$$Y^{mis} = Y_i(1 - W_i) = (1 - W_i) Y_i(1) + W_i Y_i(0) = \begin{cases} Y_i(0) & \text{if } W_i = \\ Y_i(1) & \text{if } W_i = \end{cases}$$

# Causal Estimands: Remarks

- Causal effects defined by potential outcomes
- Because causal effects are defined by comparing potential outcomes, they are well defined irrespective of the treatment actually taken
  - ✓ The definition of a causal effect does not depend on which treatment is observed
- Causal effects, in their definition, do not relate to coefficients of models

  They are not necesserily summaries of individual causal effects (think about difference in medians and median causal effects)

# What we are able to observe

| Units | Covariates $X$ | Treatment $Z$ | Potential Outcomes $Y(1)$ | $Y(0)$ | Unit-level Causal Effects |
|---|---|---|---|---|---|
| 1 | $X_1$ | 1 | $Y_1(1)$ | ? | ? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $i$ | $X_i$ | 0 | ? | $Y_i(0)$ | ? |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $N$ | $X_N$ | 1 | $Y_N(1)$ | ? | ? |

# Causal Estimands, Potential Outcomes and Observed Outcomes

- We observe at most half of all potential outcomes, and none of the unit-level causal effects
  - ✓ Given any treatment assigned to an individual unit, the potential outcome associated with any alternate treatment is missing
- There is an inferential problem associated with assessing causal effects: causal estimands are not identifiable without further assumptions
- A key role is played by the assignment mechanism
  - ✓ The assignment mechanism determines which units get which treatments or, equivalently, which potential outcomes are realized and which are not

# Perfect Doctor

| | Potential Outcomes | | | Observed Data | | | |
|---|---|---|---|---|---|---|---|
| Unit | $Y_i(0)$ | $Y_i(1)$ | $Y_i(1) - Y_i(0)$ | $W_i$ | $Y_i(0)$ | $Y_i(1)$ | $Y_i^{obs}$ |
| 1 | 1 | 7 | 6 | 1 | ? | 7 | 7 |
| 2 | 6 | 5 | −1 | 0 | 6 | ? | 6 |
| 3 | 1 | 5 | 4 | 1 | ? | 5 | 5 |
| 4 | 8 | 7 | −1 | 0 | 8 | ? | 8 |
| | 4 | 6 | 2 | | 7 | 6 | |
| | True Averages | | | | Observed Averages | | |

Observed mean difference between the treated and control outcome: −1

- The simple difference-in-means estimator does not return a valid estimate of the true causal effect (2): What went wrong?

# Implicit Imputation of Missing Potential Outcomes

| | Potential Outcomes | | | Observed Data | | | |
|---|---|---|---|---|---|---|---|
| Unit | $Y_i(0)$ | $Y_i(1)$ | $Y_i(1) - Y_i(0)$ | $W_i$ | $Y_i(0)$ | $Y_i(1)$ | $Y_i^{obs}$ |
| 1 | 1 | 7 | 6 | 1 | 7 | 7 | 7 |
| 2 | 6 | 5 | −1 | 0 | 6 | 6 | 6 |
| 3 | 1 | 5 | 4 | 1 | 7 | 5 | 5 |
| 4 | 8 | 7 | −1 | 0 | 8 | 6 | 8 |
| | 4 | 6 | 2 | | 7 | 6 | |
| | True Averages | | | | Observed Averages | | |

# Perfect Doctor: Remarks

- According to the doctor, half of the patients do better under the control treatment (standard medical procedure)
- The process by which treatment was assigned carries a lot of information!
  - ✓ The assignment depends on both $Y_i(0)$ and $Y_i(1)$ for each unit (of course, in reality we will unlikely to have any such perfect doctor)
- Key: the assignment mechanism (Rubin, 1978) – the probabilistic rule that decides which unit gets assigned to which treatment
- The key identifying assumptions in causal inference are on the assignment mechanism

# Outline

# Assignment Mechanism

Informally . . .

- The assignment mechanism is the process that determines which units receive which treatments, e.g., which units receive active treatment and which receive control
  - ✓ The assignment mechanism determines potential outcomes are realized and thus can be observed, and, conversely, which potential outcomes are missing

Formally. . .

- The assignment mechanism is a (row-exchangeable) function (taking on value in $[0, 1]$) that assigns probabilities to all $2^N$ possible values for the the $N-$dimensional vector of assignments, $\boldsymbol{W}$, given the science $[\boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)]$:

$$\Pr\left(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right)$$

  such that the probabilities sum to 1:

$$\sum_{\boldsymbol{W} \in \{0,1\}^N} \Pr\left(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right)$$

# Assignment Mechanism: Example

- $N = 2$
- There are $2^2 = 4$ possible values for $\boldsymbol{W}$

$$\boldsymbol{W} = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

- $\Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1))$ such that

$$\Pr\left( \begin{pmatrix} 0 \\ 0 \end{pmatrix} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1) \right) + \Pr\left( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1) \right) +$$
$$+ \Pr\left( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1) \right) + \Pr\left( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1) \right) = 1$$

  - ✓ Example

$$\Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \frac{1}{4} \quad \text{for } \boldsymbol{W} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right.$$

# Unit-Level Assignment Probability

- The unit-level assignment probability for unit $i$ is the sum the probabilities across all possible assignment vectors $\boldsymbol{W}$ for which
  unit $i$ is assigned to the active treatment, $W_i = 1$:

$$\Pr\left(W_i = 1 \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right) = \sum_{\boldsymbol{W}: W_i = 1} \Pr\left(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right)$$

- The probability that unit $i$ is assigned to the control treatment is

$$\Pr\left(W_i = 0 \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right) = 1 - \Pr\left(W_i = 1 \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right)$$

# Restriction on the Assignment Mechanism

- Individualistic assignment: This limits the dependence of a particular unit's assignment probability on the values of covariates and potential outcomes for other units
  - ✓ If the assignment mechanism is individualistic, each unit-level assignment probability depends only on the row of the Science corresponding to that unit

$$\Pr\left(W_i = 1 \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right) = \Pr\left(W_i = 1 \mid \boldsymbol{X}_i, Y_i(0), Y_i(1)\right)$$

- Probabilistic assignment: This requires the assignment mechanism to imply a non-zero probability for each treatment value, for every unit

$$0 < \Pr\left(W_i = 1 \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right) < 1 \qquad \text{for all } i$$

- Unconfounded assignment: This disallows dependence of the assignment mechanism on the potential outcomes

# Unconfounded Assignment Mechanism

- An assignment mechanism is unconfounded if it does not depend on potential outcomes

$$\Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \Pr(\boldsymbol{W} \mid \boldsymbol{X})$$

- The plausibility of unconfoundedness often depends on the covariates collected
  - ✓ An assignment mechanism is confounded or unconfounded given a particular set of covariates, $\boldsymbol{X}$
- Given individualistic assignment, the combination of probabilistic and unconfounded assignment is referred to as strongly ignorable treatment assignment *(Rubin and Rosenbaum, 1983)*

# Individualistic and Unconfounded Assignment Mechanism

- In a individualistic and unconfounded assignment mechanism the unit level assignment probability is

$$\Pr\left(W_i = 1 \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right) = \Pr\left(W_i = 1 \mid \boldsymbol{X}_i\right)$$

- In an individualistic and unconfounded assignment mechanism the unit level assignment probability coincides with the so-called propensity score

$$e(\boldsymbol{X}_i) = \Pr\left(W_i = 1 \mid \boldsymbol{X}_i\right)$$

- An individualistic and unconfounded assignment mechanism is proportional to the product of the propensity scores

$$\Pr\left(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)\right) = c \prod_{i=1}^{N} e(\boldsymbol{X}_i)^{W_i} \left[1 - e(\boldsymbol{X}_i)\right]^{1-W_i}$$

# Ignorable Assignment

- An assignment mechanism is ignorable if the it does not depend on the missing outcomes:

$$p(W_i = 1|X_i, Y_i(0), Y_i(1)) = p(Y_i = 1|X_i, Y_i^{obs})$$

- The terminology is corresponding to the missing data mechanism in the missing data literature (Rubin, 1976).
- The meaning of "ignorability" becomes more apparent when we talk about Bayesian inference of causal effect – meaning the assignment mechanism drops out from the data likelihood in estimating the causal effects
- An unconfounded assignment is always ignorable, but not vice versa (one example is the sequential randomized experiment)
- In most cases in practice, the difference between unconfoundedness and ignorability is negligible, and these two are used exchangeably

# Known and Controlled Assignment Mechanism

- The researcher knows and controls the functional form of the assignment mechanism

- Known and Controlled Assignment Mechanism: the assignment to the treatment for each unit is determined by a coin flip

# A Classification of Assignment Mechanisms

- Randomized experiment: Probabilistic and has a known functional form that is controlled by the researcher
  - ✓ Classical randomized experiment: Randomized experiment with an assignment mechanism that is individualistic, and unconfounded
- Observational study: The functional form of the assignment mechanism is unknown
  - ✓ Regular assignment mechanism: Assignment mechanism that is individualistic, probabilistic and unconfounded
  - ✓ Irregular assignment mechanism: Assignment mechanism where at least one regularity restriction does not hold
    - Instrumental variables, Regression discontinuity designs, Intermediate variables
- Classical randomized experiments are regular assignment mechanisms (individualistic, probabilistic and unconfounded) with a known functional form that is controlled by the researcher

# A Classification of Assignment Mechanisms: Recap

| Restriction on assignment mechanisms | Randomized Experiments | Classical Randomized Experiments | Observational study: Regular AM | Observational study: Irregular AM |
|---|---|---|---|---|
| Individualistic | ? | ✓ | ✓ | ? |
| Probabilistic | ✓ | ✓ | ✓ | ? |
| Unconfounded | ? | ✓ | ✓ | ? |
| Known + controlled | ✓ | ✓ | × | × |

# Role of Randomization

- Randomization does ...
  - ✓ Balance observed covariates: $W_i \perp\!\!\!\perp \boldsymbol{X}_i$
  - ✓ Balance unobserved covariates: $W_i \perp\!\!\!\perp \boldsymbol{U}_i$
  - ✓ Balance potential outcomes, i.e. guarantee unconfoundedness:

$$W_i \perp\!\!\!\perp (Y_i(0), Y_i(1))$$

- Under randomization, average causal effects are (nonparametrically) identified, because we can show

$$\Pr(Y_i(w)) = \Pr(Y_i^{obs} \mid W_i = w), \quad w = 0, 1$$

- Within the potential outcome framework with assumptions, under randomization, association does imply causation

$$\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)] = \mathbb{E}[Y_i^{obs} \mid W_i = 1] - \mathbb{E}[Y_i^{obs} \mid W_i = 0]$$

# Classical Randomized Experiments

- Classical randomized experiment: Assignment mechanism that is individualistic, probabilistic, unconfounded and has a known functional form that is controlled by the researcher:

$$\Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = c \prod_{i=1}^{N} e(\boldsymbol{X}_i)^{W_i} \left[1 - e(\boldsymbol{X}_i)\right]^{1-W_i}$$

where $e(\boldsymbol{X}_i) = \Pr(W_i = 1 \mid \boldsymbol{X}_i)$ is the propensity score

# Types of Classical Randomized Experiments

- Four classical randomized experiments
  1. Bernoulli trials
  2. Completely randomized experiments
  3. Stratified randomized experiments (randomized blocks)
  4. Paired randomized experiments
- The key difference between the four types of classical randomized experiments is in the set of assignment vectors $\boldsymbol{W} = [W_1, \ldots, W_N]$, where $W_i \in \{0, 1\}$, with positive probability
- $\mathbb{W} \subseteq \{0, 1\}^N$: set of assignment vectors $\boldsymbol{W}$ with positive probability
- The purpose of restricting the assignment mechanism is to eliminate vectors that are less desirable for estimating causal effects

# Bernoulli Trials

- A Bernoulli trial is a classical randomized experiment with an assignment mechanism such that the assignments for all units are independent
- Assignment mechanism of a Bernoulli trial:

$$\Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \prod_{i=1}^{N} e(\boldsymbol{X}_i)^{W_i} \left[1 - e(\boldsymbol{X}_i)\right]^{1-W_i}$$

  where $0 < e(\boldsymbol{X}_i) = \Pr(W_i = 1 \mid \boldsymbol{X}_i) < 1$ for all $i$
- The unit assignment probabilities may vary with the units' covariate values (rare)
- When the unit assignment probabilities are constant across units: $e(\boldsymbol{X}_i) = e$ for all $i = 1, \ldots, N$, the assignment to the treatment for each unit is determined by a possibly unfair coin flip
  - ✓ In case of fair coin flips:

$$e(\boldsymbol{X}_i) = e = 0.5 \qquad \text{and} \qquad \Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \left(\frac{1}{2}\right)^N$$

- The number of possible assignments is $2^N$

# Completely Randomized Experiments

- Assignment mechanism of a completely randomized experiment where $N_t$ out of $N$ units are randomly chosen to receive the treatment and $N_c = N - N_t$ i the number of control

$$\Pr(\boldsymbol{W} \mid \boldsymbol{X}, \boldsymbol{Y}(0), \boldsymbol{Y}(1)) = \begin{cases} \binom{N}{N_t}^{-1} = \left[\frac{N!}{N_t!(N-N_t)!}\right]^{-1} & \text{if } \sum_{i=1}^{N} W_i = N_t \\ 0 & \text{Otherwise} \end{cases}$$

- The unit assignment probability (propensity score) is constant across all units: $e(\boldsymbol{X}_i) = \frac{N_t}{N}$ for $i = 1, \ldots, N$
- Simplest case: $N_t = N_c = N/2$ and $e(\boldsymbol{X}_i) = 1/2$
- The number of assignment vectors with positive probability is $\binom{N}{N_t}$

# Stratified Randomized Experiments

- In a stratified randomized experiment, the experimental units are first partitioned into strata (or blocks) that are similar with respect to one or more covariates

  - ✓ The experimental units are partitioned into $J$ strata (or blocks)

  - ✓ Completely randomized experiments are conducted within each block $j$ with $N_{j,t}$ out of $N_j$ units are randomly chosen to receive the treatment, $j = 1, \ldots, J$

  - ✓ Assignments are independent across blocks

- The number of assignment vectors with positive probability is $\prod_{j=1}^{J} \binom{N_j}{N_{j,t}}$

# Pairwise (Paired) Randomized Experiments

- Paired randomized experiments are a special case of stratified trials, with 2 units within each block: one assigned to the active treatment and the other to control
  - ✓ $J = N/2$, $N_j = 2$ and $N_{j,t} = 1$, for $j = 1, \ldots, J = N/2$
- In a paired randomized experiment,
  - ✓ Units are first matched into pairs of similar units
  - ✓ Within each pair, randomize which unit is treated
- The number of assignment vectors with positive probability is $2^{N/2}$

# Classical Randomized Experiments: A comparison

- For a fixed population size, $N$, the assignment mechanisms defining Bernoulli, completely, stratified and pairwise randomized experiments give positive probability to a gradually decreasing number of distinct assignment vectors
- A completely randomized design will tend to be more informative than a Bernoulli trial
  - ✓ An assignment vector with all, or almost all, units assigned to one of the treatment levels is typically not as informative as an assignment vector with more balance between the number of treated and control units
- Stratified/Pairwise randomized experiments may be dramatically beneficial with respect to completely randomized experiments, by improving the efficiency of the design
- The stronger the correlation between potential outcomes and the covariates defining the strata/pairs, the higher the benefits of stratification/pairing are in terms of precision of the inferences for treatment effects

# Classical Randomized Experiments: A comparison

- Pairwise randomized experiments ensure the highest efficiency of the design in terms of expected precision, but raise complications in the analysis phase

  - ✓ Issues in the estimation of the sampling variance of treatment effect estimators

- Pairwise randomized experiments with respect to stratified randomized experiments with two treated and two control units per stratum generally leads to modest improvement in efficiency

- Advise: Opt for a stratified randomized experiment with at least two treated and two control units in each stratum, blocking what one can and randomize what one cannot

# References


Dominici F., Bargagli-Stoffi, FJ., Mealli, F.

Fisher, R. A. (1925). Statistical Methods for Research Workers. UK: Oliver and Boyd.

Fisher, R. A. (1935). The Design of Experiments (1st ed.), UK: Oliver and Boyd.

Hernan, MA and Robins, JM. Causal Inference. (2020). Boca Raton: Chapman & Hall/CRC.

Holland, P. (1986). Statistics and causal inference (with discussion). JASA 81, 945-970

Imbens, G. W. (2020). Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics. J. of Econ. Lit. 58, 4, 1129-79.

Imbens, G. W. and Rubin, D. B. (2015). Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge Univ. Press, New York.

# References


Li, F, Mealli, F. (2014). A conversation with Donald B. Rubin. *Statistical Science*. 29(3), 439-457.

Neyman J. (1990). On the application of probability theory to agricultural experiments: Essay on Principles, Section 9. Masters Thesis. Portions translated into English by D. Dabrowska and T. Speed (1990). Statistical Science. 1923(5): 465-472

Pearl, J. (2009). Causality. Cambridge University Press.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. J. Educ. Psychol. 66 688-701.

Rubin, D. B. (1976). Inference and missing data. Biometrika 63 581-592.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. Annals of Statistics. 6 34-58.

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test comment. JASA 75(371), 591-593.