Asking Good Causal Questions: Positivity and Consistency

Cornell STSCI / INFO / ILRST 3900
Fall 2023
causal3900.github.io

12 Sep 2023

Learning goals for today

At the end of class, you will be able to ask good causal questions.

Good causal questions

- involve treatments that exist (positivity assumption)
- involve precise treatments (consistency assumption)
- with clarity about interference (consistency assumption)

After class:

- Hernán and Robins 2020 Chapter 3
- Optional: Hernán, M. 2016.
  "Does water kill? A call for less casual causal inferences."
  Annals of Epidemiology 26(10):674–680.

Good causal questions involve

treatments that exist

1. Treatments that exist

Employer 1	Employer 2
100 employees	200 employees
Face-to-face interaction	Work in individual offices
75% randomized to vaccine 25% randomized to no vaccine	50% randomized to vaccine 50% randomized to no vaccine

How do you estimate the average effect over all 300 employees?

1. Treatments that exist

Employer 1	Employer 2
100 employees	200 employees
Face-to-face interaction	Work in individual offices
100% randomized to vaccine 0% randomized to no vaccine	50% randomized to vaccine 50% randomized to no vaccine

How do you estimate the average effect over all 300 employees?

1. Treatments that exist

If units are exchangeable given a confounder L,
then to estimate $E(Y^a)$ we need **positivity** to hold

$$\mathrm{P}(A = a \mid \vec{L} = \vec{\ell}) > 0$$

Some treatments simply do not exist in some populations.

Some treatments simply do not exist in some populations.

Source: Wikimedia A, B, C

Some treatments simply do not exist in some populations.

Source: Wikimedia A, B, C

Would the bulbs in Ithaca bloom if it did not freeze all winter?

Some treatments simply do not exist in some populations.

Source: Wikimedia A, B, C

Would the bulbs in Ithaca bloom if it did not freeze all winter?

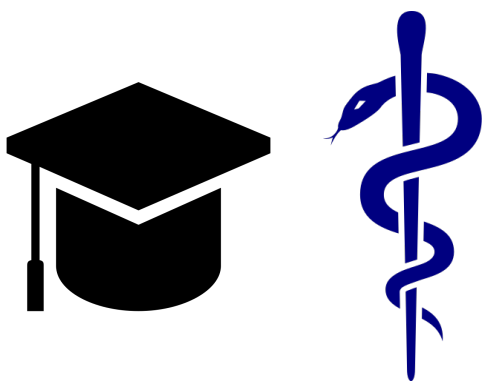
Confounder L	Ithaca
Treatment a	Did not freeze
Outcome Y^a	Blooms?

Some treatments simply do not exist in some populations.

Source: Wikimedia A, B, C

Would the bulbs in Ithaca bloom if it did not freeze all winter?

Confounder L	Ithaca
Treatment a	Did not freeze
Outcome Y^a	Blooms?



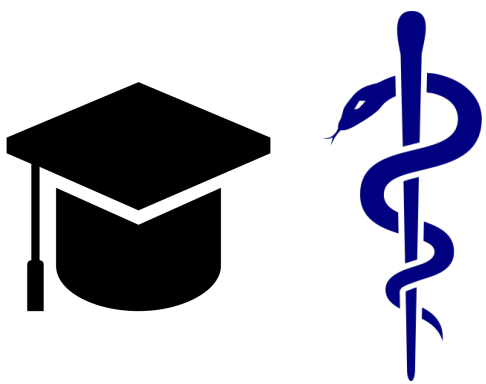
Source: Wikimedia A and B

Some treatments simply do not exist in some populations.

Source: Wikimedia A, B, C

Would the bulbs in Ithaca bloom if it did not freeze all winter?

Confounder L	Ithaca
Treatment a	Did not freeze
Outcome Y^a	Blooms?



Source: Wikimedia A and B

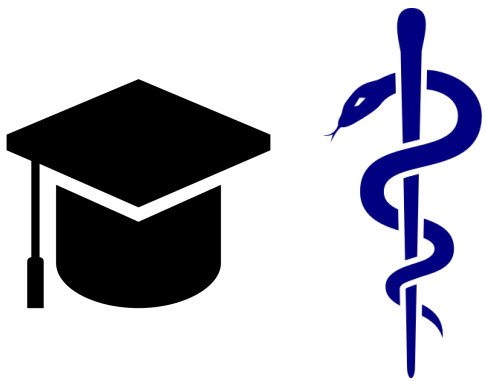
Sarah has no MD training. Would Sarah earn more money if she were a surgeon?

Some treatments simply do not exist in some populations.

Source: Wikimedia A, B, C

Would the bulbs in Ithaca bloom if it did not freeze all winter?

Confounder L	Ithaca
Treatment a	Did not freeze
Outcome Y^a	Blooms?



Source: Wikimedia A and B

Sarah has no MD training. Would Sarah earn more money if she were a surgeon?

Confounder L	No MD training
Treatment a	Surgeon
Outcome Y^a	Earnings

1. Treatments that exist

We can choose causal questions so that positivity holds.

$$\mathrm{P}(A = a \mid \vec{L} = \vec{\ell}) > 0$$

- in each population subgroup $\vec{L} = \vec{\ell}$
- only study treatment values a that can actually happen

Good causal questions involve

precise treatments

Consistency.

$$Y = Y^A$$

1. holds for precise treatments
2. holds with clarity about interference among units

2. Precise treatments

Imagine you are a high school counselor.

A statistician tells you

> The probability of receiving a BA in 6 years would be higher if a student initially enrolled in the State University of New York instead of a community college
>
> $$\mathrm{P}\left(\mathrm{BA}^{\text{Enroll in SUNY}}\right) > \mathrm{P}\left(\mathrm{BA}^{\text{Enroll in Community College}}\right)$$

How would you advise students?

2. Precise treatments

SUNY Educational Opportunity Program

2. Precise treatments

SUNY Educational Opportunity Program

11

6-year graduation rate

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK — 83%

Stony Brook University — 78%

University at Buffalo
The State University of New York — 74%

UNIVERSITY AT ALBANY
State University of New York — 66%

2. Precise treatments

The treatment value
`Enroll in SUNY`
is not sufficiently precise

6-year graduation rate	
Binghamton University, State University of New York	83%
Stony Brook University	78%
University at Buffalo, The State University of New York	74%
University at Albany, State University of New York	66%

2. Precise treatments

The treatment value
`Enroll in SUNY`
is not sufficiently precise

$$\text{BA}^{\text{Binghamton}} \neq \text{BA}^{\text{Stony Brook}} \neq \text{BA}^{\text{Buffalo}} \neq \text{BA}^{\text{Albany}}$$

	6-year graduation rate
Binghamton University, State University of New York	83%
Stony Brook University	78%
University at Buffalo, The State University of New York	74%
University at Albany, State University of New York	66%

2. Precise treatments

The treatment value
`Enroll in SUNY`
is not sufficiently precise

$$\text{BA}^{\text{Binghamton}} \neq \text{BA}^{\text{Stony Brook}} \neq \text{BA}^{\text{Buffalo}} \neq \text{BA}^{\text{Albany}}$$

To advise the student,
a precise treatment
is more helpful

	6-year graduation rate
Binghamton University, State University of New York	83%
Stony Brook University	78%
University at Buffalo, The State University of New York	74%
University at Albany, State University of New York	66%

2. Precise treatments

Consistency assumption: $Y = Y^A$

More credible when A is very precise

- it is clear how to run a hypothetical experiment
- is is clear how to inform policy

Example:

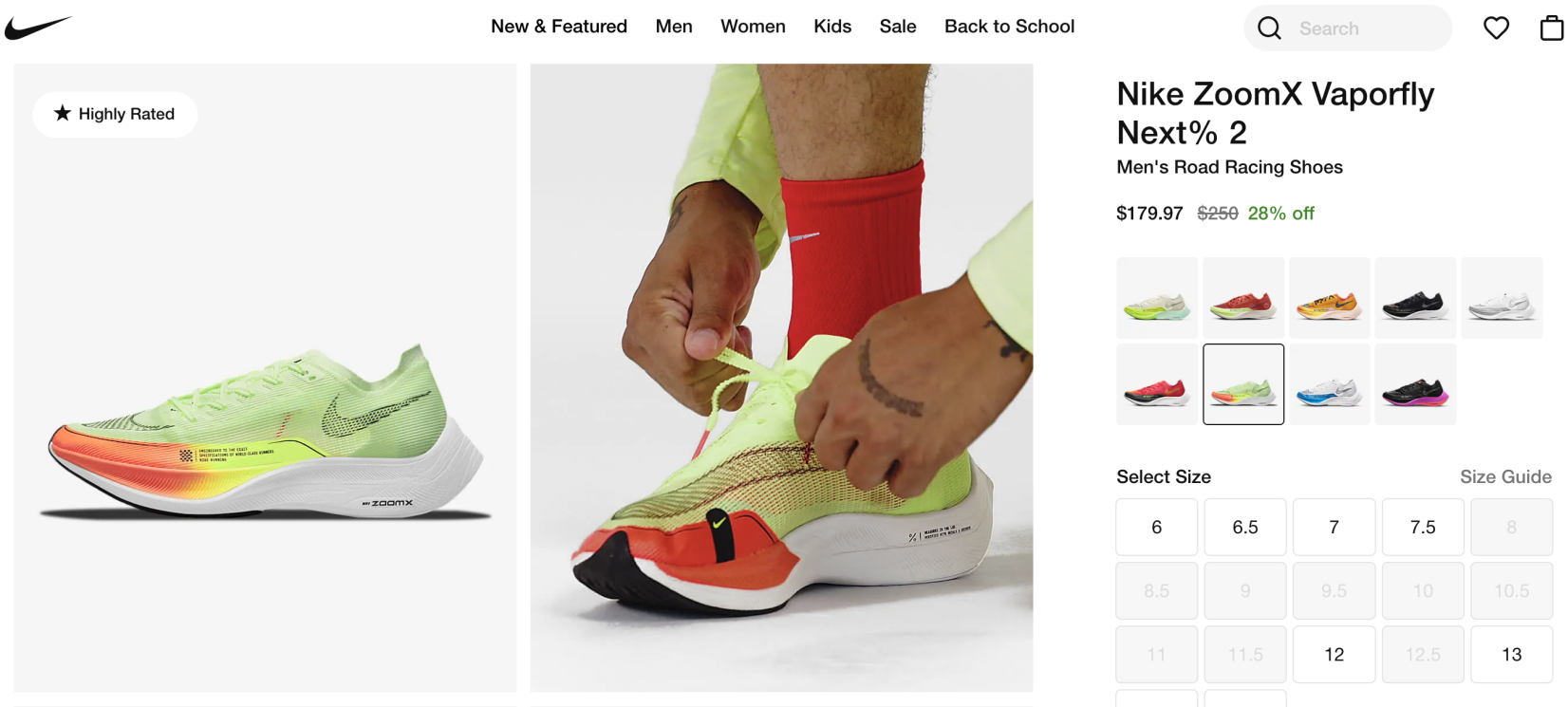
- if $a =$ SUNY, then Y^a is vague.
  To which SUNY should you send the student?
- if $a =$ Binghamton, then Y^a is clearer

A good read:
Hernán, M. 2016.
"Does water kill? A call for less casual causal inferences."
Annals of Epidemiology 26(10):674–680.

Good causal questions involve

clarity about interference

3. With clarity about interference



[1]

[1] Image source: Nike

3. With clarity about interference

You and a friend race in your normal shoes.

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.
You barely lose.

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.
You barely lose.

$$Y_{\text{You}} = \text{Lose}$$

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.
You barely lose.

$$Y_{\text{You}} = \text{Lose}$$

What if you had the springy shoes?

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.
You barely lose.

$$Y_{\text{You}} = \text{Lose}$$

What if you had the springy shoes?

$$Y_{\text{You}}^{\text{You wear springy shoes}} = \text{Win}$$

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.
You barely lose.

$$Y_{\text{You}} = \text{Lose}$$

What if you had the springy shoes?

$$Y_{\text{You}}^{\text{You wear springy shoes}} = \text{Win}$$

But what if your friend also wears them?

3. With clarity about interference

You and a friend race in your normal shoes.
It is extremely close.
You barely lose.

$$Y_{\text{You}} = \text{Lose}$$

What if you had the springy shoes?

$$Y_{\text{You}}^{\text{You wear springy shoes}} = \text{Win}$$

But what if your friend also wears them?

$$Y_{\text{You}}^{\text{You wear springy shoes, Your friend wears springy shoes}} = \text{Lose}$$

$$Y_{\text{You}}^{\text{You wears springy shoes, Your friend wear normal shoes}} = \text{Win}$$

Good causal questions: In math

We should study treatments that exist (positivity)

$$\mathrm{P}(A = a \mid \vec{L} = \vec{\ell}) > 0$$

with potential outcomes that are well-defined (consistency)

$$Y = Y^A$$

Well-defined potential outcomes involve precise treatments

$$\mathrm{BA}^{\text{Binghamton}} \quad \text{instead of} \quad \mathrm{BA}^{\text{SUNY}}$$

and incorporate interference when it exists

$$Y^{a_{\text{you}}, a_{\text{your friend}}} \quad \text{instead of} \quad Y^{a_{\text{you}}}$$

Learning goals for today

At the end of class, you will be able to ask good causal questions.

Good causal questions

- involve treatments that exist (positivity assumption)
- involve precise treatments (consistency assumption)
- with clarity about interference (consistency assumption)

After class:

- Hernán and Robins 2020 Chapter 3
- Optional: Hernán, M. 2016.
  "Does water kill? A call for less casual causal inferences."
  Annals of Epidemiology 26(10):674–680.