Causal Inference

Rachel Gordon STAT 370 Presentation 2

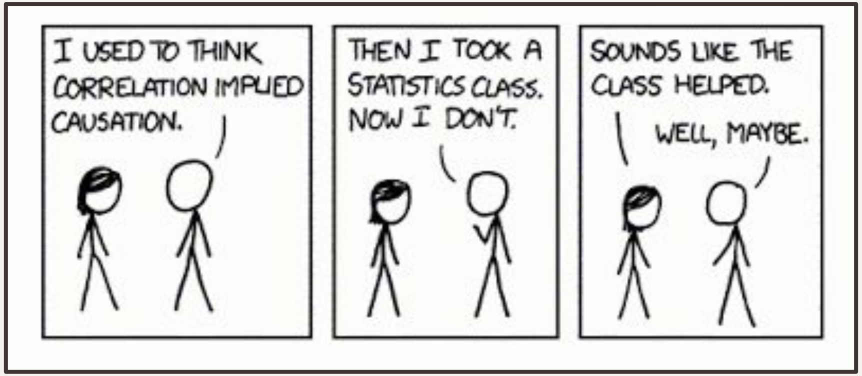
Presentation Outline

01
Background
What is the goal of causal inference?

02
Methods
How do we draw causal conclusions mathematically?

03
Challenges
What issues prevent us from drawing sound conclusions?

04
Big Picture
Why is this concept important?



“Correlation does not imply causation.”

What is Causal Inference?

<ins>Goal</ins>: to answer the question of "why" something happens

Consists of...

- Assumptions
- Study designs
- Estimation strategies

...that allow us to draw causal conclusions from the data

Judea Pearl

- Known for development of Bayesian networks
- Awarded the Turing Award in 2011 for “development of a calculus for probabilistic and causal reasoning”
- Author of several works including *The Book of Why*

“While probabilities encode our beliefs about a static world, causality tells us whether and how probabilities change when the world changes, be it by intervention or by act of imagination.”

Inferring Causation from Different Types of Data

Randomized Experimentation

Intervention in a controlled environment

Observational Studies

No imposed change in the environment

Assumptions of Causal Inference

01

Exchangeability

Treated and untreated individuals are exchangeable

02

Positivity

Probability of receiving every level of treatment is positive for every individual

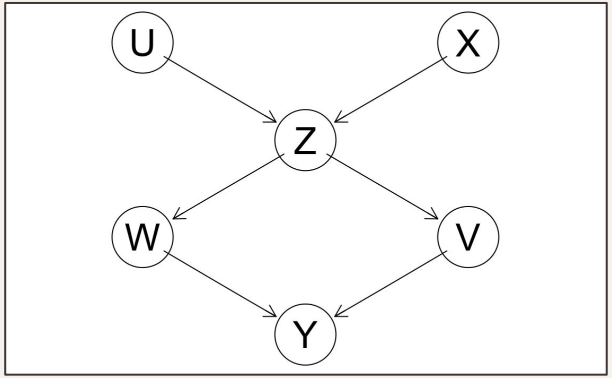
03

Consistency

Potential outcome that corresponds to the treatment that the individual actually received is “factual”

Causal Models & Do-Calculus

- Interventional versus observational conditional probabilities
- Causal model: diagram representing causal relationships within a system or population
    - Involves prior knowledge and assumptions
    - Often visualized using directed acyclic graphs (DAGs)
    - Causal Bayesian Networks
- Three inference rules for probability distributions based on the causal diagram



<u>Rules of Do-Calculus</u>:

1. Insertion/deletion of observations

$$P(Y \mid do(X), Z, W) = P(Y \mid do(X), Z)$$

If W is irrelevant to Y

2. Action/observation exchange

$$P(Y \mid do(X), Z) = P(Y \mid X, Z)$$

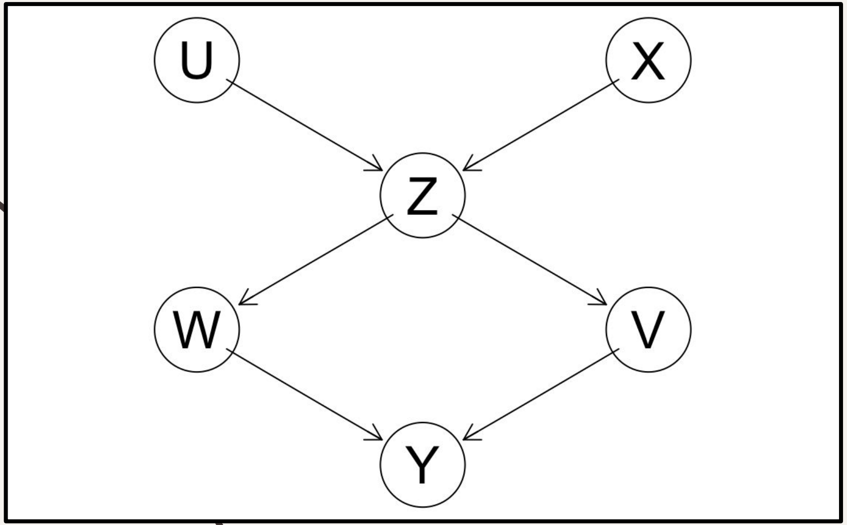
If Z blocks all back-door paths from X to Y

3. Insertion/deletion of actions

$$P(Y \mid do(X)) = P(Y)$$

If there is no causal path from X to Y

d-separation



- Two variables are <u>d-separated</u> by a set of variables if conditioning on all members in this set blocks all paths between the two nodes
- Collider: A node that is influenced by two or more variables
  - Conditioning on a collider unblocks that path
- If two nodes are d-separated given Z we assume they are conditionally independent given Z

Confounding

- Confounders: variables that differ between the treatment and control groups and influence the outcome
- For measured confounders we can make the assumption of conditional exchangeability
    - Look only at individuals who take a single value for the confounding variable
- Counterfactual: a probabilistic answer to a “what would have happened if” question
    - Hypothetical, subjective, untestable, and unfalsifiable

Why Does Causal Inference Matter?

- Solving problems and answering questions about the world
- A missing piece for developing AGI
- Philosophical applications like the notion of epistemic agency and free will

Sources

https://www.inference.vc/causal-inference-3-counterfactuals/

https://ftp.cs.ucla.edu/pub/stat_ser/r402.pdf

https://www.rebeccabarter.com/blog/2017-07-05-confounding/

https://www.inference.vc/untitled/

https://ftp.cs.ucla.edu/pub/stat_ser/r402.pdf

https://causalnex.readthedocs.io/en/latest/04_user_guide/04_user_guide.html#:~:text=Bayesian%20Network%20consists%20of%20a,each%20of%20the%20random%20variables.

https://deepmind.com/blog/article/Causal_Bayesian_Networks

https://www.r-bloggers.com/2019/11/an-introduction-to-causal-inference/

https://www.youtube.com/watch?v=aA-gTNxy1rw

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4912021/

Sources

https://www.sciencedirect.com/topics/social-sciences/causal-inference#:~:text=Causal%20inference%20refers%20to%20an,causal%20conclusions%20based%20on%20data.

https://towardsdatascience.com/implementing-causal-inference-a-key-step-towards-agi-de2cde8ea599

https://towardsdatascience.com/causal-inference-thats-not-a-b-testing-theory-practical-guide-f3c824ac9ed2

https://medium.com/@seema.singh/why-correlation-does-not-imply-causation-5b99790df07e

https://www.cs.ubc.ca/labs/lci/mlrg/slides/doCalc.pdf

https://www.youtube.com/watch?v=pEBI0vF45ic

https://amturing.acm.org/award_winners/pearl_2658896.cfm

https://www.goodreads.com/book/show/36204378-the-book-of-why

https://towardsdatascience.com/conditional-independence-the-backbone-of-bayesian-networks-85710f1b35b

Questions