Causality: Explanation versus Prediction

Department of Government
London School of Economics and Political Science

What did we learn about during MT?

New territory. . .

By the end of today you should be able to:

- Identify what makes for a causal relationship
- Distinguish causation from correlation/association
- Begin to analyse research problems using counterfactual thinking

The broad story arc for LT

- Causal inference!
  - Generating causal theories and expectations
  - Making comparisons
  - Statistical methods useful for causal inference
  - (Quasi-)Experimentation

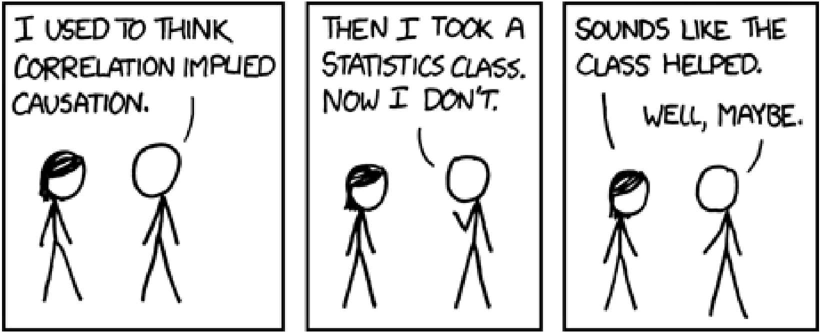
The broad story arc for LT

- Causal inference!
  - Generating causal theories and expectations
  - Making comparisons
  - Statistical methods useful for causal inference
  - (Quasi-)Experimentation

- Developing your research proposals
  - One-on-ones w/ Thomas
  - Literature review
  - Due: 21 March at 5:00pm

Write for 1 minute

What makes something a *cause*?


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Physical causality

- Action and reaction
- Features: Observable and deterministic
- Example:
    - Picture a ball resting on top of a hill
    - What happens if I push the ball?

Physical causality

- Action and reaction
- Features: Observable and deterministic
- Example:
  - Picture a ball resting on top of a hill
  - What happens if I push the ball?
- Physical causality is easy to see

Correlation I

- Correlation is the non-independence of two variables for a set of observations

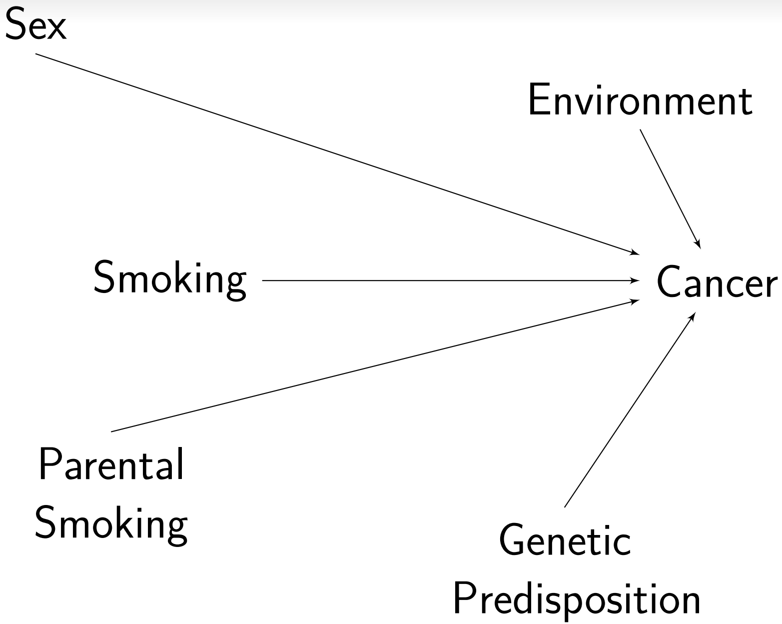
Correlation II

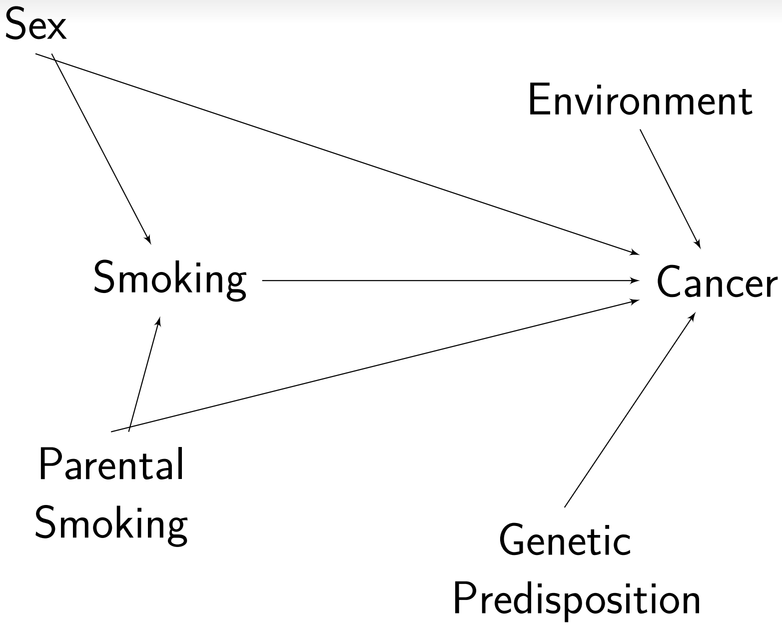
- *Observation*: A case or unit (e.g., person, country)
- *Variable*: A dimension that describes an obseration (e.g., income)
- *Independence*: Variables are unrelated to one another
  - Independent: Height and value on a fair dice roll
  - Non-independent: Height and weight

Correlation III

- Synonyms: correlation, covariation, relationship, association
  - "Effect" is frequently used to mean correlation
  - We'll reserve that term for a *causal effect*
- Any correlation is a potential cause
  - X might cause Y
  - Y might cause X
  - X and Y might be caused by Z
  - X and Y might cause Z
  - There may be no causal relationship

Smoking $\longrightarrow$ Cancer


Sex
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition


Sex
Environment
Smoking
Cancer
Parental
Smoking
Genetic
Predisposition

Causal Graphs I

- Visual representation of (possible) causal relationships
  - Called "directed, acyclic graph" (DAG)

Causal Graphs I

- Visual representation of (possible) causal relationships
  - Called “directed, acyclic graph” (DAG)
- Causality flows between variables, which are represented as “nodes”

Causal Graphs I

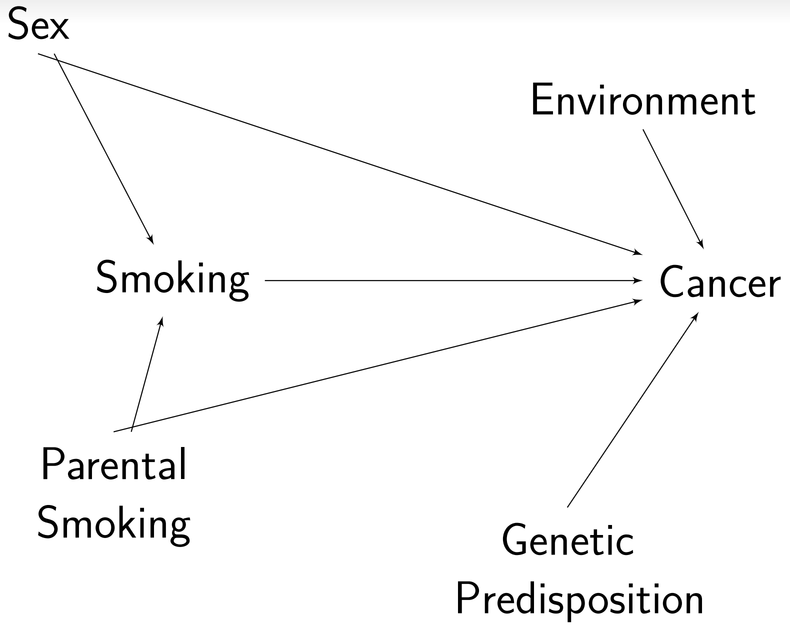
- Visual representation of (possible) causal relationships
  - Called "directed, acyclic graph" (DAG)
- Causality flows between variables, which are represented as "nodes"
- Variables are causally linked by arrows

Causal Graphs I

- Visual representation of (possible) causal relationships
    - Called "directed, acyclic graph" (DAG)
- Causality flows between variables, which are represented as "nodes"
- Variables are causally linked by arrows
- Causality only flows *forward*

Causal Graphs I

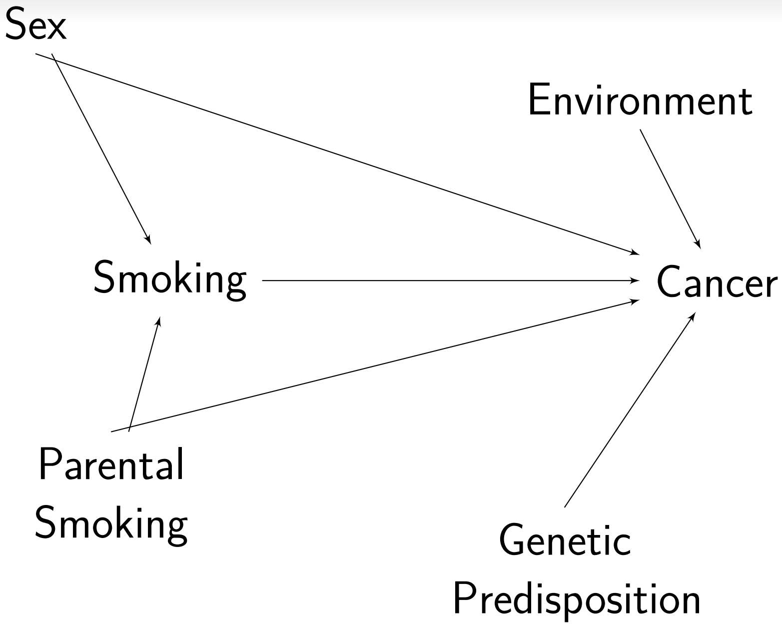
- Visual representation of (possible) causal relationships
  - Called "directed, acyclic graph" (DAG)
- Causality flows between variables, which are represented as "nodes"
- Variables are causally linked by arrows
- Causality only flows *forward*
- Nodes creating a "backdoor path" from X to Y are confounds


Sex
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition

4 (or 5) principles of causality

4 (or 5) principles of causality

1. Correlation

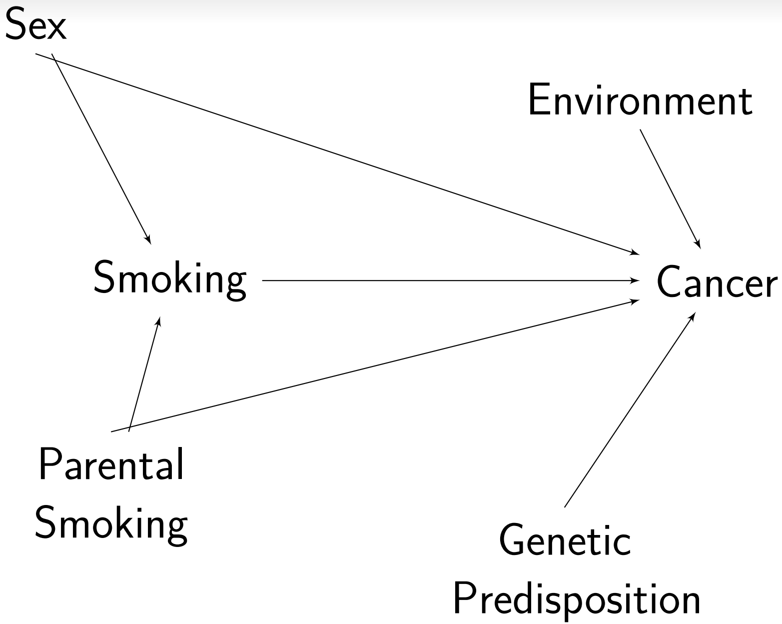

Sex
Environment
Smoking
Cancer
Parental
Smoking
Genetic
Predisposition

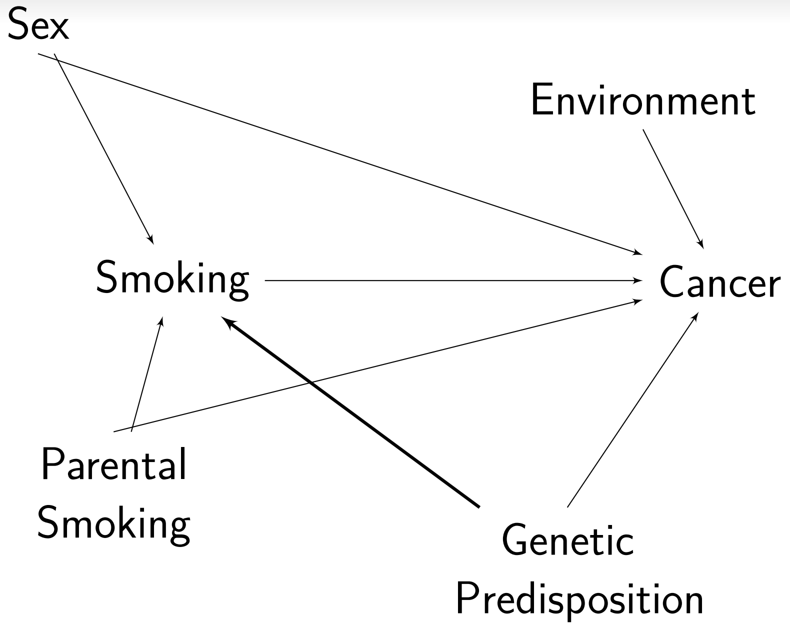
4 (or 5) principles of causality

1. Correlation
2. Nonconfounding

4 (or 5) principles of causality

1. Correlation
2. Nonconfounding
3. Direction (“temporal precedence”)


Sex
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition

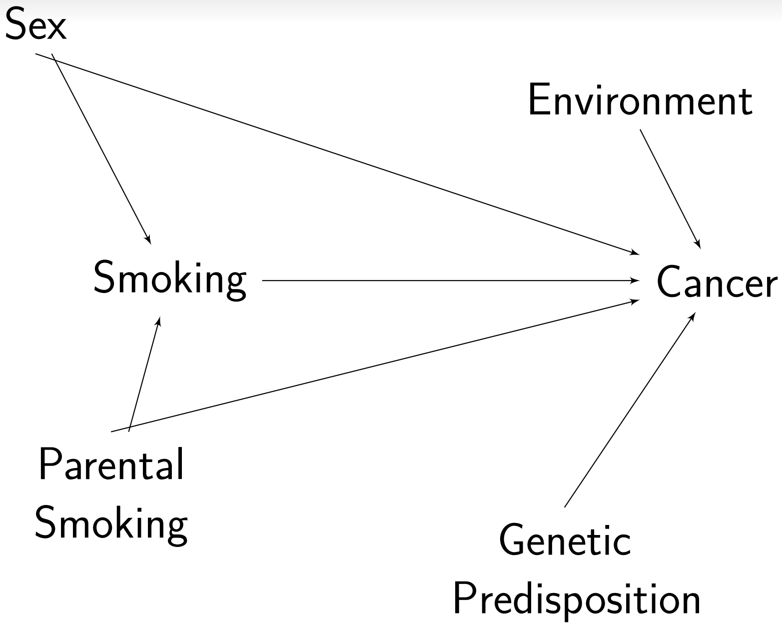

Sex
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition

4 (or 5) principles of causality

1. Correlation
2. Nonconfounding
3. Direction ("temporal precedence")

4 (or 5) principles of causality

1. Correlation
2. Nonconfounding
3. Direction (“temporal precedence”)
4. Mechanism

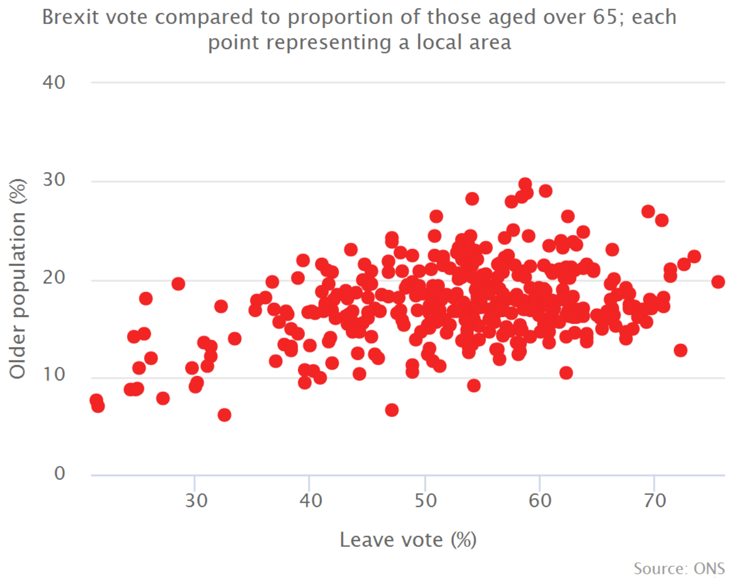

Sex
Environment
Smoking
Cancer
Parental
Smoking
Genetic
Predisposition

4 (or 5) principles of causality

1. Correlation
2. Nonconfounding
3. Direction ("temporal precedence")
4. Mechanism

4 (or 5) principles of causality

1. Correlation
2. Nonconfounding
3. Direction (“temporal precedence”)
4. Mechanism
5. (Appropriate level of analysis)



Source: *The Telegraph*. 27 June 2016. `http://www.telegraph.co.uk/news/2016/06/24/eu-referendum-how-the-results-compare-to-the-uks-educated-old-an/`

Questions?

Counterfactual Thinking

- Causal inference involves inferring *what would have happened* in a counterfactual reality *where the potential cause took on a different value*
- *Counterfactual*: relating to what has not happened or is not the case

"A Christmas Carol"

- 1843 novel by Charles Dickens
- Ebenezer Scrooge is shown his own future by the "Ghost of Christmas Yet to Come"
- Has the choice to either:
    - stay on current path (one counterfactual), or
    - change his ways (take a different counterfactual)

Dickensian Causal Inference

- *Causal effect*: The difference between two "potential outcomes"
  - The outcome that occurs if $X = x_1$
  - The outcome that occurs if $X = x_2$
- The causal effect of Scrooge's lifestyle is seen in the *difference(s)* between two potential futures

Fundamental problem of causal inference

Fundamental problem of causal inference

We can only observe any given unit in one reality!

Two solutions![1]

1. Scientific Solution
   - All units are identical
   - Each can provide a perfect counterfactual
   - Common in, e.g., agriculture, biology

[1] From Holland

Two solutions![1]

1. Scientific Solution
   - All units are identical
   - Each can provide a perfect counterfactual
   - Common in, e.g., agriculture, biology
2. Statistical Solution
   - Units are not identical
   - Random exposure to a potential cause
   - Effects measured on average across units
   - Known as the "Experimental ideal"

[1] From Holland

Mill's methods[2]

- Agreement
- Difference
- Agreement and Difference
- Residue
- Concomitant variations

[2] Discussed in Holland

Mill's Method of Difference

If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance save one in common, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or cause, or an necessary part of the cause, of the phenomenon.

"Rerum cognoscere causas"

- Causal inference is meant to help "explain" the world

"Rerum cognoscere causas"

- Causal inference is meant to help "explain" the world
  - Other notions of explain (e.g., description)

"Rerum cognoscere causas"

- Causal inference is meant to help "explain" the world
  - Other notions of explain (e.g., description)
  - Explanation may or may not involve mechanistic claims (see LT Week 5)

"Rerum cognoscere causas"

- Causal inference is meant to help "explain" the world
  - Other notions of explain (e.g., description)
  - Explanation may or may not involve mechanistic claims (see LT Week 5)
- Causation is deterministic at the unit level!

"Rerum cognoscere causas"

- Causal inference is meant to help "explain" the world
  - Other notions of explain (e.g., description)
  - Explanation may or may not involve mechanistic claims (see LT Week 5)
- Causation is deterministic at the unit level!
- Counterfactual approaches to causal inference are "forward" in nature

Prediction is not causation.
Causation is not prediction.

Prediction is not causation.
Causation is not prediction.

Why are these distinct?

The Experimental Ideal

A randomized experiment, or randomized control trial is:

> *The observation of units after, and possibly before, a randomly assigned intervention in a controlled setting, which tests one or more precise causal expectations*

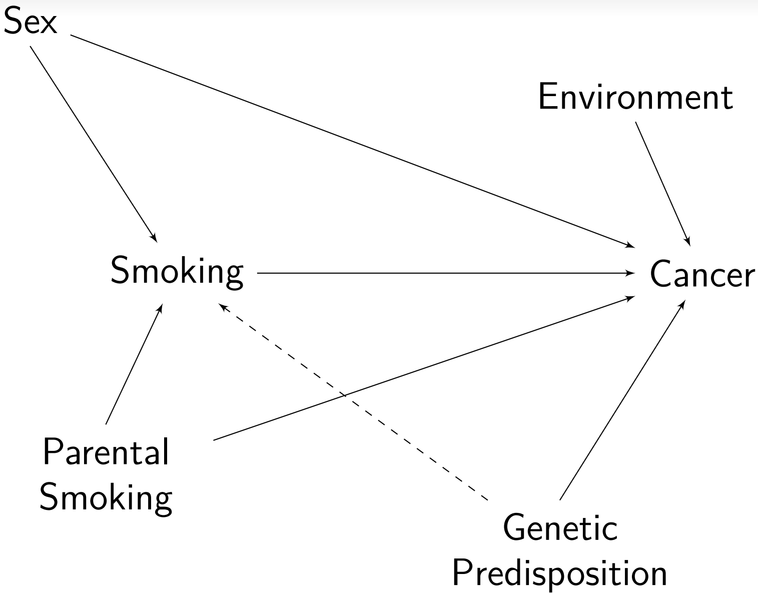
This is Holland's "statistical solution" to the fundamental problem of causal inference

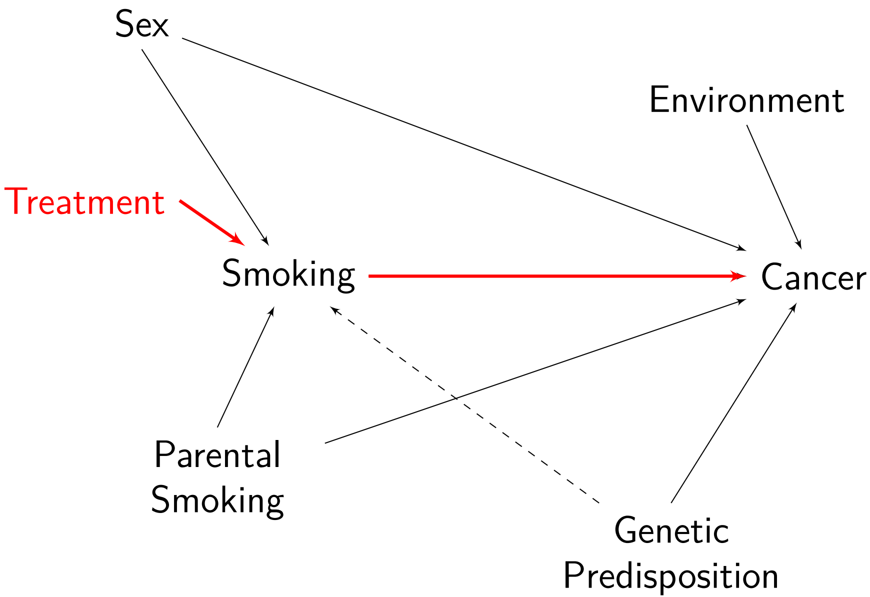
Random Assignment

- A physical process of randomization
  - Breaks the "selection process"
  - Units only take value of $X = x$ because of assignment

Random Assignment

- A physical process of randomization
  - Breaks the "selection process"
  - Units only take value of $X = x$ because of assignment
- This means:
  - Treatment groups, on average, provide in sight into counterfactual "potential" outcomes
  - Randomization means potential outcomes are balanced between groups, so no confounding


Sex
Environment
Smoking
Cancer
Parental Smoking
Genetic Predisposition


Sex
Environment
Treatment
Smoking
Cancer
Parental Smoking
Genetic Predisposition

Experimental Inference I

- Causal inference is a comparison of two *potential outcomes*

Experimental Inference I

- Causal inference is a comparison of two *potential outcomes*
- A potential outcome is the value of the outcome (Y) for a given unit (i) after receiving a particular version of the treatment (X)

Experimental Inference I

- Causal inference is a comparison of two *potential outcomes*
- A potential outcome is the value of the outcome (Y) for a given unit (i) after receiving a particular version of the treatment (X)
- Each unit has multiple *potential* outcomes (Y_{0i}, Y_{1i}), but we only observe one of them

Experimental Inference I

- Causal inference is a comparison of two *potential outcomes*
- A potential outcome is the value of the outcome (Y) for a given unit (i) after receiving a particular version of the treatment (X)
- Each unit has multiple *potential* outcomes (Y_{0i}, Y_{1i}), but we only observe one of them
- A *causal effect* is the difference between these (e.g., $Y_{X=1} - Y_{X=0}$), all else constant

Experimental Inference II

- We cannot see individual-level causal effects
  - We want to know: $TE_i = Y_{1i} - Y_{0i}$

Experimental Inference II

- We cannot see individual-level causal effects
  - We want to know: $TE_i = Y_{1i} - Y_{0i}$
- We can see *average causal effects*
  - Ex.: Average difference in cancer between those who do and do not smoke
  - $ATE_{naive} = E[Y_{1i}|X = 1] - E[Y_{0i}|X = 0]$

Experimental Inference II

- We cannot see individual-level causal effects
    - We want to know: $TE_i = Y_{1i} - Y_{0i}$
- We can see *average causal effects*
    - Ex.: Average difference in cancer between those who do and do not smoke
    - $ATE_{naive} = E[Y_{1i}|X = 1] - E[Y_{0i}|X = 0]$
- Is this what we want to know?

Experimental Inference II

- We cannot see individual-level causal effects
  - We want to know: $TE_i = Y_{1i} - Y_{0i}$
- We can see *average causal effects*
  - Ex.: Average difference in cancer between those who do and do not smoke
  - $ATE_{naive} = E[Y_{1i}|X = 1] - E[Y_{0i}|X = 0]$
- Is this what we want to know?
  - Yes, if X randomized

Experimental Inference II

- We cannot see individual-level causal effects
  - We want to know: $TE_i = Y_{1i} - Y_{0i}$
- We can see *average causal effects*
  - Ex.: Average difference in cancer between those who do and do not smoke
  - $ATE_{naive} = E[Y_{1i}|X = 1] - E[Y_{0i}|X = 0]$
- Is this what we want to know?
  - Yes, if X randomized
  - Yes, if all confounds controlled

Preview of next week

- What is a “scientific literature”?
- How do we accumulate scientific evidence?

Mill's Methods

Agreement

If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon.

Difference

If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur, have every circumstance save one in common, that one occurring only in the former; the circumstance in which alone the two instances differ, is the effect, or cause, or an necessary part of the cause, of the phenomenon.

Agreement and Difference

If two or more instances in which the phenomenon occurs have only one circumstance in common, while two or more instances in which it does not occur have nothing in common save the absence of that circumstance; the circumstance in which alone the two sets of instances differ, is the effect, or cause, or a necessary part of the cause, of the phenomenon.

Residue

Subduct from any phenomenon such part as is known by previous inductions to be the effect of certain antecedents, and the residue of the phenomenon is the effect of the remaining antecedents.

Concomitant variations

Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation.