Technical Track
Session I: Causal Inference

Plamen Nikolov
Sarajevo, Bosnia and Herzegovina
2009
Outline

- Motivation
- Defining Causality
- Causal Inference as a Form of Induction
- Counterfactual Theories
- Rubin Causal Model
- ATE and TOT
- Summary
Motivation

- The research questions that motivate most studies in the health sciences are causal in nature.
- For example:
  - What is the efficacy of a given drug a given population?
  - What fraction of deaths from a given disease could have been avoided by a given treatment or policy?
Motivation

- The most challenging empirical questions in economics also involve causal-effect relationships:
  - Does school decentralization improve schools quality?
  - Does one more year of education causes higher income?
  - Does conditional cash transfers causes better health outcomes in children?
Definitions

- **Merriam-Webster Dictionary**: Something that brings about a result especially a person or thing that is the agent of bringing something about.

- **KJ Rothman**: An event, condition, or characteristic without which the disease would not have occurred.

- **M Susser**: Something that makes a difference.
An intuitive definition of cause

- Ian took the pill on Sept 1, 2003
  - Five days later, he died

- Had Ian not taken the pill on Sept 1, 2003 (all others things being equal)
  - Five days later, he would have been alive

- Did the pill cause Ian’s death?
An intuitive definition of cause

- Jim did not take the pill on Sept 1, 2002
  - Five days later, he was dead.

- Had Jim taken the pill on Sept 1, 2002 (all others things being equal)
  - Five days later, he would have been alive

- Did the pill cause Jim’s survival?
Hill Criteria (1965)

- Strength of Association
- Temporality
- Consistency
- Theoretical Plausibility
- Coherence
- Specificity in the causes
- Dose Response Relationship
- Experimental Evidence
- Analogy
Human reasoning for causal inference

- We compare (often only mentally)
  - the outcome when action $A$ is present with
  - the outcome when action $A$ is absent
  - all other things being equal

- If the two outcomes differ, we say that the action $A$ has a causal effect
  - causative or preventive

- In epidemiology, $A$ is commonly referred to as exposure or treatment
Motivation

Interest in these questions is motivated by:

- Policy concerns
  - Do public programs reduce poverty?
- Theoretical considerations
- Problems facing individual decision makers
Intuition of the problem: a hypothetical example

- Assume we want to estimate the effects of vaccination programs. After some research, we find out that a highly motivated NGO has provided vaccines to all newly born children since 2005 in district A of some developing country. We find out that there is a (baseline) Demographic and Health Survey (DHS) before the intervention and one after: the time line looks as follows:
Vaccines and Infant Mortality

After some online research, you find a brochure of the NGO that states “Due to our intervention, infant mortality has dropped from 122 in 2004 to 97 in 2007”.

Is that convincing? Why or why not?
Vaccines and Infant Mortality

“Simple differences” over time are hard to interpret; it could be that

a. the program actually was causal for the full effect
b. other factors happened in the country and region over time that made the difference
c. the program actually had a negative effect, but was dominated by other things going on

Absolute changes are weak evidence for causal effects -> much more convincing to show that the vaccinated area did better than similar (surrounding areas): Did mortality decline more or less in district A than in other districts over the same time?
Standard Statistical Analysis

- **Tools:** likelihood and other estimation techniques
- **Aim:** to infer parameters of a distribution from samples drawn of that distribution.
- **Uses:** With the help of such parameters, one can:
  - Infer association among variables,
  - Estimate the likelihood of past and future events,
  - Update the likelihood of events in light of new evidence or new measurement.
- **Examples:** correlation, dependence, association, risk ratio, conditional independence.
Causal Analysis

- Demarcation line:
  - A statistical concept is any concept that can be defined in terms of a distribution of observed variables
  - A causal concept is any concept concerning changes in variables that cannot be defined from the distribution alone.

- Causal Analysis goes one step further than Standard Statistical Analysis:
- With the help of such aspects, one can infer
  - the likelihood of events under *static conditions*, (as in Standard Statistical Analysis)
  - and also the dynamics of events under *changing conditions*. 
Causation versus Correlation

- Standard statistical analysis/ probability theory:
  - The word “cause” is not in its vocabulary
  - Allows us to say is that two events are mutually correlated, or dependent (if we find one, we can expect to find the other)

- This is not enough for policy makers
  - They look rationales for policy decisions: if we do XXX, then will we get YYY due to XXX only?
  - Hence we must supplement the language of probability with a vocabulary for causality.
Counterfactual Theory

- Only one of many theories of causality
- It is the most dominant theory however
  - Proposed by Weber (1906) and Lewis (1973)
  - Relies on “what if” comparisons

- In the Lewis Counterfactual Theory two postulates are important for the definition of cause:
  - If “X were to occur, then Y would occur” is true.
  - “If X were not to occur, then Y would not occur either” is true.
Generalizations of counterfactual theory

(a) Causal effects in a subset of the population
   - The difference in probabilities of the outcome under two treatments (e.g. with or without a pill) is positive.

(b) Non dichotomous outcome and exposure
   - Levels of outcome and exposure

(c) Non deterministic counterfactual outcomes
   - Does an exposure cause an outcome (e.g. death or survival) or does it change the probabilities (e.g. 0.9 of death)

(d) Interference
   - Subject's counterfactual outcome does not depend on other subjects' exposure

(e) Time-varying exposures
   - Point v.s. time-sequence of exposure
Counterfactual Theories

- Neyman (1923)
  - Effects of point exposures in randomized experiments

- Rubin (1974)
  - Effects of point exposures in randomized and observational studies

- Robins (1986)
  - Effects of time-varying exposures in randomized and observational studies
The Rubin Causal Model

- Define the population by $U$. Each unit in $U$ is denoted by $u$.

- For each $u \in U$, there is an associated value $Y(u)$ of the variable of interest $Y$, which we call: the response variable.

- Rubin takes the position that causes are only those things that could be treatments in hypothetical experiments.
For simplicity, we assume that there are just two causes or level of treatment.

Let $D$ be a variable that indicates the cause to which each unit in $U$ is exposed:

$$D = \begin{cases} 
1 & \text{if unit } u \text{ is exposed to treatment} \\
0 & \text{if unit } u \text{ is exposed to control}
\end{cases}$$

In a controlled study, $D$ is constructed by the experimenter.

In an uncontrolled study, $D$ is determined by factors beyond the experimenter’s control.
The response $Y$ is potentially affected by whether $u$ receives treatment or not.

Thus, we need two response variables:

$Y_1(u)$ is the outcome if unit $u$ is exposed to treatment

$Y_0(u)$ is the outcome if unit $u$ is exposed to control
\[ D = \begin{cases} 
1 & \text{if unit } u \text{ is exposed to treatment} \\
0 & \text{if unit } u \text{ is exposed to control} 
\end{cases} \]

\( Y_1(u) \) is the outcome if unit \( u \) is exposed to treatment
\( Y_0(u) \) is the outcome if unit \( u \) is exposed to control
\[ \downarrow \]

Then, the outcome of each unit \( u \) can be written as:
\[ Y(u) = D Y_1(u) + (1 - D) Y_0(u) \]

Note: This definition assumes that the treatment status of one unit does not affect the potential outcomes of other units.
Definition: For every unit $u$, treatment causes the effect

$$\delta_u = Y_1(u) - Y_0(u)$$

Fundamental Problem of Causal Inference:
For a given $u$, we observe either $Y_1(u)$ OR $Y_0(u)$
We cannot observe the value of $Y_1(u)$ and $Y_0(u)$ on the same unit $u$
$\Rightarrow$ it is impossible to observe the effect of treatment on $u$ by itself.

Issue: We do not have the counterfactual evidence for $u$
i.e. what would have happened to $u$ in the absence of treatment.
Given that the treatment effect for a single unit $u$ cannot be observed, we aim to identify the average treatment effect for the population $U$ (or for sub-populations).

The average treatment effect $ATE$ over $U$ (or sub-populations of $U$):

$$ TE_u = \delta_u = Y_1(u) - Y_0(u) $$

$$ ATE_U = E_U \ [Y_1(u) - Y_0(u)] $$

$$ = E_U[Y_1(u)] - E_U[Y_0(u)] $$

$$ = \bar{Y}_1 - \bar{Y}_0 $$

$$ = \bar{\delta} \quad (1) $$
The statistical solution replaces the impossible-to-observe treatment effect of $t$ on a specific unit $u$ with the possible-to-estimate average treatment effect of $t$ over a population $U$ of such units.

Although $E_U(Y_1)$ and $E_U(Y_0)$ cannot both be calculated, they can be estimated.

Most econometrics methods attempt to construct from observational data consistent estimates of

$$E_U(Y_1) = \bar{Y}_1 \quad \text{and} \quad E_U(Y_0) = \bar{Y}_0$$
So we are trying to estimate:

\[
ATE_U = E_U[Y_1(u)] - E_U[Y_0(u)]
= \bar{Y}_1 - \bar{Y}_0
\]  

(1)

Consider the following simple estimator of \( ATE_U \):

\[
\hat{\delta} = [\hat{Y}_1 \mid D = 1] - [\hat{Y}_0 \mid D = 0]
\]  

(2)

- equation (1) is defined for the whole population,
- equation (2) is an estimator to be evaluated on a sample drawn from that population
Lemma: If we assume that

\[ [\bar{Y}_1 \mid D = 1] = [\bar{Y}_1 \mid D = 0] \]

and

\[ [\bar{Y}_0 \mid D = 1] = [\bar{Y}_0 \mid D = 0] \]

then

\[ \hat{\delta} = [\hat{Y}_1 \mid D = 1] - [\hat{Y}_0 \mid D = 0] \]

is a consistent estimator of

\[ \bar{\delta} = \bar{Y}_1 - \bar{Y}_0 \]
Thus, a sufficient condition for the simple estimator to consistently estimate the true $ATE$ is that:

\[
[\bar{Y}_1 \mid D = 1] = [\bar{Y}_1 \mid D = 0]
\]

and

\[
[\bar{Y}_0 \mid D = 1] = [\bar{Y}_0 \mid D = 0]
\]

The average outcome under treatment $\bar{Y}_1$ is the same for the treatment ($D=1$) and the control ($D=0$) groups.

The average outcome under control $\bar{Y}_0$ is the same for the treatment ($D=1$) and the control ($D=0$) groups.
When will those conditions be satisfied?

- It is sufficient that treatment assignment $D$ be uncorrelated with the potential outcome distributions of $Y_0$ and $Y_1$.
  - Intuitively: there can be no correlation between
    - Whether someone gets the treatment
    - How much that person potentially benefits from the treatment

- The easiest way to achieve this uncorrelatedness is through random assignment of treatment.
Another way of looking at it

- After a bit of algebra, it can be shown that:

\[
\hat{\delta} = \bar{\delta} + \left( \bar{[Y_0 \mid D = 1]} - \bar{[Y_0 \mid D = 0]} \right) + (1 - \pi) \left( \bar{\delta}_{D=1} - \bar{\delta}_{D=0} \right).
\]

- \(\hat{\delta}\) is the simple estimator of the impact.
- \(\bar{\delta}\) is the true impact.
- \(\bar{[Y_0 \mid D = 1]} - \bar{[Y_0 \mid D = 0]}\) is the Baseline Difference.
- \((1 - \pi) \left( \bar{\delta}_{D=1} - \bar{\delta}_{D=0} \right)\) is the Treatment Heterogeneity.
Another way of looking at it (in words)

- There are two sources of biases that need to be eliminated from estimates of causal effects from observational studies.
  - Baseline difference. (selection bias)
  - Treatment Heterogeneity.

- Most of the methods available only deal with selection bias
Treatment on the Treated

- \( ATE \) is not always the parameter of interest.

- Often, it is the average treatment effect for the treated that is of substantive interest:

\[
TOT = E [Y_1(u) - Y_0(u) | D = 1] = E [Y_1(u) | D = 1] - E [Y_0(u) | D = 1]
\]

- TOT: what is the average effect of the treatment per unit actually receiving it?
Treatment on the Treated

If we need to estimate TOT

\[ TOT = E \left[ Y_1(u) \mid D = 1 \right] - E \left[ Y_0(u) \mid D = 1 \right] \]

Then the simple estimator (2)

\[ \hat{\delta} = \left[ \hat{Y}_1 \mid D = 1 \right] - \left[ \hat{Y}_0 \mid D = 0 \right] \]

consistently estimates TOT if:

\[ \left[ \bar{Y}_0 \mid D = 1 \right] = \left[ \bar{Y}_0 \mid D = 0 \right] \]

“no baseline difference between the treated and control groups”
References