

Causality (UTC 2729)

Wee Yen Zhe

AY23/24 S2

Directed Acyclic Graphs

Directed Acyclic Graphs (DAGs) describe all causal relationships relevant to the effect of D on Y . They provide a graphical representation of a chain of causal effects.

- Arrows (\rightarrow) represent causal effects between two random variables.
- $D \rightarrow Y$ (direct effect).
- $D \rightarrow X \rightarrow Y$ (mediator), capturing a sequence of events originating with D .
- Backdoor path: $D \leftarrow X \rightarrow Y$ (confounder), as X jointly determines D and Y .

If there are no colliders along backdoor paths, then it is an open backdoor path (introduces bias and creates non-causal correlation). To close an open backdoor path:

1. Condition on the confounder.
2. Create a collider along the backdoor path.

No backdoor paths satisfy the backdoor criterion that isolates a causal effect.

- $D \rightarrow X \leftarrow Y$ (Collider) closes a backdoor path.
- Do not control for colliders, as it may open up a backdoor path.

Collider Bias

Collider bias results from bad controls and selection bias. Bad controls occur when the outcome has been a collider linking treatment to the outcome of interest (e.g., $D \rightarrow O \leftarrow A \rightarrow Y$).

5 Methods for Causation

1. Method of agreement.
2. Method of difference (similar to the idea of causation and comparison among counterfactuals).
3. Joint method.
4. Method of concomitant variation.
5. Method of residuals.

Potential Outcomes

The causal effect is defined as a comparison between two states of the world: actual vs. counterfactual; it expresses causality in terms of counterfactuals. Treat i for a unit Y .

- Y^0 : potential outcome of a unit not receiving treatment.
- Y^1 : potential outcome of a unit receiving treatment.

Switching equation:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$$

Where D_i is 1 if the unit receives treatment and 0 if it does not.

- $Y_i = Y_i^1$ if the unit receives treatment.
- $Y_i = Y_i^0$ if the unit does not receive treatment.

Causal effect:

$$Y_i^1 - Y_i^0$$

Average Treatment Effect (ATE)

The expected value of the causal effect $\mathbb{E}[Y_i^1 - Y_i^0]$, can be estimated as:

$$ATE = p \cdot ATT + (1 - p) \cdot ATU$$

- Average Treatment Effect for the Treatment group (ATT): population mean treatment effect for the group of units assigned the treatment in the first place.

$$ATT = \mathbb{E}[\theta_i | D_i = 1] = \mathbb{E}[Y_i^1 | D_i = 1] - \mathbb{E}[Y_i^0 | D_i = 1]$$

- Average Treatment Effect for the Control group (ATU): population mean treatment effect for units assigned to the control group.

$$ATU = \mathbb{E}[\theta_i | D_i = 0] = \mathbb{E}[Y_i^1 | D_i = 0] - \mathbb{E}[Y_i^0 | D_i = 0]$$

Simple difference in means:

$$\mathbb{E}[Y^1 | D = 1] - \mathbb{E}[Y^0 | D = 0]$$

Selection Bias

Selection bias refers to inherent bias between the treated and untreated groups if both groups did not receive treatment.

Heterogeneous Treatment Effect Bias

Heterogeneous treatment effect bias represents the different returns one can obtain from the treatment group for the two groups, multiplied by the share of the population that is in the untreated group.

Independence Assumption

The independence assumption implies that treatment was given randomly to two parties with no regard to potential outcomes for an individual. Check for independence by ensuring the mean potential outcome for both treated and untreated groups is the same for both groups. This can eliminate both selection bias and heterogeneous treatment effect bias.

- Independence: two groups of units have the same potential outcome on average in the population.

$$\mathbb{E}[Y^1|D = 1] - \mathbb{E}[Y^1|D = 0] = 0$$

$$\mathbb{E}[Y^0|D = 1] - \mathbb{E}[Y^0|D = 0] = 0$$

Stable Unit Treatment Value Assumption (SUTVA)

SUTVA assumes that the potential outcomes for each treated unit are the same homogeneous doses.

Randomization Inference

Randomization inference justifies standard errors and appeals to the large sample properties of an estimator. Fisher's sharp null hypothesis states that no unit in the data had a causal effect.

Matching

Conditional Independence Assumption (CIA)

CIA implies that randomization occurred only conditional on observable characteristics. The expected values of Y^1 and Y^0 are equal for the treatment and control groups for each value of X .

- CIA conditions satisfy the backdoor criterion and represent a situation of selection on observables.

- Covariate: a variable that satisfies the backdoor criterion and is usually a random variable assigned to individual units before treatment (must not be a collider).

To estimate a causal effect, need:

1. CIA
2. Probability of treatment to be between 0 and 1 for each stratum (common support)

Example: Whether or not being in first class made someone more likely to survive the Titanic incident. Potential confounder: women and children are more likely to be in first class. Solution: use age and gender to close backdoor paths and satisfy the backdoor criterion.

Curse of Dimensionality

If we condition on too specific a stratum, we may not have the specific examples required (specific age, gender, and class of room).

Matching

Matching estimates the outcome by inputting potential outcomes by conditioning on confounding, observed covariates (filling in missing potential outcomes for each treatment unit using the closest control group unit to the treatment group unit for a confounder variable).

Notion of Exchangeability

A covariate is balanced when the mean of the covariates is the same.

Regression Discontinuity (RD)

RD eliminates selection bias with an assignment variable X , an observable confounder. Estimate potential outcomes change smoothly as a function of the running variable through the cutoff. The only thing that causes the outcome to change abruptly at the cutoff variable is the treatment.

- Causal effects can be identified through units whose scores are in a close neighborhood around some cutoff.
- Compare people above and below the cutoff, then estimate the local average treatment effect as we do not have common support (extrapolation required).
- Running variable: refers to the assignment variable itself (has to be quantitative).

Sharp RD

The probability of treatment goes from 0 to 1 at the cutoff; treatment is deterministic. Sharp RD estimation can be interpreted as the average causal effect of the treatment as the running variable approaches the cutoff in the limit.

Fuzzy RD

The probability of treatment discontinuously increases at the cutoff; treatment is probabilistic.

Local Average Treatment Effect

Local because it is only for running variables near the cutoff. The continuity assumption implies expected outcomes would be continuous even across the cutoff threshold if there had been no treatment (rules out omitted variable bias at the cutoff).

Challenges

Violations to RD include:

1. Assignment rule known in advance.
2. Agents are interested in adjusting.
3. Agents have time to adjust.
4. Cutoff is endogenous to factors that independently cause potential outcomes to shift.
5. Non-random distribution of the running variable.

Instrumental Variables (IV)

Typically used to address omitted variable bias, measurement error, and simultaneity. Given a backdoor path between D and Y , U , which is assumed to be unobserved: $D \leftarrow U \rightarrow Y$ and $Z \rightarrow D \rightarrow Y$.

When Z varies, D varies, which causes Y to change. Y is only varying because D has varied - Z affects Y only through D .

- Hypothetical situation: D consists of people making choices, and sometimes these choices affect Y and sometimes these are merely correlated with changes in Y due to unobserved changes in U . Z then induces a shock to some people in D to make different decisions. Then, we can observe the causal effect of D on Y for this subgroup of people (“compliers”).

Requirements for IV Strategy

1. Stable unit treatment value assumption: potential outcomes for each person are unrelated to treatment status of other individuals.
2. Independence assumption: “as good as random assignment” assumption (IV is independent of potential outcomes and potential treatment assignments).
3. Exclusion restriction: any effect of Z on Y must be via effect of Z on D .
4. First stage: Z has to have a statistically significant effect on average probability of treatment.
5. Monotonicity assumption: IV operates in the same direction on all individual units (guarantees a weighted average of the underlying causal effects of the affected group).

Local Average Treatment Effect (LATE)

The average causal effect of D on Y for those whose treatment status was changed by instrument Z (only measures causal effect of compliers in heterogeneous treatment effects). For homogeneous treatment effects, compliers would have had the same treatment effects as non-compliers.

Popular IV Designs

- Lottery.
- Judge fixed effects design.
- Bartik instruments.

Lottery: randomized trials - whether one was offered treatments.

Differences in Differences (DiD) Design

The idea is to find an instance where a consequential treatment was given to some people but denied to others - also known as a natural experiment because it is based on naturally occurring variation in a treatment variable that affects only some units over time.

- Why not ATE: ATE only works if treatment had been randomized, but choices made are endogenous to potential outcomes (people would choose what is most beneficial for them).
- Works on two pertinent differences: simple before-and-after difference (eliminates unit-specific fixed effects), and then differences the differences from the two groups (hence the term, difference in differences).

Parallel Trends Assumption

DiD only works with the parallel trends assumption - the idea that there are no unobserved factors that could also lead to a change in the outcome, and that the only change is in the intervention to a variable. DiD would simplify to ATT if the expected difference in outcome for both groups (treated and non-treated) over a period of time is zero had both groups not been treated (arisen from selection bias).

Triple Differences

Inclusion of variable-specific time change to the original “generic” intervention change.

Two-way Fixed Effects with Differential Timing

Units receive treatments at different points in time.