Causality

Graphical Causal Models

Alireza Nouri

Sharif University

2024

1. Causality vs. Association
2. Graphical Model for Confounding Bias

Causality vs. Association

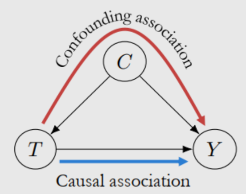
- **Association:**
  - *Definition*: Relationship between two variables, without implying cause.
  - *Example*: Ice cream sales and drowning incidents.

Causality vs. Association

- **Association:**
  - *Definition*: Relationship between two variables, without implying cause.
  - *Example*: Ice cream sales and drowning incidents.
- **Causality:**
  - *Definition*: One event directly influences another.
  - *Example*: Smoking causes lung cancer.

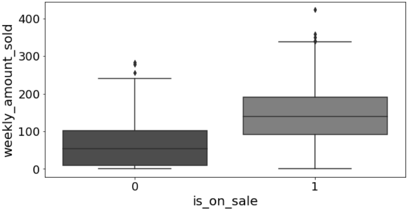
Causality vs. Association

- **Association:**
  - *Definition*: Relationship between two variables, without implying cause.
  - *Example*: Ice cream sales and drowning incidents.
- **Causality:**
  - *Definition*: One event directly influences another.
  - *Example*: Smoking causes lung cancer.
- **Key Difference:**
  - Association identifies a relationship, causality confirms the influence.



Causality vs. Association Example

	store	weeks_to_xmas	avg_week_sales	is_on_sale	weekly_amount_sold
0	1	3	12.98	1	219.60
1	1	2	12.98	1	184.70
2	1	1	12.98	1	145.75
3	1	0	12.98	0	102.45
4	2	3	19.92	0	103.22
5	2	2	19.92	0	53.73



Individual Treatment Effect with the do(.) Operator

Definition: The do(.) operator allows you to express the **individual treatment effect**, which measures the impact of a treatment on the outcome for an individual unit i.

Formal Representation:

$$\tau_i = Y_i \mid \mathrm{do}(T = t_1) - Y_i \mid \mathrm{do}(T = t_0)$$

Understanding the Individual Treatment Effect

Interpretation:

- τ_i represents the effect of switching from treatment t_0 to t_1 for unit i.
- This effect is calculated as the difference in the outcome of unit i under treatment t_1 compared to t_0.

In Words: "The effect, τ_i, of going from treatment t_0 to t_1 for unit i is the difference in the outcome of that unit under t_1 compared to t_0."

Addressing the Fundamental Problem of Causal Inference

Problem Statement:

- The Fundamental Problem of Causal Inference: You can never know the **individual treatment effect** τ_i because only one of the potential outcomes is observable.

Moving Forward:

- Despite this limitation, there are other causal quantities that can be estimated from data.
- One key quantity is the **Average Treatment Effect (ATE)**.

Average Treatment Effect (ATE)

Definition of ATE:

- The ATE is defined as the expected value of the individual treatment effects:
$$\text{ATE} = E[\tau_i],$$
- Alternatively, it can be expressed as:
$$\text{ATE} = E[Y_{1i} - Y_{0i}],$$
- Or using the do(.) operator:
$$\text{ATE} = E[Y \mid \text{do}(T = 1)] - E[Y \mid \text{do}(T = 0)],$$

Interpretation: The ATE represents the average impact of the treatment T across all units.

Estimating the ATE from Data

Estimation Approach:

- Although individual effects τ_i are unknown, the ATE can be estimated using sample averages from the data.
- This involves replacing the expectation with sample means.

Key Insight: While individual impacts vary, the ATE provides a meaningful average effect of the treatment across a population.

Confounding Bias in ATE: Understanding the Pitfalls with an Example

	i	y0	y1	t	x	y	te
0	1	200	220	0	0	200	20
1	2	120	140	0	0	120	20
2	3	300	400	0	1	300	100
3	4	450	500	1	0	500	50
4	5	600	600	1	0	600	0
5	6	600	800	1	1	800	200

The table shows data with unit identifier i, outcome y, potential outcomes y_0 and y_1, treatment indicator t, and covariate x (time until Christmas). Observations are from one week before and during Christmas.

Example Analysis of Confounding Bias

Example Analysis:

- With complete information on AmountSold_0 and AmountSold_1, calculating causal quantities is straightforward.
- The **Average Treatment Effect (ATE)** is:

$$\text{ATE} = \frac{20 + 20 + 100 + 50 + 0 + 200}{6} = 65$$

This indicates an average increase of 65 units in sales due to the treatment.

Data and Real-World Challenges: Part 1

	i	y0	y1	t	x	y	te
0	1	200.0	NaN	0	0	200	NaN
1	2	120.0	NaN	0	0	120	NaN
2	3	300.0	NaN	0	1	300	NaN
3	4	NaN	500.0	1	0	500	NaN
4	5	NaN	600.0	1	0	600	NaN
5	6	NaN	800.0	1	1	800	NaN

Misinterpreting Association as Causation:

- Directly comparing the mean sales of treated vs. untreated groups to estimate the Average Treatment Effect (ATE) can be misleading.

Data and Real-World Challenges: Part 2

Example Calculation:

$$\text{ATE} = \frac{500 + 600 + 800}{3} - \frac{200 + 120 + 300}{3} = 426.67$$

Key Points:

- This method overlooks differences between groups.
- Treated businesses might have inherently higher sales potential.
- Proper causal inference methods are needed to account for these differences and accurately estimate the treatment effect.

Association vs. Causation

Association:

- **Measured by**:

$$E[Y \mid T = 1] - E[Y \mid T = 0]$$

where $E[Y \mid T = 1]$ is the average outcome for treated units and $E[Y \mid T = 0]$ is for untreated units.

Causation:

- **Measured by**:

$$E[Y1 - Y0] \text{ or } E[Y \mid \mathrm{do}(T = 1)] - E[Y \mid \mathrm{do}(T = 0)]$$

which is the difference in potential outcomes under treatment and no treatment.

Key Point: Association may not accurately reflect causation due to inherent differences between treated and control groups.

The Bias Equation

Bias Equation:

$$
\begin{aligned}
E[Y \mid T=1] - E[Y \mid T=0] &= E[Y1 \mid T=1] - E[Y0 \mid T=0] \\
&= \underbrace{E[Y1 - Y0 \mid T=1]}_{\text{ATT}} \\
&\quad + \underbrace{E[Y0 \mid T=1] - E[Y0 \mid T=0]}_{\text{BIAS}}
\end{aligned}
$$

Implications:

- **ATT (Average Treatment Effect on the Treated)**: Reflects the treatment effect for those who received treatment.
- **Bias**: Represents the difference in outcomes between treated and control groups, regardless of treatment.

Introduction to Randomized Experiments

Randomized Experiments:

- **Purpose:** To determine causal relationships between variables.
- **Key Components:**
  - **Random Assignment:** Participants are randomly assigned to treatment or control groups.
  - **Treatment Implementation:** The treatment is administered to the treatment group.
  - **Outcome Measurement:** Outcomes are measured and compared between groups.
- **Objective:** To isolate the effect of the treatment and control for confounding variables.

Measuring Causation

Randomized Experiments and Causation:

- **Control for Confounding:** Random assignment isolates the treatment effect.
- **Average Treatment Effect (ATE):**

$$\mathrm{ATE} = \mathrm{E}[Y1 - Y0] \tag{1}$$

  where:
  - **Y1** = Outcome with treatment
  - **Y0** = Outcome without treatment
- **Example:**
  - **Treatment Group:** Receives a new drug
  - **Control Group:** Receives a placebo
  - Compare average outcomes to measure the effect of the drug.

Example Calculation

Calculation of ATE:

- **Treatment Group Average Reduction:** $\bar{Y}1 = 10$ mmHg
- **Control Group Average Reduction:** $\bar{Y}0 = 5$ mmHg
- **Average Treatment Effect (ATE):**

$$\text{ATE} = \bar{Y}1 - \bar{Y}0 = 10 - 5 = 5 \text{ mmHg} \tag{2}$$

- **Interpretation:** The drug reduces blood pressure by an average of 5 mmHg compared to the placebo.

Weaknesses and Summary

Weaknesses of Randomized Experiments:

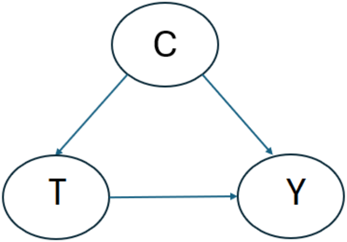
- **Ethical Concerns:** May be unethical to deny treatment.
- **Practical Constraints:** Expensive and logistically challenging.
- **External Validity:** Results may not generalize to other settings.
- **Attrition and Noncompliance:** Dropouts can bias results.
- **Limited Scope:** Often focuses on short-term outcomes.

Summary:

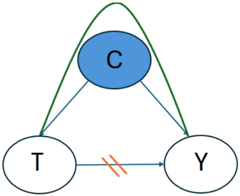
- **Randomized Experiments:** A powerful method for causal inference.
- Controls for confounding factors and provides clear treatment effect measurement.
- **Limitations:** Ethical issues, practical constraints, and potential biases.

1. Causality vs. Association
2. Graphical Model for Confounding Bias

Common Cause Graph



d-separation (Backdoor Path)



- **Mathematical Representation**:

$$T \perp\!\!\!\perp Y \mid C$$

  This notation means that T and Y are conditionally independent given C.

- **Explanation**:
  - In the graph $T \leftarrow C \rightarrow Y$, conditioning on C blocks the path between T and Y.
  - Therefore, T and Y are d-separated by C, which implies:

$$P(T, Y \mid C) = P(T \mid C) \cdot P(Y \mid C)$$

Identifying Causal Effects with CIA

- **Concept**: The Conditional Independence Assumption (CIA) enables us to identify causal effects using observable data.
- **Approach**: If treatment is random within groups defined by X, we can compare treated and untreated groups within each X group.

Adjustment Formula

Formula:

$$\text{ATE} = \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]\right]$$

$$\text{ATE} = \sum_x \{(\mathbb{E}[Y \mid T = 1, X = x] - \mathbb{E}[Y \mid T = 0, X = x])\, P(X = x)\}$$

- **Explanation**: The ATE is the weighted average of differences between treated and control groups within each X group.

Conditionality Principle

- **Key Idea**: By conditioning on X, we can identify the average treatment effect (ATE) as a weighted average of in-group differences.
- **Importance**: Conditioning on X blocks non-causal paths, making causal quantities like ATE identifiable.

Backdoor Adjustment

- **Definition**: The process of closing backdoor paths by adjusting for confounders.
- **Purpose**: This adjustment ensures that the causal effect is correctly identified, removing bias from non-causal associations.

An example from Backdoor Path

	profits_prev_6m	consultancy	profits_next_6m
0	1.0	0	1.0
1	1.0	0	1.1
2	1.0	1	1.2
3	5.0	0	5.5
4	5.0	1	5.7
5	5.0	1	5.7

Graphical Model

Calculation of Average Treatment Effect (ATE)

Given the data:

- $\mathbb{E}[Y \mid T = 1, X = 1.0] = 1.2$
- $\mathbb{E}[Y \mid T = 0, X = 1.0] = 1.05$
- $\mathbb{E}[Y \mid T = 1, X = 5.0] = 5.7$
- $\mathbb{E}[Y \mid T = 0, X = 5.0] = 5.5$

Formula for ATE:

$$\text{ATE} = \frac{1}{2}\left(\mathbb{E}[Y \mid T = 1, X = 1.0] - \mathbb{E}[Y \mid T = 0, X = 1.0]\right)$$

$$+\frac{1}{2}\left(\mathbb{E}[Y \mid T = 1, X = 5.0] - \mathbb{E}[Y \mid T = 0, X = 5.0]\right)$$

$=0.175$

References

- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.
- Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Hernán, M. A., & Robins, J. M. (2020). *Causal Inference: What If*. Chapman & Hall/CRC.
- Rosenbaum, P. R. (2002). *Observational Studies*. Springer.
- Angrist, J. D., & Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.