Causal Models

Brady Neal

causalcourse.com
The Identification-Estimation Flowchart

Causal Estimand → Identification → Statistical Estimand → Estimation → Estimate
The Identification-Estimation Flowchart

- Causal Estimand
- Identification
- Statistical Estimand
- Estimation
- Causal Model
- Data
- Estimate
The Identification-Estimation Flowchart

Causal Estimand → Identification

Statistical Estimand → Estimation

Causal Model

Data
The $do$-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation
The *do*-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation
Conditioning vs. intervening

The \textit{do}-operator
Conditioning vs. intervening

Population
Conditioning vs. intervening

Population

Subpopulations

\[ T = 0 \]

\[ T = 1 \]

The *do*-operator
Conditioning vs. intervening

Population

Subpopulations

Conditioning

The *do*-operator
Conditioning vs. intervening

- Population
- Subpopulations: $T = 0$, $T = 1$
- Conditioning: $T = 1$
- Intervening: $do(T = 1)$
- $T = 0$
- $do(T = 0)$

The *do*-operator
Some notation and terminology

The *do*-operator
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \]
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \]
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t)) \]
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t)) \]

Average treatment effect (ATE):

\[ \mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)] \]
Some notation and terminology

Interventional distributions:
\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t)) \]

Average treatment effect (ATE):
\[ \mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)] \]

Observational \hspace{2cm} Interventional
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangledown P(y \mid do(t)) \]

Average treatment effect (ATE):

\[ \mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)] \]

Observational

\[ P(Y, T, X) \]

Interventional
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t)) \]

Average treatment effect (ATE): 

\[ \mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)] \]

Observational \hspace{1cm} Interventional

\[ P(Y, T, X) \hspace{2cm} P(Y \mid do(T = t)) \]
Some notation and terminology

Interventional distributions:

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t)) \]

Average treatment effect (ATE):

\[ \mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)] \]

Observational

\[ P(Y, T, X) \]

\[ P(Y \mid T = t) \]

Interventional

\[ P(Y \mid do(T = t)) \]
Some notation and terminology

**Interventional distributions:**

\[ P(Y(t) = y) \triangleq P(Y = y \mid do(T = t)) \triangleq P(y \mid do(t)) \]

**Average treatment effect (ATE):**

\[ \mathbb{E}[Y \mid do(T = 1)] - \mathbb{E}[Y \mid do(T = 0)] \]

**Observational**

\[ P(Y, T, X) \]

\[ P(Y \mid T = t) \]

**Interventional**

\[ P(Y \mid do(T = t)) \]

\[ P(Y \mid do(T = t), X = x) \]

The *do*-operator
Identifiability

The \textit{do}-operator

\begin{center}
\begin{tikzpicture}
    \node (causal_estimand) at (0, 0) {Causal Estimand};
    \node (statistical_estimand) at (0, -2) {Statistical Estimand};
    \node (causal_model) at (2, 0) {Causal Model};
    \path[->] (causal_estimand) edge node {Identification} (statistical_estimand);
    \path[->] (statistical_estimand) edge (causal_model);
\end{tikzpicture}
\end{center}
Identifiability

\[ P(y \mid do(t)) \]

Causal Estimand

Identification

Statistical Estimand

Causal Model

The \textit{do}-operator
Identifiability

\[ P(y \mid do(t)) \]

Causal Estimand

\[ P(y \mid t) \]

Statistical Estimand

Identification

Causal Model

The \textit{do}-operator
Identifiability

Accessible via experiment

\[ P(y \mid do(t)) \]

Causal Estimand

Identification

\[ P(y \mid t) \]

Statistical Estimand

Causal Model

The \textit{do}-operator
Identifiability

Accessible via experiment

\[ P(y \mid do(t)) \]

Causal Estimand

Identification

\[ P(y \mid t) \]

Statistical Estimand

Accessible via observational data

Causal Model

The \textit{do}-operator
Identifiability

Accessible via experiment

\[ P(y \mid do(t)) \]

\[ P(y \mid t) \]

Accessible via observational data

Causal Estimand

\[ P(y \mid do(t)) \]

Identification

Causal Model

No confounding

Statistical Estimand

Accessible via observational data

The do-operator
Identifiability

Accessible via experiment

\[ P(y \mid do(t)) \]

Causal Estimand

Identification

\[ P(y \mid t) \]

Statistical Estimand

Accessible via observational data

Causal Model

confounders X

The do-operator
Identifiability

Accessible via experiment

$P(y \mid do(t))$

Causal Estimand

Identification

Statistical Estimand

Accessible via observational data

$E_X[P(y \mid t, X)]$

Causal Model

confounders $X$

The $do$-operator
The *do*-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation
Causal mechanism

\[ P(x_i \mid \text{pa}_i) \]

Main assumption: modularity
Causal mechanism

$$P(x_i | \text{pa}_i)$$

Main assumption: modularity
Modularity assumption
Modularity assumption

If we intervene on a node $X_i$, then only the mechanism $P(x_i \mid pa_i)$ changes. All other mechanisms $P(x_j \mid pa_j)$ where $i \neq j$ remain unchanged.
Modularity assumption

If we intervene on a node $X_i$, then only the mechanism $P(x_i \mid \text{pa}_i)$ changes. All other mechanisms $P(x_j \mid \text{pa}_j)$ where $i \neq j$ remain unchanged.

In other words, the causal mechanisms are modular.
Modularity assumption

If we intervene on a node $X_i$, then only the mechanism $P(x_i \mid \text{pa}_i)$ changes. All other mechanisms $P(x_j \mid \text{pa}_j)$ where $i \neq j$ remain unchanged.

In other words, the causal mechanisms are **modular**.

Many names: independent mechanisms, autonomy, invariance, etc.
Modularity assumption: more formal
Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all $i$, we have the following:
Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all $i$, we have the following:

1. If $i \not\in S$, then $P(x_i \mid \text{pa}_i)$ remains unchanged.
Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all $i$, we have the following:

1. If $i \not\in S$, then $P(x_i \mid \text{pa}_i)$ remains unchanged.

2. If $i \in S$, then $P(x_i \mid \text{pa}_i) = 1$ if $x_i$ is the value that $X_i$ was set to by the intervention; otherwise, $P(x_i \mid \text{pa}_i) = 0$. 
Modularity assumption: more formal

If we intervene on a set of nodes $S \subseteq [n]$, setting them to constants, then for all $i$, we have the following:

1. If $i \not\in S$, then $P(x_i \mid pa_i)$ remains unchanged.

2. If $i \in S$, then $P(x_i \mid pa_i) = 1$ if $x_i$ is the value that $X_i$ was set to by the intervention; otherwise, $P(x_i \mid pa_i) = 0$. consistent with the intervention
Manipulated graphs

Observational data

Main assumption: modularity
Manipulated graphs

Main assumption: modularity
Manipulated graphs

Observational data

Interventional data

Main assumption: modularity
Manipulated graphs

Observational data

Interventional data

Main assumption: modularity
Manipulated graphs

Main assumption: modularity
Manipulated graphs

Main assumption: modularity
Manipulated graphs

Main assumption: modularity
What would it mean if modularity is violated?

Main assumption: modularity
What would it mean if modularity is violated?

Intervention on $T$ not only changes $P(T \mid \text{pa}(T))$
What would it mean if modularity is violated?

Intervention on $T$ not only changes $P(T \mid \text{pa}(T))$

but also changes other mechanisms such as $P(T_2 \mid \text{pa}(T_2))$
Truncated factorization

Recall the Bayesian network factorization:

\[ P(x_1, \ldots, x_n) = \prod_i P(x_i \mid \text{pa}_i) \]
Truncated factorization:

$$P(x_1, \ldots, x_n \mid do(S = s)) = \prod_i P(x_i \mid pa_i)$$
Truncated factorization:

\[ P(x_1, \ldots, x_n \mid do(S = s)) = \prod_{i \not\in S} P(x_i \mid pa_i) \]
Truncated factorization:

\[ P(x_1, \ldots, x_n \mid do(S = s)) = \prod_{i \not\in S} P(x_i \mid pa_i) \]

if \( x \) is consistent with the intervention.
Truncated factorization:

\[ P(x_1, \ldots, x_n \mid do(S = s)) = \prod_{i \notin S} P(x_i \mid \text{pa}_i) \]

if \( x \) is consistent with the intervention.

Otherwise,

\[ P(x_1, \ldots, x_n \mid do(S = s)) = 0 \]
Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$

Main assumption: modularity
Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$

Bayesian network factorization: $P(y, t, x) = P(x) \cdot P(t \mid x) \cdot P(y \mid t, x)$
Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$

Bayesian network factorization: $P(y, t, x) = P(x) P(t \mid x) P(y \mid t, x)$

Truncated factorization: $P(y, x \mid do(t)) = P(x) P(y \mid t, x)$

Main assumption: modularity
Simple identification via truncated factorization

Goal: identify $P(y \mid do(t))$

Bayesian network factorization: $P(y, t, x) = P(x) P(t \mid x) P(y \mid t, x)$

Truncated factorization: $P(y, x \mid do(t)) = P(x) P(y \mid t, x)$

Marginalize: $P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x)$

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x) \]

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x) \]

\[ P(y \mid do(t)) \neq P(y \mid t) \]

Main assumption: modularity
Association vs. causation revisited

\[
P(y \mid \text{do}(t)) = \sum_x P(y \mid t, x) P(x)
\]

\[
P(y \mid \text{do}(t)) \neq P(y \mid t)
\]

\[
\sum_x P(y \mid t, x) P(x)
\]

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x) \]

\[ P(y \mid do(t)) \neq P(y \mid t) \]

\[ \sum_x P(y \mid t, x) P(x \mid t) \]
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x) \]

\[ P(y \mid do(t)) \neq P(y \mid t) \]

\[ \sum_x P(y \mid t, x) P(x \mid t) = \sum_x P(y, x \mid t) \]

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x) \]

\[ P(y \mid do(t)) \neq P(y \mid t) \]

\[ \sum_x P(y \mid t, x) P(x \mid t) = \sum_x P(y, x \mid t) = P(y \mid t) \]

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x) \]

\[ P(y \mid do(t)) \neq P(y \mid t) \]

\[ \sum_x P(y \mid t, x) P(x \mid t) = \sum_x P(y, x \mid t) = P(y \mid t) \]

Main assumption: modularity
Association vs. causation revisited

$$P(y \mid do(t)) = \sum_x P(y \mid t, x) P(x)$$

$$P(y \mid do(t)) \neq P(y \mid t)$$

$$\sum_x P(y \mid t, x) P(x \mid t) = \sum_x P(y, x \mid t) = P(y \mid t)$$

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid \text{do}(t)) = \sum_x P(y \mid t, x) P(x) \]

\[ P(y \mid \text{do}(t)) \neq P(y \mid t) \]

\[ \sum_x P(y \mid t, x) P(x \mid t) = \sum_x P(y, x \mid t) = P(y \mid t) \]

Main assumption: modularity
Association vs. causation revisited

\[ P(y \mid do(t)) = \sum_x P(y \mid t, x) \, P(x) \]

\[ P(y \mid do(t)) \neq P(y \mid t) \]

\[ \sum_x P(y \mid t, x) \, P(x \mid t) = \sum_x P(y, x \mid t) \]

\[ = P(y \mid t) \]

Main assumption: modularity
The \textit{do}-operator

Main assumption: modularity

\textbf{Backdoor adjustment}

Structural causal models

A complete example with estimation
Blocking backdoor paths

\[ P(Y \mid t) \]

\[ W_2 \]

\[ W_1 \]

\[ C \]

\[ W_3 \]

\[ T \]

\[ M \]

\[ Y \]
Blocking backdoor paths

\[ P(Y \mid t) \]

\[ W_2 \]

\[ W_1 \]

\[ C \]

\[ W_3 \]

\[ T \]

\[ M \]

\[ Y \]

Causal association
Blocking backdoor paths

\[ P(Y \mid t) \]

Causal association
Causal association

$P(Y \mid t)$

Backdoor adjustment
Blocking backdoor paths

\[ P(Y \mid t) \]

\[ \begin{align*}
& \text{Causal association} \\
& \text{Backdoor adjustment}
\end{align*} \]
Blocking backdoor paths

\[ P(Y \mid t) \]

- Backdoor adjustment

\[ P(Y \mid \text{do}(t)) \]

- Causal association
Causal association

Backdoor adjustment
Blocking backdoor paths

\[ P(Y \mid t, c, w_1) \]

\[ P(Y \mid do(t)) \]

Causal association

Backdoor adjustment
Causal association

\[ P(Y \mid t, c, w_1, w_3) \]

\[ P(Y \mid do(t)) \]
Backdoor criterion and backdoor adjustment
Backdoor criterion and backdoor adjustment

A set of variables $W$ satisfies the backdoor criterion relative to $T$ and $Y$ if the following are true:
Backdoor criterion and backdoor adjustment

A set of variables $W$ satisfies the backdoor criterion relative to $T$ and $Y$ if the following are true:

1. $W$ blocks all backdoor paths from $T$ to $Y$
2. 
Backdoor criterion and backdoor adjustment

A set of variables $W$ satisfies the backdoor criterion relative to $T$ and $Y$ if the following are true:

1. $W$ blocks all backdoor paths from $T$ to $Y$
2. $W$ does not contain any descendants of $T$
Backdoor criterion and backdoor adjustment

A set of variables $W$ satisfies the backdoor criterion relative to $T$ and $Y$ if the following are true:

1. $W$ blocks all backdoor paths from $T$ to $Y$
2. $W$ does not contain any descendants of $T$

Given the modularity assumption and that $W$ satisfies the backdoor criterion, we can identify the causal effect of $T$ on $Y$:

$$P(y \mid do(t)) = \sum_w P(y \mid t, w) P(w)$$
Proof of backdoor adjustment

\[ P(y \mid do(t)) = \sum_w P(y \mid t, w) P(w) \]
Proof of backdoor adjustment

\[ P(y \mid do(t)) \]

Example graph:
Proof of backdoor adjustment

\[ P(y \mid do(t)) = \sum_{w} P(y \mid do(t), w) P(w \mid do(t)) \]

Example graph:

![Diagram showing variables T, W, Y and their relationships with arrows]

Backdoor adjustment
Proof of backdoor adjustment

\[ P(y \mid do(t)) = \sum_w P(y \mid do(t), w) P(w \mid do(t)) \]

\[ = \sum_w P(y \mid t, w) P(w \mid do(t)) \]

Example graph:
Proof of backdoor adjustment

\[ P(y \mid \text{do}(t)) = \sum_w P(y \mid \text{do}(t), w) P(w \mid \text{do}(t)) \]

\[ = \sum_w P(y \mid t, w) P(w \mid \text{do}(t)) \]

\[ = \sum_w P(y \mid t, w) P(w) \]

Example graph:
Backdoor criterion as d-separation

Backdoor adjustment
Backdoor criterion as $d$-separation

1. Blocks all backdoor paths from $T$ to $Y$.

2. 

---

**Backdoor adjustment**

Brady Neal
Backdoor criterion as d-separation

1. Blocks all backdoor paths from $T$ to $Y$.
2. 

\[ G \]

\begin{itemize}
  \item $W$ blocks all backdoor paths from $T$ to $Y$.
  \item $M_1$, $M_2$, $W_1$, $W_2$, $W_3$, $X_1$, $X_2$, $X_3$, and $Y$ are nodes in the graph.
  \item There are causal associations from $T$ to $M_1$, $M_1$ to $M_2$, $M_2$ to $Y$, and non-causal associations from $X_1$ to $T$, $X_2$ to $M_1$, and $X_3$ to $Y$.
\end{itemize}
Backdoor criterion as d-separation

1. \( W \) blocks all backdoor paths from \( T \) to \( Y \)
2. \( W \) does not contain any descendants of \( T \)
Backdoor criterion as d-separation

1. Blocks all backdoor paths from $T$ to $Y$
2. Does not contain any descendants of $T$
Backdoor criterion as d-separation

1. Blocks all backdoor paths from $T$ to $Y$
2. Does not contain any descendants of $T$
Backdoor criterion as d-separation

1. Blocks all backdoor paths from $T$ to $Y$
2. Does not contain any descendants of $T$

$$Y \perp_{G_T} T \mid W$$
Question:
How does this backdoor adjustment relate to the adjustment formula we saw in the potential outcomes lecture?

Backdoor adjustment:
\[ P(y \mid do(t)) = \sum_w P(y \mid t, w) P(w) \]

Adjustment formula from before:
\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_W [\mathbb{E}[Y \mid T = 1, W] - \mathbb{E}[Y \mid T = 0, W]] \]
Question: How does this backdoor adjustment relate to the adjustment formula we saw in the potential outcomes lecture?

Backdoor adjustment:

\[ P(y \mid do(t)) = \sum_w P(y \mid t, w) P(w) \]

Adjustment formula from before:

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_W [\mathbb{E}[Y \mid T = 1, W] - \mathbb{E}[Y \mid T = 0, W]] \]

Section 4.4.1 of the ICI book
The *do*-operator

Main assumption: modularity

Backdoor adjustment

**Structural causal models**

A complete example with estimation
Structural equations
Structural equations

The equals sign does not convey any causal information.
Structural equations

The equals sign does not convey any causal information.

\[ B = A \]  means the same thing as  \[ A = B \]
Structural equations

The equals sign does not convey any causal information.  

\[ B = A \quad \text{means the same thing as} \quad A = B \]

Structural equation for A as a cause of B:  

\[ B := f(A) \]
Structural equations

The equals sign does not convey any causal information.

$$B = A \quad \text{means the same thing as} \quad A = B$$

Structural equation for A as a cause of B:

$$B := f(A)$$

$$B := f(A, U)$$
Structural equations

The equals sign does not convey any causal information.

\[ B = A \quad \text{means the same thing as} \quad A = B \]

Structural equation for A as a cause of B:

\[ B := f(A) \]

\[ B := f(A, U) \]
Causal mechanisms and direct causes revisited

\[ P(x_i \mid \text{pa}_i) \]

Structural causal models
Causal mechanisms and direct causes revisited

Causal mechanism for $X_i$

$X_i := f(A, B, \ldots)$
Causal mechanisms and direct causes revisited

Causal mechanism for $X_i$

$X_i := f(A, B, \ldots)$

Direct causes of $X_i$
Structural causal models (SCMs)

\[ B := f_B(A, U_B) \]
\[ M := f_C(A, B, U_C) \]
\[ D := f_D(A, C, U_D) \]
Structural causal models (SCMs)

\[ B := f_B(A, U_B) \]

\[ M : C := f_C(A, B, U_C) \]

\[ D := f_D(A, C, U_D) \]
Structural causal models (SCMs)

\[ B := f_B(A, U_B) \]
\[ M := f_C(A, B, U_C) \]
\[ D := f_D(A, C, U_D) \]
Structural causal models (SCMs)

\[ B := f_B(A, U_B) \]
\[ M := C := f_C(A, B, U_C) \]
\[ D := f_D(A, C, U_D) \]
Structural causal models (SCMs)

\[
B := f_B(A, U_B) \\
M := f_C(A, B, U_C) \\
D := f_D(A, C, U_D)
\]

**SCM Definition**

A tuple of the following sets:

1. A set of endogenous variables
2. A set of exogenous variables
3. A set of functions, one to generate each endogenous variable as a function of the other variables

Exogenous variables

Endogenous variables
Interventions

SCM (model)

\[ M : \]

\[ T := f_T(X, U_T) \]
\[ Y := f_Y(X, T, U_Y) \]
Interventions

SCM (model)

\[ M : \]
\[ T := f_T(X, U_T) \]
\[ Y := f_Y(X, T, U_Y) \]

Interventional SCM (submodel)

\[ M_t : \]
\[ T := t \]
\[ Y := f_Y(X, T, U_Y) \]
Interventions

SCM (model)

\[ M : \]
\[ T := f_T(X, U_T) \]
\[ Y := f_Y(X, T, U_Y) \]

Interventional SCM (submodel)

\[ M_t : \]
\[ T := t \]
\[ Y := f_Y(X, T, U_Y) \]
Modularity assumption for SCMs

Consider an SCM $M$ and an interventional SCM $M_t$ that we get by performing the intervention $do(T = t)$. The modularity assumption states that $M$ and $M_t$ share all of their structural equations except the structural equation for $T$, which is $T := t$ in $M_t$. 
Modularity assumption for SCMs

Consider an SCM $M$ and an interventional SCM $M_t$ that we get by performing the intervention $do(T = t)$. The modularity assumption states that $M$ and $M_t$ share all of their structural equations except the structural equation for $T$, which is $T := t$ in $M_t$.

\[
M : \quad T := f_T(X, U_T) \\
Y := f_Y(X, T, U_Y)
\]

\[
M_t : \quad T := t \\
Y := f_Y(X, T, U_Y)
\]
Why not condition on descendants of treatment: blocking causal association
Why not condition on descendants of treatment: blocking causal association
Why not condition on descendants of treatment: inducing new post-treatment association

Collider bias
Why not condition on descendants of treatment: inducing new post-treatment association

Collider bias
Why not condition on descendants of treatment: inducing new post-treatment association

Collider bias
Why not condition on descendants of treatment: inducing new post-treatment association

Collider bias

Rule: don’t condition on post-treatment covariates
Inducing new pretreatment association (M-bias)
Inducing new \textbf{pretreatment} association (M-bias)

\begin{center}
\begin{tikzpicture}
  \node [draw, circle] (z1) at (0,0) {$Z_1$};
  \node [draw, circle] (z2) at (0,1) {$Z_2$};
  \node [draw, circle] (z3) at (1,0) {$Z_3$};
  \node [draw, circle] (t) at (-0.5,0) {$T$};
  \node [draw, circle] (y) at (0.5,0) {$Y$};
  \draw [->] (z1) -- (z2);
  \draw [->] (z2) -- (z3);
  \draw [->] (t) -- (z2);
  \draw [->] (z3) -- (y);
\end{tikzpicture}
\end{center}

See \textbf{Elwert \& Winship (2014)} for many real examples of collider bias
Questions:
1. What are the nonparametric structural equations for this causal graph?
2. What are the endogenous and exogenous variables in this causal graph?
3. What is collider bias?
The *do*-operator

Main assumption: modularity

Backdoor adjustment

Structural causal models

A complete example with estimation
Problem: effect of sodium intake on blood pressure
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality.
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
- Epidemiological example taken from Luque-Fernandez et al. (2018)
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

• Epidemiological example taken from Luque-Fernandez et al. (2018)
• Outcome Y: (systolic) blood pressure (continuous)
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)
- Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
• Epidemiological example taken from Luque-Fernandez et al. (2018)
• Outcome Y: (systolic) blood pressure (continuous)
• Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
• Covariates
  • W age
  • Z amount of protein excreted in urine
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

• Epidemiological example taken from Luque-Fernandez et al. (2018)
• Outcome Y: (systolic) blood pressure (continuous)
• Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
• Covariates
  • W age
  • Z amount of protein excreted in urine
• Simulation: so we know the “true” ATE is 1.05
The causal graph

- Sodium intake
- Blood pressure
- Amount of protein excreted in urine
- Age

A complete example with estimation
Identification

A complete example with estimation
Identification

Causal estimand: \( \mathbb{E}[Y \mid do(t)] \)
Identification

Causal estimand: $\mathbb{E}[Y \mid do(t)]$

Statistical estimand from last week: $\mathbb{E}_{W,Z}[Y \mid t, W, Z]$
Identification

Causal estimand: \[ \mathbb{E}[Y \mid do(t)] \]

Statistical estimand from last week:
\[ \mathbb{E}_{W,Z}[Y \mid t, W, Z] \]
Identification

Causal estimand: \( \mathbb{E}[Y \mid do(t)] \)

Statistical estimand from last week: \( \mathbb{E}_{W,Z}[Y \mid t, W, Z] \)
Identification

Causal estimand:
\[
E[Y | do(t)]
\]

Statistical estimand from last week:
\[
E_{W,Z}[Y | t, W, Z]
\]

Statistical estimand from causal graph:
\[
E_W E[Y | t, W]
\]
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$
Estimation of ATE

True ATE: \( \mathbb{E} [Y(1) - Y(0)] = 1.05 \)

Identification: \( \mathbb{E} [Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]] \)
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[ \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X] \right]$}

Estimation:
Estimation of ATE

True ATE: \( \mathbb{E}[Y(1) - Y(0)] = 1.05 \)

Identification: \( \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]] \)

Estimation: \( \frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]] \)
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Model (linear regression)
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Estimates:
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Model (linear regression)

Estimates:

$X = \{\} \text{ (naive): 5.33}$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Model (linear regression)

Estimates:

$X = \{\}$ (naive): $5.33 - 1.05 = 4.28$\% error
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Model (linear regression)

Estimates:

$X = \{\}$ (naive): $5.33$  \hspace{1cm} $\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): 0.85
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[ \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X] \right]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Estimates:

$X = \{\}$ (naive): $5.33$ $\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): $0.85$ $19\%$ error
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_x [ \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Model (linear regression)

Estimates:
- $X = \{\} \ (\text{naive}): 5.33 \quad \frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\% \ \text{error}$
- $X = \{W, Z\} \ (\text{last week}): 0.85 \quad 19\% \ \text{error}$
- $X = \{W\} \ (\text{unbiased}): 1.0502$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_i [\mathbb{E}[Y \mid T = 1, X = x_i] - \mathbb{E}[Y \mid T = 0, X = x_i]]$

Model (linear regression)

Estimates:

$X = \{\}$ (naive): $5.33$  \quad $\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$ error

$X = \{W, Z\}$ (last week): $0.85$  \quad $19\%$ error

$X = \{W\}$ (unbiased): $1.0502$  \quad $0.02\%$ error
M-bias

A complete example with estimation
M-bias

A complete example with estimation
M-bias

A complete example with estimation
M-bias

A complete example with estimation