# Week 7

Wentao Gao

# Framework

**Potential outcomes(Rubin)**

**Causation in Graphs(Pearl)**

# Potential outcomes(Rubin)

# Fundamental Problem

- It is impossible to observe all potential outcomes for a given individual.

- Example:
  if, Get a dog, observing happiness after getting a dog, Y(1);
  or, not get a dog, observing happiness , Y(0);
  However, we cannot get and not get at the same time.

- It is fundamental because we cannot observe both Y(1) and Y(0) at the same time. Then we cannot observe causal effect Y(1) – Y(0)

# Potential outcome & counterfactual outcome

- In potential outcome framework. We also called potential outcomes as counterfactual outcomes. However, it is quite mess.

- Because the potential outcome Y(t) will never become counter unless we observe another potential outcome Y(t_). The potential outcome that is observed is sometimes referred to as a factual.

- Note that there are no counterfactuals or factuals until the outcome is observed. Before that, there are only potential outcomes average treatment effect (ATE):

$$\tau \triangleq \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[Y(1) - Y(0)]\,,$$

$$\text{ATE}\ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$$ (Binary)

associational difference:

$$\mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0].$$

**They are not equal due to confounding**

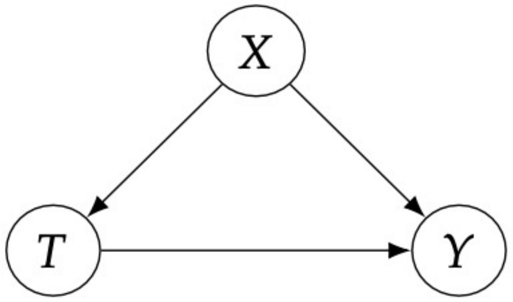



**Figure 2.1:** Causal structure of $X$ confounding the effect of $T$ on $Y$.

# What assumption(s) would make it so that the ATE is simply the associational difference?

| $i$ | $T$ | $Y$ | $Y(1)$ | $Y(0)$ | $Y(1) - Y(0)$ |
|---|---|---|---|---|---|
| 1 | 0 | 0 | ? | 0 | ? |
| 2 | 1 | 1 | 1 | ? | ? |
| 3 | 1 | 0 | 0 | ? | ? |
| 4 | 0 | 0 | ? | 0 | ? |
| 5 | 0 | 1 | ? | 1 | ? |
| 6 | 1 | 1 | 1 | ? | ? |

what makes it valid to calculate the ATE by taking the average of the $Y(0)$ column, ignoring the question marks, and subtracting that from the average of the $Y(1)$ column, ignoring the question marks?

**Assumption 2.1** (Ignorability / Exchangeability)

$$(Y(1), Y(0)) \perp\!\!\!\perp T$$

## Another perspective on this assumption is that of exchangeability

$$\mathbb{E}[Y(1)|T=0] = \mathbb{E}[Y(1)|T=1]$$

$$\mathbb{E}[Y(0)|T=1] = \mathbb{E}[Y(0)|T=0]$$

$$\mathbb{E}[Y(1)|T=t] = \mathbb{E}[Y(1)]$$

$$\mathbb{E}[Y(0)|T=t] = \mathbb{E}[Y(0)]$$

**This assumption is key to causal inference because it allows us to reduce the ATE to the associational difference:**

$$\begin{aligned}\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] &= \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \\ &= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]\end{aligned}$$

**However, there is no reason to expect that the groups are the same in all relevant variables other than the treatment.**

**Assumption 2.2** (Conditional Exchangeability / Unconfoundedness)

$$(Y(1), Y(0)) \perp\!\!\!\perp T \mid X$$

*Backdoor adjustment is a way to achieve Conditional Exangeability/ Unconfoundedness*

**We do not have exchangeability in the data because $X$ is a common cause of $T$ and $Y$.**

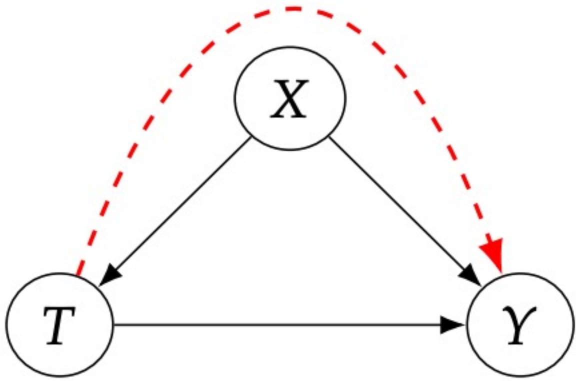



**Figure 2.3:** Causal structure of $X$ confounding the effect of $T$ on $Y$. We depict the confounding with a red dashed line.

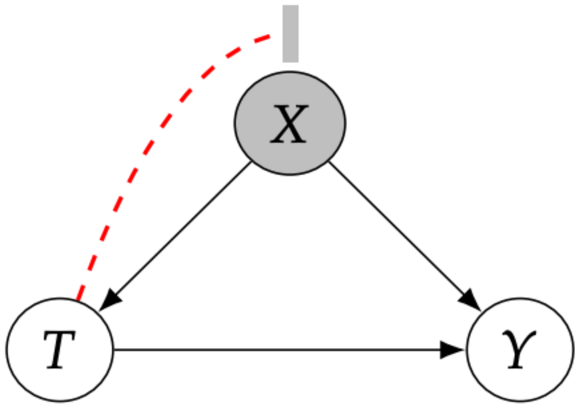



**Figure 2.4:** Illustration of conditioning on $X$ leading to no confounding.

**However, we do have conditional exchangeability in the data.**

$$\begin{aligned}\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) \mid X] \\ &= \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]\end{aligned}$$

# Adjustment Formula

**Theorem 2.1** (Adjustment Formula) *Given the assumptions of unconfoundedness, positivity, consistency, and no interference, we can identify the average treatment effect:*

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

**Assumption 2.3** (Positivity / Overlap / Common Support) *For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),*

$$0 < P(T = 1 \mid X = x) < 1$$

For discrete covariates and outcome, adjustment formula can be rewritten as follows:

$$\sum_x P(X = x) \left( \sum_y y \, P(Y = y \mid T = 1, X = x) - \sum_y y \, P(Y = y \mid T = 0, X = x) \right) \tag{2.10}$$

Then, applying Bayes' rule, this can be further rewritten:

$$\sum_x P(X = x) \left( \sum_y y \, \frac{P(Y = y, T = 1, X = x)}{P(T = 1 \mid X = x) P(X = x)} - \sum_y y \, \frac{P(Y = y, T = 0, X = x)}{P(T = 0 \mid X = x) P(X = x)} \right) \tag{2.11}$$

# No interference means that my outcome is unaffected by anyone else's treatment

**Assumption 2.4** (No Interference)

$$Y_i(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = Y_i(t_i)$$

# Consistency is the assumption that the outcome we observe *Y* is actually the potential outcome under the observed treatment *T*

**Assumption 2.5** (Consistency) *If the treatment is $T$, then the observed outcome $Y$ is the potential outcome under treatment $T$. Formally,*

$$T = t \implies Y = Y(t) \tag{2.12}$$

*We could write this equivalently as follow:*

$$Y = Y(T) \tag{2.13}$$

# Randomized Experiment: One perspective

**Definition 5.1** (Covariate Balance) *We have covariate balance if the distribution of covariates X is the same across treatment groups. More formally,*

$$P(X \mid T=1) \stackrel{d}{=} P(X \mid T=0) \tag{5.1}$$

Randomization implies covariate balance, across all covariates, even unobserved ones.

**Randomization makes causation equal to association**

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

# Another perspective: In DAG

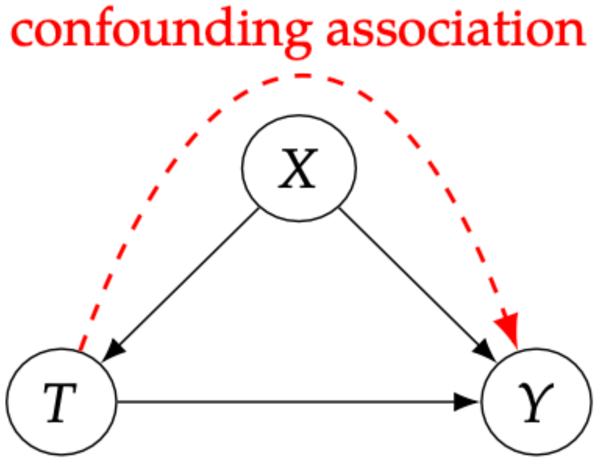



**Figure 5.1:** Causal structure of $X$ confounding the effect of $T$ on $Y$.

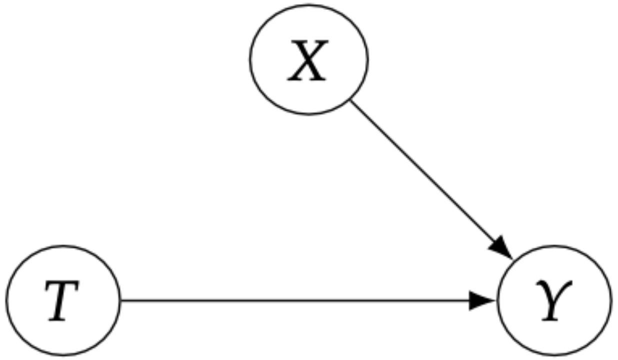


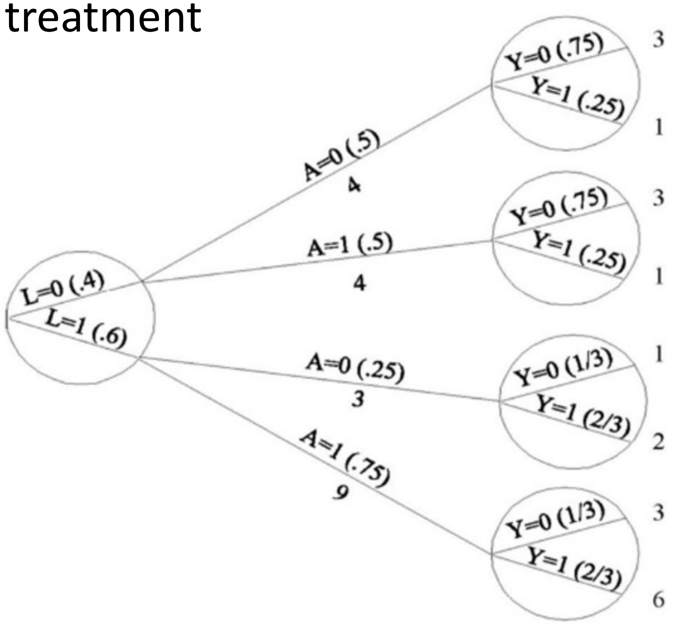

**Figure 5.2:** Causal structure when we randomize treatment.

$X$ is a confounder of the effect of $T$ on $Y$. Non-causal association flows along the backdoor path $T \leftarrow X \rightarrow Y$.

However, if we randomize $T$, something magical happens:
$T$ no longer has any causal parents, as we depict in Figure 5.2.

$$P(Y \mid do(T = t)) = P(Y \mid T = t)$$

The circles contain the bifurcations defined by non treatment variables.




# Inverse probability weighting

$$\frac{\Pr[Y^{a=1}=1]}{\Pr[Y^{a=0}=1]}$$

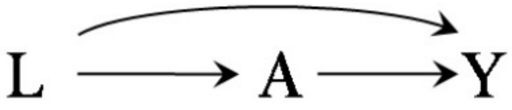
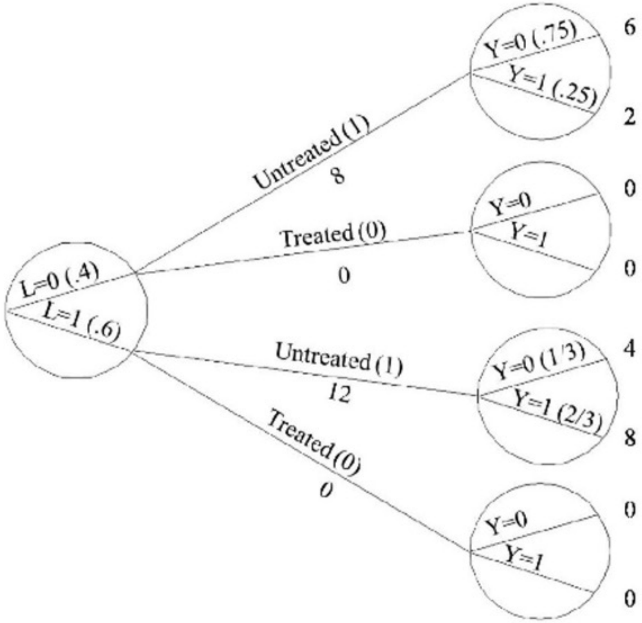
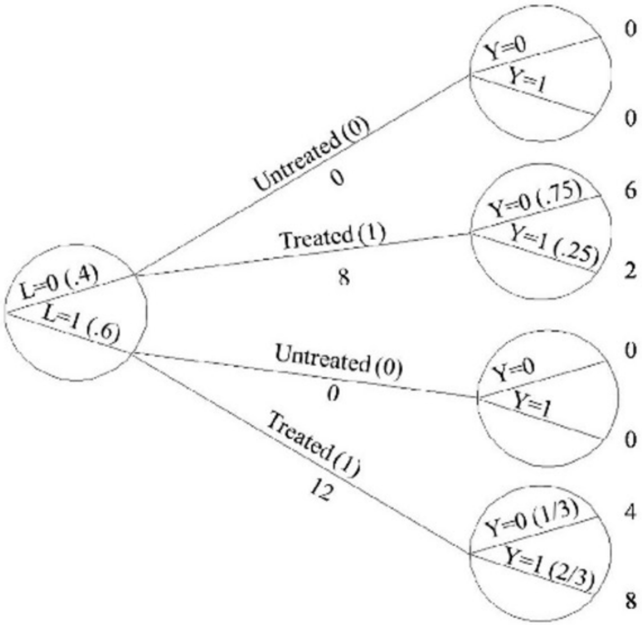



Figure 6.1



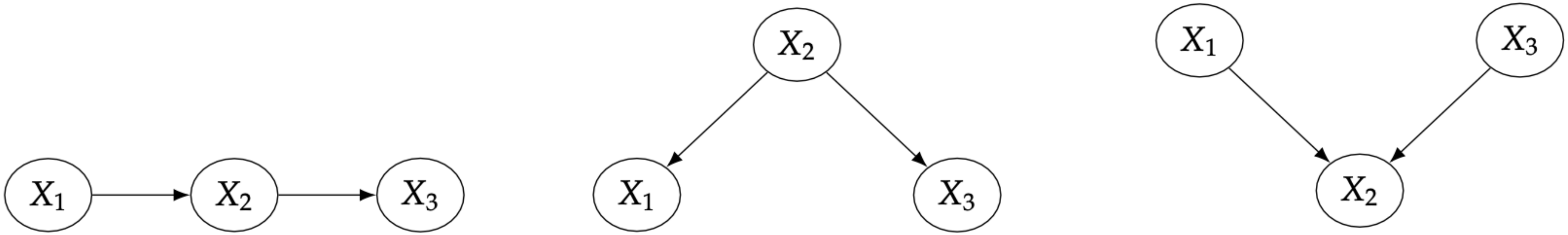
# Causation in Graphs(Pearl)

**Assumption 3.2** (Minimality Assumption)

1. *Given its parents in the DAG, a node $X$ is independent of all its non-descendants (Assumption 3.1).*
2. *Adjacent nodes in the DAG are dependent.*[3]

**Assumption 3.3** ((Strict) Causal Edges Assumption) *In a directed graph, every parent is a direct cause of all its children.*

# Basic Graphs




**(a)** Chain **(b)** Fork **(c)** Immorality

**Figure 3.9:** Basic graph building blocks

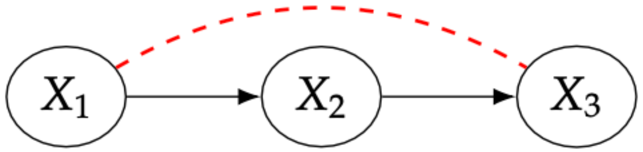



**Figure 3.12:** Chain with flow of association drawn as a dashed red arc.

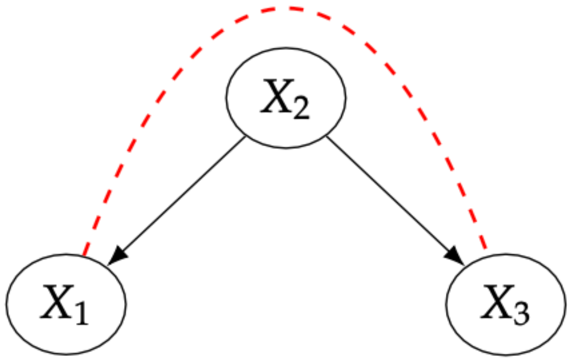



**Figure 3.13:** Fork with flow of association drawn as a dashed red arc.

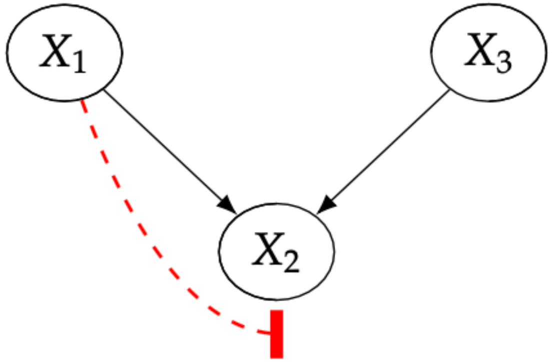



**Figure 3.16:** Immorality with association blocked by a collider.

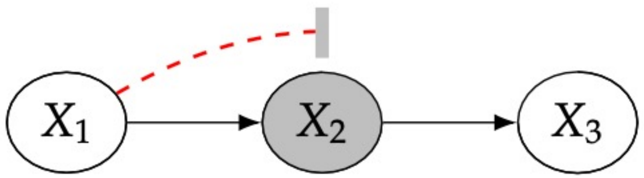



**Figure 3.14:** Chain with association blocked by conditioning on $X_2$.

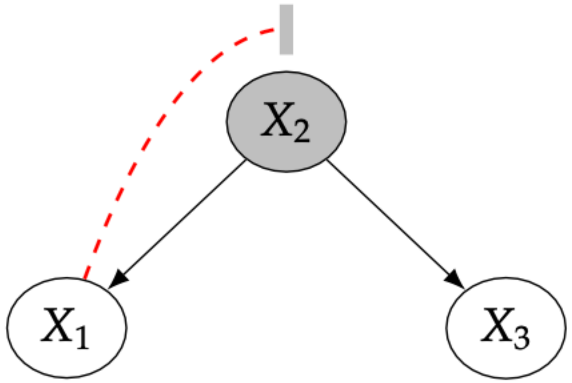



**Figure 3.15:** Fork with association blocked by conditioning on $X_2$.

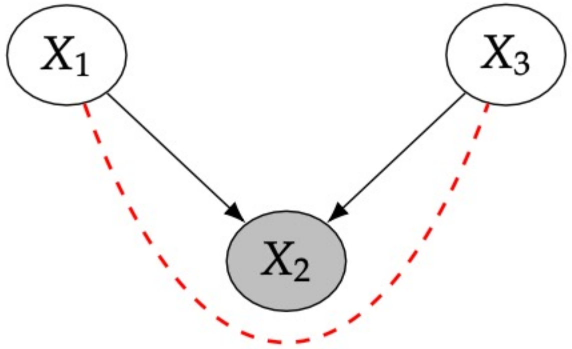
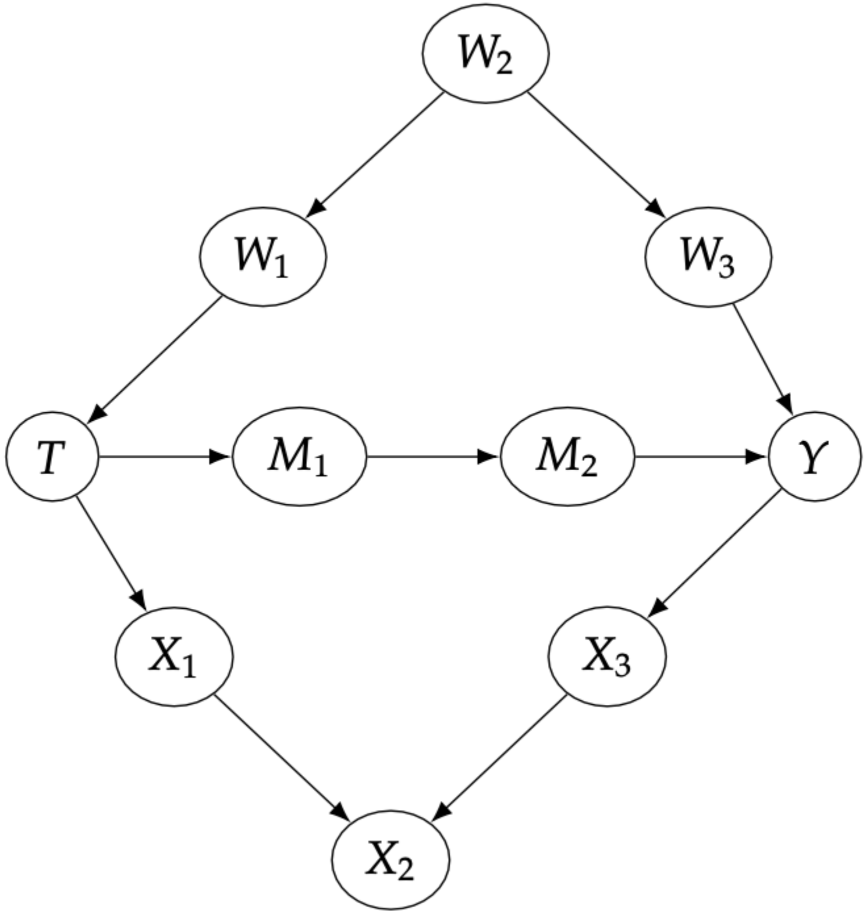


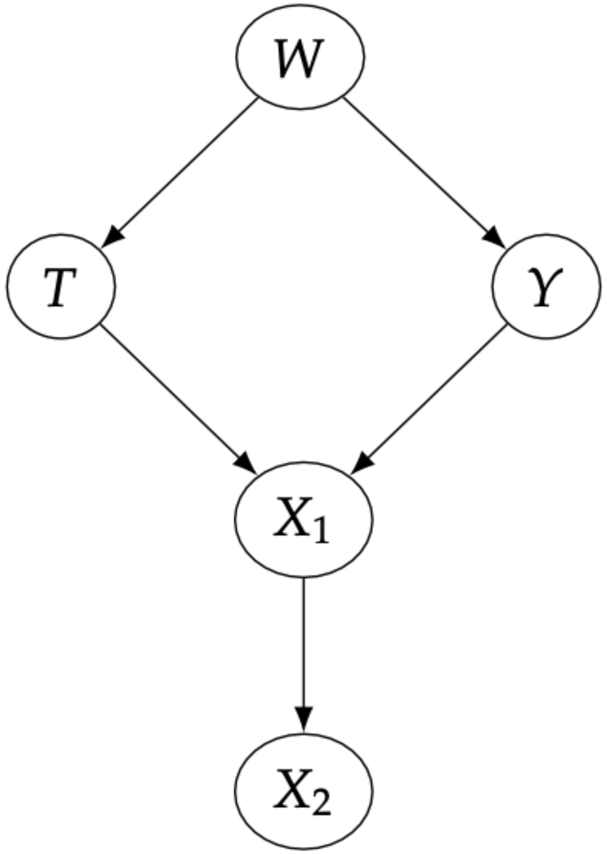

**Figure 3.17:** Immorality with association unblocked by conditioning on the collider.

**Definition 3.4** (d-separation) *Two (sets of) nodes X and Y are d-separated by a set of nodes Z if all of the paths between (any node in) X and (any node in) Y are blocked by Z [16].*




**(a)**

1. Are $T$ and $Y$ d-separated by the empty set?
2. Are $T$ and $Y$ d-separated by $W_2$?
3. Are $T$ and $Y$ d-separated by $\{W_2, M_1\}$?
4. Are $T$ and $Y$ d-separated by $\{W_1, M_2\}$?
5. Are $T$ and $Y$ d-separated by $\{W_1, M_2, X_2\}$?
6. Are $T$ and $Y$ d-separated by $\{W_1, M_2, X_2, X_3\}$?




**(b)**

1. Are $T$ and $Y$ d-separated by the empty set?
2. Are $T$ and $Y$ d-separated by $W$?
3. Are $T$ and $Y$ d-separated by $\{W, X_2\}$?

# Flow of Association and Causation

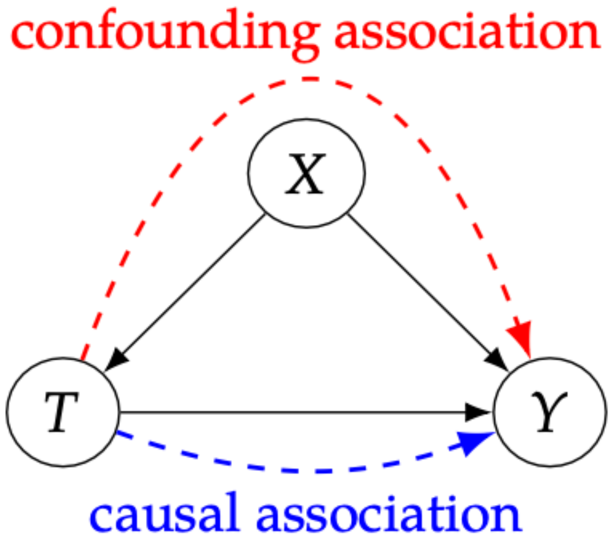



**Figure 3.20:** Causal graph depicting an example of how confounding association and causal association flow.

# The *do*-operator and Interventional Distributions

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

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

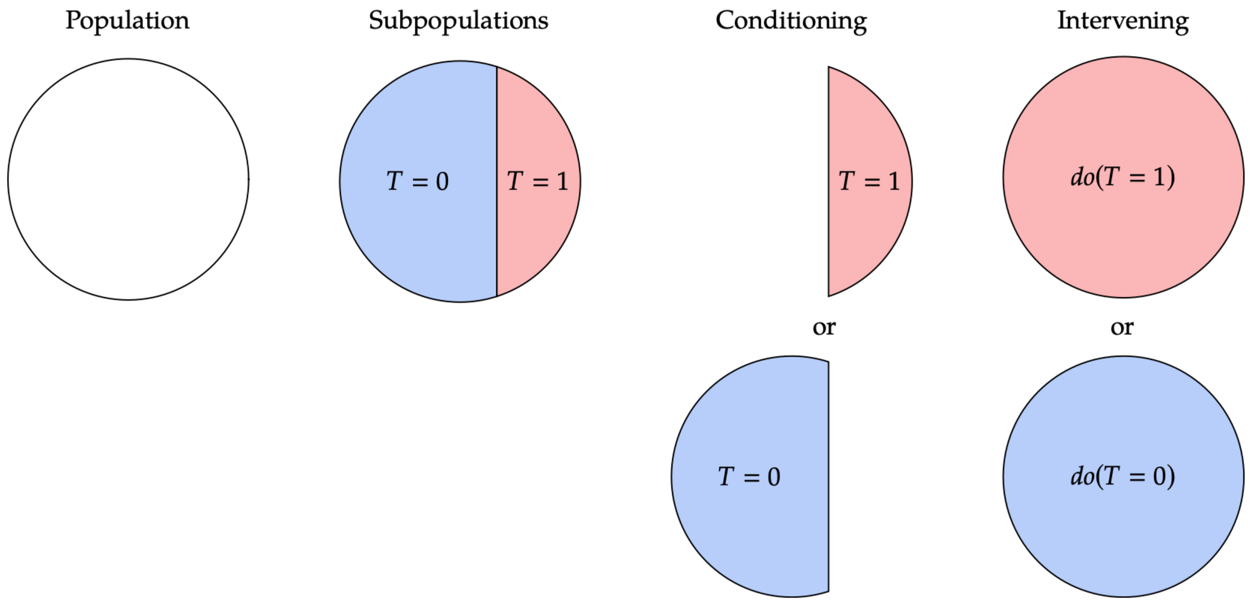



**Figure 4.2:** Illustration of the difference between conditioning and intervening

# The Backdoor Adjustment

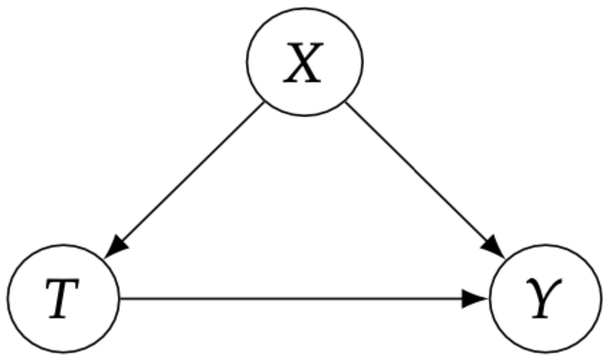



**Figure 4.5:** Simple causal structure where $X$ counfounds the effect of $T$ on $Y$ and where $X$ is the only confounder.

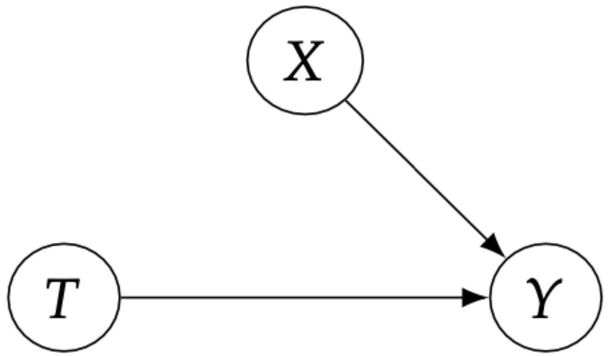



**Figure 4.6:** Manipulated graph that results from intervening on $T$, when the original graph is Figure 4.5.

**Definition 4.1** (Backdoor Criterion) *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 $T$.*

# Proof

$$
\begin{aligned}
P(y \mid do(t)) &= \sum_w P(y \mid do(t), w)\, P(w \mid do(t)) \\
&= \sum_w P(y \mid t, w)\, P(w \mid do(t)) \\
&= \sum_w P(y \mid t, w)\, P(w)
\end{aligned}
$$

# Relationship to d-separation

W is the set we use to make t and y are d separated in another path.

# Relation to Potential Outcomes

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_W\left[\mathbb{E}[Y \mid T = 1, W] - \mathbb{E}[Y \mid T = 0, W]\right]$$ Potential outcome adjustment formula

$$\mathbb{E}[Y \mid do(t)] = \sum_w \mathbb{E}[Y \mid t, w]\, P(w)$$ Backdoor adjustment formula

$$\mathbb{E}[Y \mid do(t)] = \mathbb{E}_W \mathbb{E}[Y \mid t, W]$$

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

$$(Y(1), Y(0)) \perp\!\!\!\perp T \mid W$$ Conditional exangeability

using graphical causal models, we know how to choose a valid $W$: we simply choose $W$ so that it satisfies the backdoor criterion.