Chapter 2

Policy Evaluation I: Identification

2.1 Motivation

A large portion of research in causality is concerned with what may be broadly called *policy evaluation*. Policy evaluation is concerned with determining the effect of an action, or a sequence of actions, on a particular outcome or set of outcomes. For example, in economics, we might ask:

"What is the average increase in lifetime earnings for individuals with a bachelors degree compared to only high school education?"

In healthcare, we might ask:

"For a patient with kidney stones, does surgery or medicine have a higher success rate in eliminating symptoms after one month?"

Abstractly, these questions involve estimating an **average treatment effect**

$$\mathbb{E}[Y \mid \texttt{do}(A = 1)] - \mathbb{E}[Y \mid \texttt{do}(A = 0)],$$

the difference in the average outcome Y of two interventions, $\texttt{do}(A = 1)$ and $\texttt{do}(A = 0)$.

Such questions may be extended in a variety of ways. First, we may care about continuous-valued treatments. For example, in economics, we might ask:

"What is the average increase in GDP for x billion dollars of investment in infrastructure?"

In healthcare, we might ask:

"What is the average reduction in tumor size for x milligrams of a PD-1 inhibitor (a common immunotherapy drug for breast cancer)?"

Abstractly, these questions involve estimating a **dose-response curve**, i.e. the function

$$f(x) = \mathbb{E}[Y \mid \texttt{do}(A = x)].$$

Second, we may care not only about population averages, but about averages for specific subpopulations. For example, in economics, we might ask:

"Given the annual income of an individual's parents, what is the average increase in lifetime earnings for each additional year of college?"

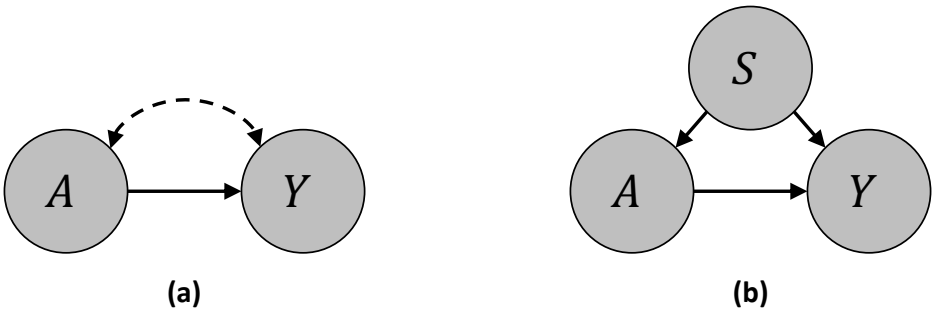

Figure 2.1: **(a)** An ADMG where the causal effect $\mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(A = a))$ is not identifiable. **(a)** An DAG where the causal effect $\mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(A = a))$ is identifiable. .

In healthcare, we might ask:

"Given the size of an individual's kidney stone, does surgery or medicine have a higher success rate?"

Abstractly, these questions involve estimating a **conditional average treatment effect**

$$f(\mathbf{s}) = \mathbb{E}[Y \mid \mathbf{S} = \mathbf{s}, \texttt{do}(A = 1)] - \mathbb{E}[Y \mid \mathbf{S} = \mathbf{s}, \texttt{do}(A = 0)]$$

The subfield of causality devoted to answering such questions may be called **treatment effect estimation** or **causal inference**. This is a very rich subfield, and we will only scratch its surface.

2.1.1 Randomized Controlled Trials

When a randomized controlled trial (RCT) can be conducted, questions about treatment effect are often conceptually easy to answer.

In a randomized controlled trial, patients are randomly assigned to treatment ($\texttt{do}(A = 1)$) or control ($\texttt{do}(A = 0)$). Thus, one obtains samples from the distributions $\mathbb{P}_{\mathcal{X}}(\mathcal{X} \mid \texttt{do}(A = 1))$ and $\mathbb{P}_{\mathcal{X}}(\mathcal{X} \mid \texttt{do}(A = 0))$, from which conditional expectations can be directly estimated using averaging.

However, even in the study of randomized controlled trials, there is a rich set of questions, for example, questions about experimental design (i.e., the assignment of patients to treatment and control) or about the use of auxiliary variables.

2.1.2 Identifiability from observational data

In this lecture, we will discuss how to answer questions about treatment effects from only observational data, i.e., data where the treatment variable has not been manipulated. This is of great interest in many fields where randomized controlled trials are infeasible due to cost or ethical concerns. For instance, in healthcare, there is a large amount of data in electronic health records (EHRs) about patients, treatments, and outcomes, which does not come from RCTs. We will now formally define the issue of identifiability of an interventional distribution from observational data.

Definition 2.1. *Let $\mathcal{G}$ be an ADMG. We say that an interventional distribution $\mathbb{P}(Y \mid \texttt{do}(\mathbf{A} = \mathbf{a}))$ is (observationally)* **identifiable** *from $\mathcal{G}$ if, for any two SCMs M_a and M_b with causal graph $\mathcal{G}$ and entailed distribution $\mathbb{P}_{\mathcal{X}}(\mathcal{X})$, we have $\mathbb{P}^{M_a}_{\mathcal{X}}(X_i \mid \texttt{do}(\mathbf{A} = \mathbf{a})) = \mathbb{P}^{M_b}_{\mathcal{X}}(X_i \mid \texttt{do}(\mathbf{A} = \mathbf{a}))$.*

don't understand

The following example shows how an interventional distribution may not be identifiable.

why Xi not Y? Aren't we identifying Y not Xi?

Example 2.1. *Let $\mathcal{G}$ be the ADMG in Figure 2.1(a). Let M_a be the structural causal model*

$$\begin{aligned} A &= \varepsilon_1 \\ Y &= A \oplus \varepsilon_1 \\ \varepsilon_1 &\sim \mathsf{Ber}(0.5) \end{aligned}$$

Let M_b be the structural causal model

$$\begin{aligned} A &= \varepsilon_1 \\ Y &= 0 \\ \varepsilon_1 &\sim \mathsf{Ber}(0.5) \end{aligned}$$

Both models have causal graph $\mathcal{G}$, and the entailed distributions $\mathbb{P}^{M_a}_{\mathcal{X}}(A, Y)$ and $\mathbb{P}^{M_b}_{\mathcal{X}}(A, Y)$ are the same. However, $\mathbb{P}^{M_a}_{\mathcal{X}}(Y = 1 \mid \mathtt{do}(A = 0)) = 0.5$, while $\mathbb{P}^{M_b}_{\mathcal{X}}(Y = 1 \mid \mathtt{do}(A = 0)) = 0$.

In contrast, if the variable confounding A and Y is observed, then we may identify $\mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a))$, as follows.

S is adjustment set

Theorem 2.1. *Assume that $\mathbb{P}_{\mathcal{X}}$ factorizes according to $\mathcal{G}$ in Figure 2.1(b). Then*

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a)) = \sum_s \mathbb{P}_{\mathcal{X}}(Y \mid A = a, S = s)\mathbb{P}_{\mathcal{X}}(S = s)$$

Proof. We have by the law of total probability that

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a)) = \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a), S = s)\mathbb{P}_{\mathcal{X}}(S = s \mid \mathtt{do}(A = a))$$

We may condition on $A = a$ without changing the probabilities, since $A = a$ with probability one:

$$= \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a), A = a, S = s)\mathbb{P}_{\mathcal{X}}(S = s \mid \mathtt{do}(A = a), A = a)$$

Since Y is not intervened and $\mathrm{pa}_{\mathcal{G}}(Y) = \{A, S\}$, we have by consistency that

$$= \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid A = a, S = s)\mathbb{P}_{\mathcal{X}}(S = s \mid \mathtt{do}(A = a))$$

Since $\mathbf{S}$ is not intervened and $\mathrm{pa}_{\mathcal{G}}(S) = \varnothing$, we have by consistency that

$$= \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid A = a, S = s)\mathbb{P}_{\mathcal{X}}(S = s)$$

□

Remark 2.1. *In general, we say that* $\mathbf{S}$ *is an* **adjustment set** *for $\mathbb{P}(Y \mid \mathtt{do}(A = a))$ if the following equation holds:*

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a)) = \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid A = a, \mathbf{S} = \mathbf{s})\mathbb{P}_{\mathcal{X}}(\mathbf{S} = \mathbf{s})$$

Our goal in this lecture will be to establish more general conditions under which an interventional distribution is identifiable from observational data. We will study two well-known identification formulas, called *backdoor adjustment* and *frontdoor adjustment*. To prove the correctness of these formulas, we will have to introduce several fundamental concepts and basic results in the study of graphical models.

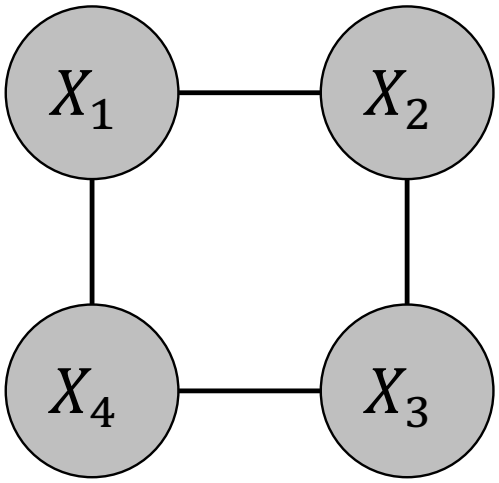

Figure 2.2: The undirected graph $\mathcal{G}$ from Example 2.2.

2.2 Factorization and Markovianity in undirected graphs

First, we introduce undirected graphical models, which we will see to be an indispensable tool in our study of directed graphical models. The next two definitions introduce a relationship between undirected graphs and probability distributions based on *factorization*.

Definition 2.2. *Let $\mathcal{G}$ be an undirected graph on nodes $\mathcal{X}$. A* **clique** *in $\mathcal{G}$ is a set $\mathbf{C} \subseteq \mathcal{X}$ such that, for each pair $X_i, X_j \in \mathbf{C}$, we have the edge $X_i - X_j$ in $\mathcal{G}$. The set of all cliques in an undirected graph is denoted $\mathcal{C}(\mathcal{G})$.*

Definition 2.3. *Let $\mathcal{G}$ be an undirected graph on nodes $\mathcal{X}$ and $\mathbb{P}_{\mathcal{X}}$ be a distribution. We say that $\mathbb{P}_{\mathcal{X}}$* **factorizes** *according to $\mathcal{G}$ if there exists an indexed set of functions $\{\phi_{\mathbf{C}}\}_{\mathbf{C} \in \mathcal{C}(\mathcal{G})}$ such that*

def of factorization

$$\mathbb{P}_{\mathcal{X}}(\mathcal{X}) = \prod_{\mathbf{C} \in \mathcal{C}(\mathcal{G})} \phi_{\mathbf{C}}(\mathbf{C})$$

Example 2.2. *Let $\mathcal{G}$ be the undirected graph in Figure 2.2. We have*

$$\mathcal{C}(\mathcal{G}) = \{\{X_1, X_2\}, \{X_2, X_3\}, \{X_3, X_4\}, \{X_4, X_1\}\}$$

Let

$$\mathbb{P}_{\mathcal{X}}(\mathcal{X}) \propto \exp\{X_1 X_2\} \exp\{X_2 X_3\} \exp\{X_3 X_4\} \exp\{X_4 X_1\}$$

$\mathbb{P}_{\mathcal{X}}$ factorizes according to $\mathcal{G}$.

We will make use of the following fundamental results about factorization:

Claim 2.1. *$\mathbf{A} \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} \mathbf{B} \mid \mathbf{S}$ if and only if there exists h_1, h_2 such that $\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{B}, \mathbf{S}) = h_1(\mathbf{A}, \mathbf{S}) h_2(\mathbf{B}, \mathbf{S})$.*

def of independence

Proof. If $\mathbf{A} \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} \mathbf{B} \mid \mathbf{S}$, then $\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{B} \mid \mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{A} \mid \mathbf{S}) \mathbb{P}_{\mathcal{X}}(\mathbf{B} \mid \mathbf{S})$. Thus,

$$\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{B}, \mathbf{S}) = \mathbb{P}_{\mathcal{S}}(\mathbf{A}, \mathbf{S}) \mathbb{P}_{\mathcal{X}}(\mathbf{B} \mid \mathbf{S}),$$

i.e., $h_1(\mathbf{A}, \mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{S})$ and $h_2(\mathbf{B}, \mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{B} \mid \mathbf{S})$.

Conversely, suppose $\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{B}, \mathbf{S}) = h_1(\mathbf{A}, \mathbf{S}) h_2(\mathbf{B}, \mathbf{S})$. Integrating both sides over $\mathbf{A}$, we have

$$\mathbb{P}_{\mathcal{X}}(\mathbf{B}, \mathbf{S}) = \alpha_1(\mathbf{S}) h_2(\mathbf{B}, \mathbf{S}) \qquad \text{for } \alpha_1(\mathbf{S}) = \int h_1(\mathbf{A}, \mathbf{S}) d\mathbf{A}$$

Similarly, we have

$$\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{S}) = \alpha_2(\mathbf{S}) h_1(\mathbf{A}, \mathbf{S}) \qquad \text{for } \alpha_2(\mathbf{S}) = \int h_2(\mathbf{B}, \mathbf{S}) d\mathbf{B}$$

Therefore,

$$\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{B}, \mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{S}) \mathbb{P}_{\mathcal{X}}(\mathbf{B}, \mathbf{S}) (\alpha_1(\mathbf{S}) \alpha_2(\mathbf{S}))^{-1}$$

Integrating both sides over $\mathbf{A}$ and $\mathbf{B}$ and re-arranging, we have $\alpha_1(\mathbf{S})\alpha_2(\mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{S})$. Plugging in this equality and dividing both sides by $\mathbb{P}_{\mathcal{X}}(\mathbf{S})$, we have

$$\mathbb{P}_{\mathcal{X}}(\mathbf{A}, \mathbf{B} \mid \mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{A} \mid \mathbf{S}) \mathbb{P}_{\mathcal{X}}(\mathbf{B} \mid \mathbf{S})$$

Which proves the result. □

Next, we introduce a relationship between undirected graphs and probability distributions based on *separation* and *conditional independence*. We will see that this relationship is closely tied to factorization.

Definition 2.4. *Let $\mathcal{G}$ be an undirected graph on nodes $\mathcal{X}$. Let $\mathbf{A}, \mathbf{B}, \mathbf{S}$ be disjoint subsets of $\mathcal{X}$. We say* **S separates A** *and* **B** *in $\mathcal{G}$ if all paths from $\mathbf{A}$ to $\mathbf{B}$ go through $\mathbf{S}$. We denote this by $\mathbf{A} \perp\!\!\!\perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{S}$. We denote the complete set of separation statements in $\mathcal{G}$ as $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G})$, i.e.,*

$$\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) := \{(\mathbf{A}, \mathbf{B}, \mathbf{S}) : \mathbf{A} \perp\!\!\!\perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{S}\}$$

Definition 2.5. *Given a distribution $\mathbb{P}_{\mathcal{X}}$, we define its* **independence model** *as*

$$\mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}}) := \{(\mathbf{A}, \mathbf{B}, \mathbf{S}) : \mathbf{A} \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} \mathbf{B} \mid \mathbf{S}\}$$

We now introduce terminology to capture the relationship between the two sets $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G})$ and $\mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$.

Definition 2.6. *Let $\mathcal{G}$ be an undirected graph on nodes $\mathcal{X}$ and $\mathbb{P}_{\mathcal{X}}$ be a distribution. We say that $\mathbb{P}_{\mathcal{X}}$ is* **Markov** *with respect to $\mathcal{G}$ if $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) \subseteq \mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$.*

Remark 2.2. *Note the direction of the inclusion: every separation in $\mathcal{G}$ implies a corresponding conditional independence in $\mathbb{P}_{\mathcal{X}}$. However, $\mathbb{P}_{\mathcal{X}}$ may include additional conditional independences, beyond the separations of $\mathcal{G}$. For example, if $\mathbb{P}_{\mathcal{X}}$ is a product distribution, then $\mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$ includes all conditional independence statements, and thus $\mathbb{P}_{\mathcal{X}}$ is Markov with respect to any undirected graph $\mathcal{G}$.*

Requiring that $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) = \mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$ is a much stronger condition which we call **faithfulness** *of $\mathbb{P}_{\mathcal{X}}$ to $\mathcal{G}$. We will return to the concept of faithfulness when we discuss structure learning.*

Example 2.3. *Let $\mathcal{G}$ be the undirected graph and $\mathbb{P}_{\mathcal{X}}$ be the distribution from Example 2.2. We have*

$$\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) = \{(X_1, X_3, \{X_2, X_4\}), (X_2, X_4, \{X_1, X_3\})\}$$

We also have

$$\mathbb{P}_{\mathcal{X}}(\mathcal{X}) = (\exp\{X_1, X_2\} \exp\{X_4 X_1\})(\exp\{X_2, X_3\} \exp\{X_3, X_4\}) = h_1(X_1, X_2, X_4) h_1(X_3, X_2, X_4)$$

Thus, $X_1 \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} X_3 \mid X_2, X_4$ by Claim 2.1. Symmetrically, $X_2 \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} X_4 \mid X_1, X_3$. So, $\mathbb{P}_{\mathcal{X}}$ is Markov with respect to $\mathcal{G}$.

The next result establishes the connection between factorization and Markovianity.

Proposition 2.1. *Suppose $\mathbb{P}_{\mathcal{X}}$ factorizes according to the undirected graph $\mathcal{G}$. Then $\mathbb{P}_{\mathcal{X}}$ is Markov with respect to $\mathcal{G}$.*

Proof. Let $\mathbf{A} \perp\!\!\!\perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{S}$. Let $\mathbf{A}'$ denote the connected components of $\mathcal{G}[\mathcal{X} \setminus \mathbf{S}]$ which contain $\mathbf{A}$. Let $\mathbf{B}' = \mathcal{X} \setminus (\mathbf{S} \cup \mathbf{A}')$, so that $\mathcal{X} = \mathbf{S} \sqcup \mathbf{A}' \sqcup \mathbf{B}'$, i.e. $\mathcal{X}$ is partitioned into 3 sets. Since $\mathbf{A}$ and $\mathbf{B}$ are separated by

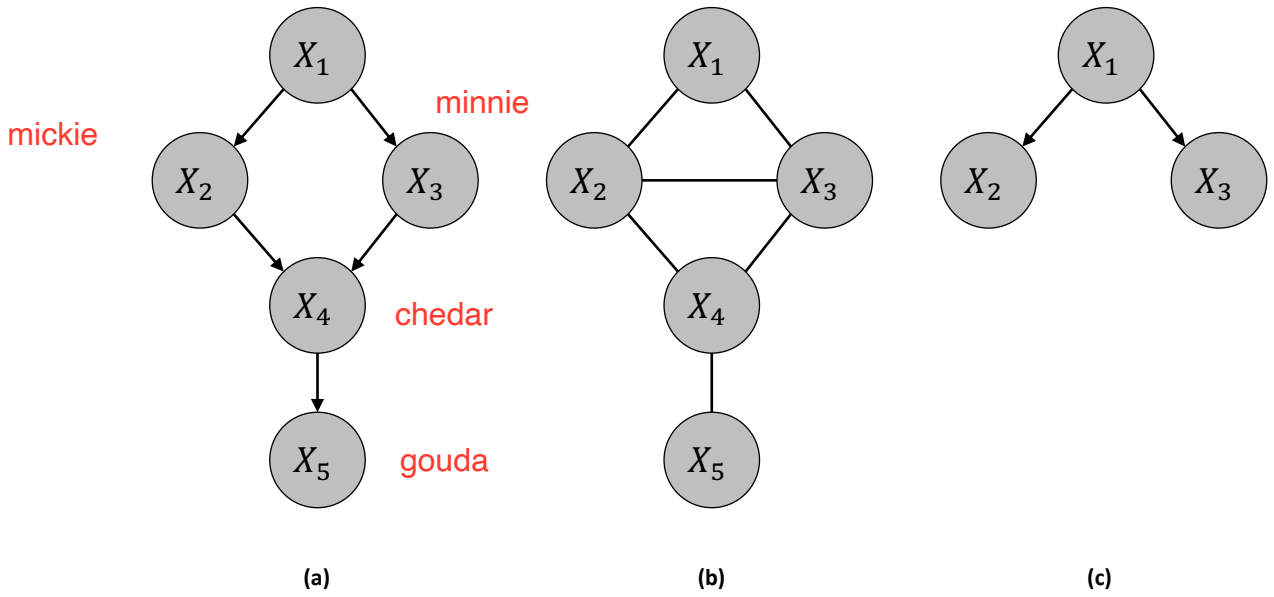

Figure 2.3: **(a)** The DAG $\mathcal{G}$. **(b)** The moral graph $\overline{\mathcal{G}}$. **(c)** The induced subgraph $\mathcal{G}[\mathbf{V}]$ for $\mathbf{V} = \{X_1, X_2, X_3\}$.

$\mathbf{S}$, any clique $\mathbf{C}$ of $\mathcal{G}$ is either a subset of $\mathbf{A}' \cup \mathbf{S}$ or of $\mathbf{B}' \cup \mathbf{S}$. Let $\mathcal{C}_1 = \{\mathbf{C} \in \mathcal{C}(\mathcal{G}) : \mathbf{C} \subseteq \mathbf{A}' \cup \mathbf{S}\}$ and $\mathcal{C}_2 = \mathcal{C}(\mathcal{G}) \setminus \mathcal{C}_1$. Thus,

$$\mathbb{P}_{\mathcal{X}}(\mathcal{X}) = \left(\prod_{\mathbf{C} \in \mathcal{C}_1} \phi_{\mathbf{C}}(\mathbf{C})\right)\left(\prod_{\mathbf{C} \in \mathcal{C}_2} \phi_{\mathbf{C}}(\mathbf{C})\right) = h_1(\mathbf{A}' \cup \mathbf{S}) h_2(\mathbf{B}' \cup \mathbf{S})$$

Thus, $\mathbf{A}' \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} \mathbf{B}' \mid \mathbf{S}$ by Claim 2.1. This further implies that $\mathbf{A} \perp\!\!\!\perp_{\mathbb{P}_{\mathcal{X}}} \mathbf{B} \mid \mathbf{S}$. □

Remark 2.3. *Proposition 2.1 says that, for undirected graphs,* factorization implies Markovianity. *Under conditions (such as positivity of $\mathbb{P}_{\mathcal{X}}$), it can be shown that the converse holds, i.e., Markovianity implies factorization. This result is the* Hammersley-Clifford Theorem, *which we will not need. See todo: theorem of Lauritzen (1996).*

2.3 Moral graphs

First, we define factorization for DAGs:

Definition 2.7. *Let $\mathcal{G}$ be a DAG on nodes $\mathcal{X}$ and $\mathbb{P}_{\mathcal{X}}$ be a distribution. We say that $\mathbb{P}_{\mathcal{X}}$* **factorizes** *according to $\mathcal{G}$ if*

$$\mathbb{P}_{\mathcal{X}}(\mathcal{X}) = \prod_{V_i \in \mathcal{S}} \mathbb{P}_{\mathcal{X}}(X_i \mid \mathrm{pa}_{\mathcal{G}}(X_i)) \tag{2.1}$$

Claim 2.2. *Let $\mathbb{P}_{\mathcal{X}}$ be the entailed distribution of a Markovian structural causal model with causal graph $\mathcal{G}$. Then $\mathbb{P}_{\mathcal{X}}$ factorizes according to $\mathcal{G}$.*

We can now connect directed graphical models with undirected graphical models via an important transformation called *moralization*.

Definition 2.8. *Let $\mathcal{G}$ be a DAG on nodes $\mathcal{X}$. The* **moral graph** *of $\mathcal{G}$, denoted $\overline{\mathcal{G}}$, is the undirected graph with the edge $X_i - X_j$ if:*

- *$X_i \to X_j$ or $X_j \to X_i$, or*

Text

- *There exists* X_k *such that* $X_i \to X_k \leftarrow X_j$

If $X_i - X_j$ *is in* $\overline{\mathcal{G}}$ *but not* $\mathcal{G}$*, then we call the edge* $X_i - X_j$ *a* **marriage**.

Example 2.4. *Let* $\mathcal{G}$ *be the DAG in Figure 2.3(a). Then* $\overline{\mathcal{G}}$ *is the graph in Figure 2.3(b).*

Finally, we establish that factorization is preserved under moralization.

Lemma 2.1. *Suppose* $\mathbb{P}_{\mathcal{X}}$ *factorizes with respect to* $\mathcal{G}$*. Then* $\mathbb{P}_{\mathcal{X}}$ *factorizes according to* $\overline{\mathcal{G}}$*.*

Proof. Let $\mathcal{C} = \{\overline{\text{pa}}_{\mathcal{G}}(X_i) \mid X_i \in \mathcal{X}\}$. By construction, $\mathcal{C} \subseteq \mathcal{C}(\overline{\mathcal{G}})$. Then

$$\mathbb{P}_{\mathcal{X}}(X) = \prod_{X_i \in \mathcal{X}} \mathbb{P}(X_i \mid \text{pa}_{\mathcal{G}}(i)) = \prod_{\mathbf{C} \in \mathcal{C}} \phi_{\mathbf{C}}(\mathbf{C}),$$

i.e., $\mathbb{P}_{\mathcal{X}}(X)$ factorizes according to $\overline{\mathcal{G}}$. □

2.4 Separation in directed graphs

As in undirected graphs, we introduce a notion of separation for directed graphical models. However, this notion of separation is significantly more complex, and will require us to build up a definition in smaller pieces. As a first step, we consider paths in DAGs, and divide nodes on the path into two types.

Definition 2.9. *Let* $\gamma = \langle \gamma_1, \ldots, \gamma_M \rangle$ *be a path in DAG. We call a node* γ_m *on this path a* **collider** *on* γ *if* $\gamma_{m-1} \to \gamma_m \leftarrow \gamma_{m+1}$*, i.e., two arrowheads "collide" at* γ_m*. Otherwise, we call* γ_m *a* **non-collider**.

Now, we define what it means for a path to be blocked at a certain node, with the conditions for being blocked differing for colliders and non-colliders.

Then why blockedness is important? because it defines "connectedness" which determines whether there's causal explanation between two vraiblaes

Definition 2.10. *Given an DAG* $\mathcal{G}$*, a set* $\mathbf{S} \subseteq \mathcal{X}$*, and a path* γ*, we call a node* γ_m *on the path a:*

- **Blocked non-collider** *if* γ_m *is a non-collider and* $\gamma_m \in \mathbf{S}$*, or a*
- **Blocked collider** *if* γ_m *is a collider and* $\overline{\text{de}}_{\mathcal{G}}(\gamma_m) \cap \mathbf{S} = \varnothing$*, i.e., neither* γ_m *nor any of its descendants belong to* $\mathbf{S}$.

Otherwise, we call the node γ_m **unblocked**.

block or non-block defined by S

Collider is important in terms of finiing "block"edness.

Now, we may define d-connection simply in terms of the path being completely unblocked.

Definition 2.11. *Let* $\mathcal{G}$ *be an DAG on nodes* $\mathcal{X}$*. Given two nodes* X_i *and* X_j *and a set* $\mathbf{S} \subseteq \mathcal{X} \setminus \{X_i, X_j\}$*, we call a path* γ *between* X_i *and* X_j *a* **d-connecting path** *if all nodes in* γ *are unblocked. We say that* X_i *and* X_j *are* **d-connected** *given* $\mathbf{S}$ *if there exists any d-connecting path.*

d-connected: exists at least one path that is d-connecting (which means all nodes in a path are not bloc

Conversely, we define d-separation as the situation where all paths are blocked given $\mathbf{S}$.

Definition 2.12. *Given an DAG* $\mathcal{G}$*, we say two nodes* X_i *and* X_j *are* **d-separated** *by a set* $\mathbf{S}$ *if they are not d-connected given* $\mathbf{S}$*. We denote this by* $X_i \perp\!\!\!\perp_{\mathcal{G}} X_j \mid \mathbf{S}$*. For disjoints sets* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{S}$*, we say* $\mathbf{A}$ *is d-separated from* $\mathbf{B}$ *given* $\mathbf{S}$ *if* $X_i \perp\!\!\!\perp_{\mathcal{G}} X_j \mid \mathbf{S}$ *for all* $X_i \in \mathbf{A}$*,* $X_j \in \mathbf{S}$*. We denote the complete set of d-separation statements in an DAG* $\mathcal{G}$ *as* $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G})$*, i.e.,*

isn't it B? (YES)

$$\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) := \{(\mathbf{A}, \mathbf{B}, \mathbf{S}) : \mathbf{A} \perp\!\!\!\perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{S}\}$$

Example 2.5. *Let $\mathcal{G}$ be the DAG in Figure 2.3(a).*

(a) *$\gamma = X_1 \to X_2 \to X_4$ is an d-connecting path from X_1 to X_4 given $\mathbf{S} = \varnothing$. This represents the fact that, knowing whether there was a genetic modification (X_1) tells us about the weight of Cheddar, through its effect on the weight of Mickey. However, γ is not an d-connecting path given $\mathbf{S} = \{X_2\}$, since X_2 is a blocked non-collider.* this happens in pset 1 pb1

(b) *$\gamma = X_2 \to X_4 \leftarrow X_3$ is not an d-connected path from X_2 to X_4 given $\mathbf{S} = \{X_1\}$, since X_4 is a blocked collider. Indeed, X_2 and X_4 are d-separated given X_1. This represents that, if we know whether there was a genetic modification, then knowing Mickey's weight tells us nothing about Minnie's weight.* isn't it X3?

(c) *However, $\gamma = X_2 \to X_4 \leftarrow X_3$ is an d-connecting path given $\mathbf{S} = \{X_1, X_4\}$, since X_4 is unblocked. This represents the fact that, if we also know Cheddar's weight, then Mickey and Minnie's weights are again related. For example, suppose we know that Cheddar has a high weight. If we find that Minnie has a low weight, this means it is more likely that Mickey has a high weight to account for Cheddar's weight. This form of reasoning is commonly called* explaining away*.*

(c) *Similarly, $\gamma = X_2 \to X_4 \leftarrow X_3$ is an d-connecting path given $\mathbf{S} = \{X_1, X_5\}$: if we know Gouda's weight instead of Cheddar's, then Mickey and Minnie's weights are still related.*

2.5 Separation implies conditional independence in directed graphs

Finally, we are ready to show that factorization according to a DAG implies Markovianity.

Definition 2.13. *Let $\mathcal{G}$ be a DAG on nodes $\mathcal{X}$ and $\mathbf{V} \subseteq \mathcal{X}$. The* **induced subgraph** *of $\mathcal{G}$ on $\mathbf{V}$, denoted $\mathcal{G}[\mathbf{V}]$, is the DAG with nodes $\mathbf{V}$ and an edge $X_i \to X_j$ if $X_i \to X_j$ is an edge in $\mathcal{G}$ and $X_i, X_j \in \mathbf{V}$.*

Example 2.6. *Let $\mathcal{G}$ be the DAG in Figure 2.3(a). Let $\mathbf{V} = \{X_1, X_2, X_3\}$. Then $\mathcal{G}[\mathbf{V}]$ is the graph in Figure 2.3(c).*

Theorem 2.2. *Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{S}$ be disjoint subsets of nodes in a DAG $\mathcal{G}$. Let $\mathbf{V} = \overline{\text{an}}_{\mathcal{G}}(\mathbf{A} \cup \mathbf{B} \cup \mathbf{S})$. Then $\mathbf{A} \perp\!\!\!\perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{S}$ implies $\mathbf{A} \perp\!\!\!\perp_{\overline{\mathcal{G}[\mathbf{V}]}} \mathbf{B} \mid \mathbf{S}$. In particular, let*

what does this bar mean - a. "moral graph"

$$\mathcal{I}^m_{\perp\!\!\!\perp}(\mathcal{G}) := \{(\mathbf{A}, \mathbf{B}, \mathbf{S}) : (\mathbf{A}, \mathbf{B}, \mathbf{S}) \in \mathcal{I}_{\perp\!\!\!\perp}(\overline{\mathcal{G}[\mathbf{V}]}), \mathbf{V} = \overline{\text{an}}_{\mathcal{G}}(\mathbf{A} \cup \mathbf{B} \cup \mathbf{S})\}$$

Then $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) \subseteq \mathcal{I}^m_{\perp\!\!\!\perp}(\mathcal{G})$.

what is the advantage of keep getting back and forth between DAG and UDG?

Proof. We prove the contrapositive: if $(\mathbf{A}, \mathbf{B}, \mathbf{S}) \notin \mathcal{I}^m_{\perp\!\!\!\perp}(\mathcal{G})$, then $(\mathbf{A}, \mathbf{B}, \mathbf{S}) \notin \mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G})$. Suppose $\mathbf{A}$ and $\mathbf{B}$ are not separated by $\mathbf{C}$ in $\overline{\mathcal{G}[\mathbf{V}]}$, i.e. there is a path γ from $\mathbf{A}$ to $\mathbf{B}$ that does not go through $\mathbf{C}$. We will use this path to show the existence of a corresponding d-connecting path in $\mathcal{G}$.

First, note that we may use any edge $X_i - X_j$ not belonging to a marriage, since these edges are in $\mathcal{G}$ and both X_i and X_j must not be in $\mathbf{S}$. For each marriage $X_i - X_j$ in γ, we have $X_i \to X_k \leftarrow X_j$ for some X_k. We have two cases:

- X_k is unblocked, i.e., $\overline{\text{de}}_{\mathcal{G}}(X_k) \cap \mathbf{S} \neq \varnothing$. Then we can replace $X_i - X_j$ with $X_i \to X_k \leftarrow X_j$.
- Otherwise, X_k must be an ancestor of some node a in $\mathbf{A}$ or $\mathbf{B}$. Then we can replace edge $X_i - X_j$ and all following edges in γ with the path $X_i \to X_k \to \ldots \to a$.

□

Example 2.7. *We illustrate the construction from Theorem 2.2 in Figure 2.4.*

Remark 2.4. *The converse of Theorem 2.2 also holds, so we actually have $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) = \mathcal{I}^m_{\perp\!\!\!\perp}(\mathcal{G})$. You will prove the converse in the first problem set. For reference, see Proposition 3.25 of Lauritzen (1996).*

Definition 2.14. *Let $\mathcal{G}$ be a DAG on nodes $\mathcal{X}$ and $\mathbb{P}_{\mathcal{X}}$ be a distribution. We say that $\mathbb{P}_{\mathcal{X}}$ is* **Markov** *with respect to $\mathcal{G}$ if $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) \subseteq \mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$.*

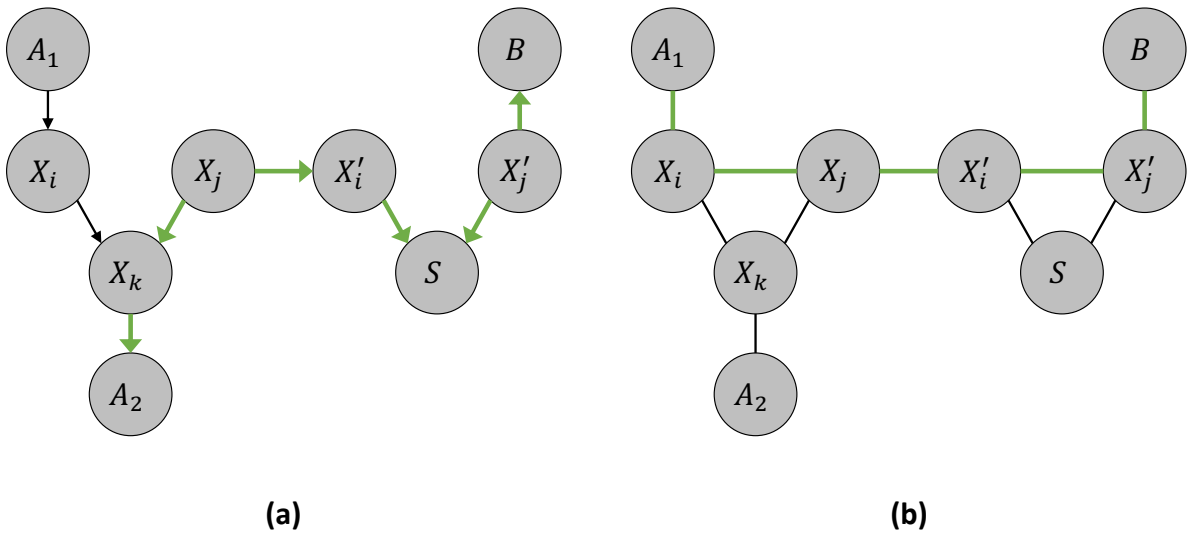

Figure 2.4: An illustration of the d-connecting path construction from Theorem 2.2. **(a)** The DAG $\mathcal{G}[\mathbf{V}]$ with $\mathbf{A} = \{A_1, \mathbf{A}_2\}$, $\mathbf{B} = \{B\}$, and $\mathbf{S} = \{S\}$. **(b)** The moral graph $\overline{\mathcal{G}[\mathbf{V}]}$ with the connecting path $A - X_i - X_j - X_i' - X_j'$. The marriage $X_i' - X_j'$ can be replace by $X_i' \to S \leftarrow X_j'$, since $S \in \mathbf{S}$ and therefore the new path is not blocked at S. The marriage $X_i - X_j$ can be replaced by the path $X_j \to X_k \to A_2$.

Theorem 2.3. *Let $\mathbb{P}_{\mathcal{X}}$ factorize according to a DAG $\mathcal{G}$. Then $\mathbb{P}_{\mathcal{X}}$ is Markov to $\mathcal{G}$, i.e., every d-separation in $\mathcal{G}$ entails a conditional independence in $\mathbb{P}_{\mathcal{X}}$.*

Proof. Let $\mathbf{A} \perp\!\!\!\perp_{\mathcal{G}} \mathbf{B} \mid \mathbf{S}$ and $\mathbf{V} = \overline{\mathrm{an}}_{\mathcal{G}}(\mathbf{A} \cup \mathbf{B} \cup \mathbf{S})$.

Step 1: Marginalize.
First, $\mathbb{P}_{\mathbf{V}}$ factorizes according to $\mathcal{G}[\mathbf{V}]$.

Step 2: Factorization in moral graph.
By Lemma 2.1, $\mathbb{P}_{\mathbf{V}}$ factorizes according to $\overline{\mathcal{G}[\mathbf{V}]}$.

Step 3: Factorization implies Markovianity in undirected graphs.
By Proposition 2.1, $\mathbb{P}_{\mathbf{V}}$ is Markov with respect to $\overline{\mathcal{G}[\mathbf{V}]}$, i.e., $\mathcal{I}_{\perp\!\!\!\perp}(\overline{\mathcal{G}[\mathbf{V}]}) \subseteq \mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathbf{V}})$. Since $\mathbf{A}, \mathbf{B}, \mathbf{S}$ are arbitrary, we have $\mathcal{I}^m_{\perp\!\!\!\perp}(\mathcal{G}) \subseteq \mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$.

Step 4: Equivalence with d-separation.
By Theorem 2.2, we have $\mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G}) \subseteq \mathcal{I}_{\perp\!\!\!\perp}(\mathbb{P}_{\mathcal{X}})$. □

Remark 2.5. *As was the case for undirected graphs, the converse of Theorem 2.3 also holds: if $\mathbb{P}_{\mathcal{X}}$ is Markov to $\mathcal{G}$, then $\mathbb{P}_{\mathcal{X}}$ factorizes according to $\mathcal{G}$. Again, we will not need this fact and do not prove it here. See Theorem 3.27 of Lauritzen (1996).*

2.6 Non-parametric identification formulas

In this section, we will discuss two well-known formulas for identifying causal effects in the non-parametric setting. We will use the following result:

Theorem 2.4 (Action-Observation Exchange)**.** *Let I be an intervention with $\mathcal{X}(I) = \mathbf{A}$. Suppose $\zeta^I \perp\!\!\!\perp_{\mathcal{G}^I} Y \mid \mathbf{S}, \mathbf{A}$. Then*

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{S} = \mathbf{s}, \mathtt{do}(\mathbf{A} = \mathbf{a})) = \mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{S} = \mathbf{s}, \mathbf{A} = \mathbf{a}).$$

Proof. $\mathbb{P}_{\mathcal{X}^I}$ factorizes according to $\mathcal{G}^I$. By Claim 1.2,

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{S} = \mathbf{s}, \texttt{do}(\mathbf{A} = \mathbf{a})) = \mathbb{P}_{\mathcal{X}^I}(Y \mid \mathbf{S} = \mathbf{s}, \zeta^I = 1)$$

By consistency and Theorem 2.3, we have

$$\begin{aligned}\mathbb{P}_{\mathcal{X}^I}(Y \mid \mathbf{S} = \mathbf{s}, \zeta^I = 1) &= \mathbb{P}_{\mathcal{X}^I}(Y \mid \mathbf{S} = \mathbf{s}, \mathbf{A} = \mathbf{a}, \zeta^I = 1) \\ &= \mathbb{P}_{\mathcal{X}^I}(Y \mid \mathbf{S} = \mathbf{s}, \mathbf{A} = \mathbf{a}) \\ &= \mathbb{P}^I_{\mathcal{X}}(Y \mid \mathbf{S} = \mathbf{s}, \mathbf{A} = \mathbf{a})\end{aligned}$$

Proving the result. □

Theorem 2.5. *Let I be an intervention with $\mathcal{X}(I) = \mathbf{A}$. Suppose $\mathbf{V} \perp\!\!\!\perp_{\mathcal{G}^I} \zeta^I \mid \mathbf{S}$. Then*

$$\mathbb{P}_{\mathcal{X}}(\mathbf{V} \mid \texttt{do}(\mathbf{A} = \mathbf{a}), \mathbf{S}) = \mathbb{P}_{\mathcal{X}}(\mathbf{V} \mid \mathbf{S})$$

Proof. First, by Claim 1.2,

$$\mathbb{P}_{\mathcal{X}}(\mathbf{V} \mid \texttt{do}(\mathbf{A} = \mathbf{a}), \mathbf{S} = \mathbf{s}) = \mathbb{P}_{\mathcal{X}^I}(\mathbf{V} \mid \zeta^I = 1, \mathbf{S} = \mathbf{s})$$

By Theorem 2.3, we have

$$\mathbb{P}_{\mathcal{X}^I}(\mathbf{V} \mid \zeta^I = 1, \mathbf{S} = \mathbf{s}) = \mathbb{P}_{\mathcal{X}^I}(\mathbf{V} \mid \zeta^I = 0, \mathbf{S} = \mathbf{s}) = \mathbb{P}_{\mathcal{X}}(\mathbf{V} \mid \mathbf{S} = \mathbf{s})$$

which gives the desired result. □

Definition 2.15. *We say that* $\mathbf{S}$ *satisfies the* **backdoor criterion** *for* $\mathbb{P}(Y \mid \texttt{do}(A = a))$ *if* $\zeta^I \perp\!\!\!\perp_{\mathcal{G}^I} Y \mid \mathbf{S}, A$ *and* $\mathbf{S} \perp\!\!\!\perp_{\mathcal{G}^I} \zeta^I$.

Remark 2.6. *The backdoor criterion can be interpreted as follows:*

- $\zeta^I \perp\!\!\!\perp_{\mathcal{G}^I} Y \mid \mathbf{S}, A$ *says that* $\mathbf{S}$ *blocks all* **backdoor paths** *from A to Y, i.e., paths with arrows into both A and Y. All paths from ζ^I to Y with edges out of A will be blocked by conditioning on A, but conditioning on A unblocks paths with edges into A. The backdoor criterion requires that* $\mathbf{S}$ *blocks these paths.*
- $\mathbf{S} \perp\!\!\!\perp_{\mathcal{G}^I} \zeta^I$ *says that* $\mathbf{S}$ *contains no descendants of A, since ζ^I is only d-connected to A and its descendants.*

Theorem 2.6 (Backdoor Adjustment)**.** *Suppose that* $\mathbf{S}$ *satisfies the backdoor criterion for* $\mathbb{P}(Y \mid \texttt{do}(A = a))$. *Then*

$$\mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(A = a)) = \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid A = a, \mathbf{S} = \mathbf{s}) \mathbb{P}_{\mathcal{X}}(\mathbf{S} = \mathbf{s})$$

Proof. We have by the law of total probability that

$$\mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(A = a)) = \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(A = a), \mathbf{S} = \mathbf{s}) \mathbb{P}_{\mathcal{X}}(\mathbf{S} = \mathbf{s} \mid \texttt{do}(A = a))$$

By Theorem 2.4, we have that

$$= \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid A = a, \mathbf{S} = \mathbf{s}) \mathbb{P}_{\mathcal{X}}(\mathbf{S} = \mathbf{s} \mid \texttt{do}(A = a))$$

By Theorem 2.5,

$$= \sum_{\mathbf{s}} \mathbb{P}_{\mathcal{X}}(Y \mid A = a, \mathbf{S} = \mathbf{s}) \mathbb{P}_{\mathcal{X}}(\mathbf{S} = \mathbf{s})$$

□

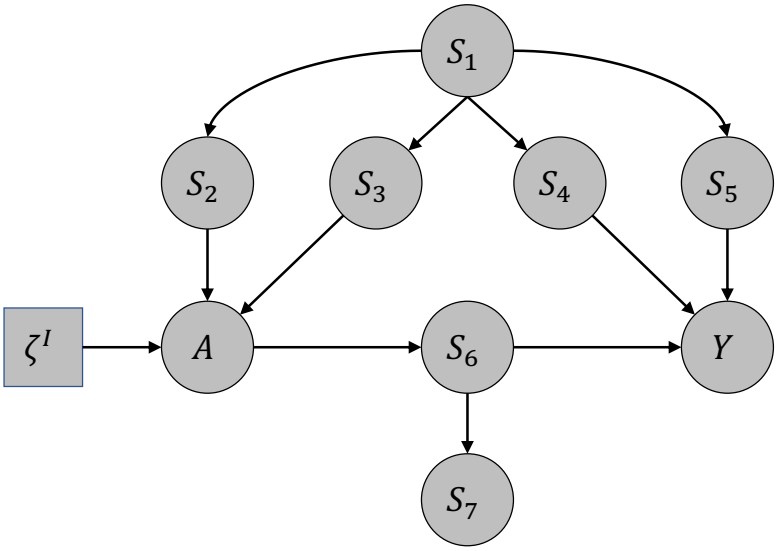

Figure 2.5: Graph for Example 2.8

Example 2.8. *Let $\mathcal{G}$ be the graph in Figure 2.5. Sets that satisfy the backdoor criterion include $\{S_2, S_3\}$, $\{S_4, S_5\}$, and $\{S_1\}$, among others. However, adding S_6 or S_7 to any of these sets will result in a violation of the backdoor criterion, since S_6 and S_7 are d-connected to ζ^I.*

Theorem 2.7 (Frontdoor Adjustment)**.** *Suppose that $\mathbb{P}_{\mathcal{X}}$ is the entailed distribution for a structural causal model with causal graph $\mathcal{G}$ in Figure 2.6. Then*

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(A = a)) = \sum_{\mathbf{m}} \mathbb{P}_{\mathcal{X}}(\mathbf{M} = \mathbf{m} \mid A = a) \sum_{a'} (\mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{M} = \mathbf{m}, A = a')\mathbb{P}_{\mathcal{X}}(A = a'))$$

We first prove the following:

Claim 2.3.

$$\mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{M} = \mathbf{m}, \mathtt{do}(A = a)) = \mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(\mathbf{M} = \mathbf{m})) \tag{2.2}$$

Proof. Let I_2 be an intervention on $\mathbf{M}$. Then, since $\zeta^{I_2} \perp\!\!\!\perp_{(\mathcal{G}^I)^{I_2}} Y \mid \mathbf{M}, A$, we have by Theorem 2.4 that

$$\begin{aligned}
\mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{M} = \mathbf{m}, \mathtt{do}(A = a)) &= \mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{M} = \mathbf{m}, A = a, \mathtt{do}(A = a)) \\
&= \mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(\mathbf{M} = \mathbf{m}), A = a, \mathtt{do}(A = a)) \\
&= \mathbb{P}_{\mathcal{X}}(Y \mid \mathtt{do}(\mathbf{M} = \mathbf{m}), \mathtt{do}(A = a))
\end{aligned}$$

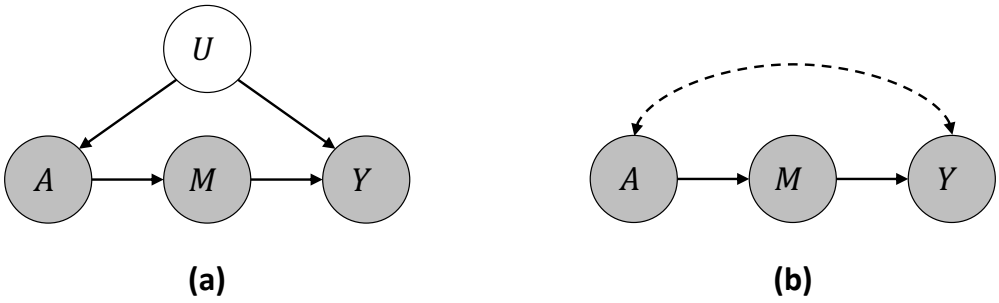

Figure 2.6: "Frontdoor" graph for Theorem 2.7

Since $Y \perp\!\!\!\perp_{(\mathcal{G}^I)^{I_2}} \zeta^I \mid \mathbf{M}$, we have by Theorem 2.5 that

$$= \mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(\mathbf{M} = \mathbf{m}))$$

□

Proof of Theorem 2.7. We have by the law of total probability that

$$\mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(A = a)) = \sum_{\mathbf{m}} \mathbb{P}_{\mathcal{X}}(\mathbf{M} = \mathbf{m} \mid \texttt{do}(A = a)) \mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{M} = \mathbf{m}, \texttt{do}(A = a))$$

Since M is not intervened, by consistency we have

$$= \sum_{\mathbf{m}} \mathbb{P}_{\mathcal{X}}(\mathbf{M} = \mathbf{m} \mid A = a) \mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{M} = \mathbf{m}, \texttt{do}(A = a))$$

Using Equation (2.2),

$$= \sum_{\mathbf{m}} \mathbb{P}_{\mathcal{X}}(\mathbf{M} = \mathbf{m} \mid A = a) \mathbb{P}_{\mathcal{X}}(Y \mid \texttt{do}(\mathbf{M} = \mathbf{m}))$$

Now, A is a backdoor adjustment set for $\mathbb{P}(Y \mid \texttt{do}(M = m))$, so we have the result. □

2.7 Additional Reading

- **The do-calculus and the ID algorithm**: Theorem 2.4 and Theorem 2.5 are special cases of the three rules of the **do-calculus** (Pearl, 1995). The do-calculus is a *complete* set of rules for determining identifiability from observational data: if a distribution $\mathbb{P}_{\mathcal{X}}(Y \mid \mathbf{S} = \mathbf{s}, \texttt{do}(A = a))$ is identifiable, then the rules of the do-calculus can be applied to transform the interventional distribution into an expression which involves only observational distributions, called an **identification formula**. Shpitser and Pearl (2006) describes a complete algorithm to conduct this transformation: the algorithm either outputs an identification formula, or provides a certificate to show that the effect is not identifiable.

- **Counterfactual identification**: Many questions, especially those involving fairness, are best framed in terms of counterfactuals instead of interventions. Shpitser and Pearl (2008) provides a method for identifying counterfactual queries, and Malinsky et al. (2019) uses single world intervention graphs to define an analogue of the do-calculus called the potential outcome calculus or the **po-calculus**.

- **Identification from interventional data**: We have been considering identification from observational data alone. However, one may consider identifying the effect of an intervention from data where different interventions have taken place. This form of identification is considered for do-interventions in Lee et al. (2020) and for soft interventions in Correa and Bareinboim (2020).

- **Partial identification**: While an interventional query might not be identifiable, one may still be able to place *bounds* on it. This is particularly important if one wishes to determine whether a treatment effect is positive, i.e. one treatment does better than another. Deriving such bounds is called **partial identification**, see Richardson et al. (2014) for a review.

Bibliography

Correa, J. and Bareinboim, E. (2020). A calculus for stochastic interventions: Causal effect identification and surrogate experiments. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 10093–10100.

Lauritzen, S. L. (1996). *Graphical models*, volume 17. Clarendon Press.

Lee, S., Correa, J. D., and Bareinboim, E. (2020). General identifiability with arbitrary surrogate experiments. In *Uncertainty in artificial intelligence*, pages 389–398. PMLR.

Malinsky, D., Shpitser, I., and Richardson, T. (2019). A potential outcomes calculus for identifying conditional path-specific effects. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 3080–3088. PMLR.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4):669–688.

Richardson, A., Hudgens, M. G., Gilbert, P. B., and Fine, J. P. (2014). Nonparametric bounds and sensitivity analysis of treatment effects. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 29(4):596.

Shpitser, I. and Pearl, J. (2006). Identification of conditional interventional distributions. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*, pages 437–444.

Shpitser, I. and Pearl, J. (2008). Complete identification methods for the causal hierarchy. *Journal of Machine Learning Research*, 9:1941–1979.