

Chapter 6: Graphical Representation of Causal Effects

Contents

1	6.1 Causal Diagrams (pp. 59-61)	1
1.1	Basic DAG Notation	2
1.2	Example: Smoking, Exercise, and Heart Disease	2
2	6.2 Causal Diagrams and Marginal Independence (pp. 61-64)	3
2.1	Independence from DAGs	3
3	6.3 Causal Diagrams and Conditional Independence (pp. 64-68)	3
3.1	Three Basic Structures	3
3.1.1	1. Chain: $X \rightarrow Z \rightarrow Y$	3
3.1.2	2. Fork: $X \leftarrow Z \rightarrow Y$	3
3.1.3	3. Collider: $X \rightarrow Z \leftarrow Y$	4
4	6.4 Positivity and Consistency in Causal Diagrams (pp. 68-69)	4
4.1	Positivity in DAGs	4
4.2	Consistency in DAGs	5
5	6.5 A Structural Classification of Bias (pp. 69-73)	5
5.1	Confounding	5
5.2	The Backdoor Criterion	5
5.3	Selection Bias	6
5.4	Measurement Bias	6
6	6.6 The Structure of Effect Modification (pp. 73-76)	6
6.1	Effect Modification in DAGs	6
6.2	Confounding vs. Effect Modification	7
7	Summary	7
8	References	8

So far, we have represented causal relationships using mathematical notation and counterfactual outcomes. This chapter introduces **causal diagrams**—visual tools for representing causal assumptions and determining which variables need to be adjusted for when estimating causal effects.

Causal diagrams are **directed acyclic graphs (DAGs)** that encode our knowledge about the causal structure of the problem. They provide an intuitive way to identify confounding, avoid selection bias, and select appropriate adjustment sets.

This chapter is based on Hernán and Robins (2020, chap. 6, pp. 59-76).

1 6.1 Causal Diagrams (pp. 59-61)

A **causal diagram** is a graph where:

- **Nodes** (vertices) represent variables

- **Directed edges** (arrows) represent direct causal effects

Definition 1.1 (Directed Acyclic Graph (DAG)). A **directed acyclic graph** is a causal diagram where: 1. All edges are directed (have arrowheads indicating the direction of causation) 2. There are no cycles (you cannot start at a node, follow the arrows, and return to that node)

1.1 Basic DAG Notation

Consider a simple causal diagram:

$L \rightarrow A \rightarrow Y$

$L \rightarrow Y$

This diagram represents:

- L causes A (arrow from L to A)
- L causes Y (arrow from L to Y)
- A causes Y (arrow from A to Y)

Key terms:

- L is a **parent** of A and Y (direct cause)
- A is a **child** of L (direct effect)
- L is an **ancestor** of Y (can reach Y by following arrows)
- Y is a **descendant** of L (can be reached from L by following arrows)

What does a DAG represent?

A DAG represents our **qualitative causal assumptions**:

- The presence of an arrow $X \rightarrow Y$ means we believe X has a direct causal effect on Y
- The absence of an arrow $X \nrightarrow Y$ means we believe X has no direct causal effect on Y (conditional on other variables in the graph)

Important: A DAG does not tell us the magnitude or sign of effects, only their presence/absence and direction.

Acyclic assumption: We assume no feedback loops (e.g., A causes Y , which causes A). This rules out dynamic systems with bidirectional causation. Extensions exist for such settings (e.g., structural equation models with cycles) but are beyond this chapter's scope.

1.2 Example: Smoking, Exercise, and Heart Disease

Example 1.1 (Simple Causal Diagram). Consider the causal relationships between:

- A : Smoking
- L : Exercise
- Y : Heart disease

Plausible causal diagram:

$A \rightarrow Y$

$L \rightarrow Y$

$L \rightarrow A$

Interpretation:

- Smoking directly causes heart disease
- Exercise directly affects heart disease (protective)
- Exercise affects smoking behavior (perhaps exercisers are less likely to smoke)

2 6.2 Causal Diagrams and Marginal Independence (pp. 61-64)

Causal diagrams encode information about statistical independence relationships.

2.1 Independence from DAGs

From a DAG, we can determine which variables are marginally independent (unconditionally).

Example 2.1 (Marginal Independence). **Diagram 1:** $A \rightarrow Y$

- A and Y are **dependent** (associated)

Diagram 2: $A \quad Y$ (no arrow)

- A and Y are **independent**

Diagram 3: $L \rightarrow A \rightarrow Y$ (chain)

- A and Y are **dependent**
- L and Y are **dependent** (through the path $L \rightarrow A \rightarrow Y$)

Diagram 4: $A \rightarrow Y \leftarrow L$ (fork)

- A and L are **independent** (no common cause, no direct connection)
- Both A and L are **dependent** with Y

Key principle: In a DAG, two variables are marginally associated if and only if there exists a path (sequence of arrows, ignoring direction) connecting them, UNLESS the path is blocked.

Blocking: A path is blocked if it contains a **collider**—a variable where two arrowheads meet (e.g., $A \rightarrow C \leftarrow B$). Paths through colliders are naturally blocked.

Later sections will formalize this with **d-separation**.

3 6.3 Causal Diagrams and Conditional Independence (pp. 64-68)

Conditioning on (stratifying by) a variable can change independence relationships.

3.1 Three Basic Structures

There are three fundamental structures in causal diagrams:

3.1.1 1. Chain: $X \rightarrow Z \rightarrow Y$

- **Marginal:** X and Y are **dependent** (path $X \rightarrow Z \rightarrow Y$)
- **Conditional on Z :** X and Y are **independent**
 - Knowing Z blocks the path from X to Y

3.1.2 2. Fork: $X \leftarrow Z \rightarrow Y$

- **Marginal:** X and Y are **dependent** (common cause Z)
- **Conditional on Z :** X and Y are **independent**
 - Controlling for the common cause blocks the association

3.1.3 3. Collider: $X \rightarrow Z \leftarrow Y$

- **Marginal:** X and Y are **independent** (no connecting path)
- **Conditional on Z :** X and Y are **dependent!**
 - Conditioning on a collider creates association (**collider bias**)

Example 3.1 (Collider Bias Example). Suppose:

- A : Natural athletic ability
- E : Training effort
- Y : Professional athlete (yes/no)

Diagram: $A \rightarrow Y \leftarrow E$

Marginal: Among the general population, natural ability and training effort are independent (some people train hard, some don't; some are naturally gifted, some aren't; these are unrelated).

Conditional on $Y = 1$ (among professional athletes): Natural ability and training effort are **negatively** associated!

Why? Among those who made it to the pros, if someone has low natural ability, they must have compensated with high training effort. Conversely, high natural ability allows one to reach the pros with less effort.

This is **collider stratification bias**.

Collider bias is subtle and dangerous:

1. **Intuition:** Conditioning on a collider induces a spurious association between its causes, even when those causes are truly independent
2. **Selection bias:** When the collider is related to study selection, we get selection bias (Chapter 8)
3. **Measurement:** Even conditioning on a descendant of a collider can induce bias (though weaker)

Practical implication: Do NOT adjust for variables that are purely effects of both treatment and outcome. This will induce bias, not remove it.

Example: In studying the effect of exercise on weight loss, do NOT adjust for “membership in a weight loss support group” if that’s an effect of both exercise and weight. Doing so would induce collider bias.

4 6.4 Positivity and Consistency in Causal Diagrams (pp. 68-69)

Causal diagrams help us understand the identifiability assumptions introduced in Chapter 3.

4.1 Positivity in DAGs

The **positivity assumption** requires that all levels of treatment occur at all levels of confounders:

$$\Pr[A = a|L = l] > 0 \quad \text{for all } a, l$$

In a DAG, positivity violations can occur when:

- Structural determinism: Certain values of L make A deterministic (e.g., if L completely determines A)
- Random violations: By chance, no individuals in the sample have certain (A, L) combinations

DAGs and positivity: The DAG itself doesn't tell us whether positivity holds—that depends on the actual data. However, DAGs can help identify situations where positivity is likely to be violated (e.g., when many arrows point into the treatment, suggesting treatment is highly determined by covariates).

4.2 Consistency in DAGs

The **consistency assumption** requires well-defined interventions:

$$Y = Y^A$$

In DAGs, consistency requires: 1. **No treatment variation:** All individuals receiving $A = a$ receive exactly the same version of treatment 2. **No interference:** One individual's treatment doesn't affect another's outcome

DAGs and consistency:

Treatment variation: The node A in the DAG must represent a single, well-defined intervention. If “treatment” actually consists of multiple versions (e.g., different exercise types), the DAG should have separate nodes for each.

Interference: Standard DAGs assume no interference. When interference exists, we need extended graphical frameworks or explicit modeling of dependencies between individuals.

5 6.5 A Structural Classification of Bias (pp. 69-73)

Causal diagrams allow us to classify different types of bias based on graph structure.

5.1 Confounding

Definition 5.1 (Confounding (DAG Definition)). **Confounding** occurs when there exists a **backdoor path** from treatment A to outcome Y :

A backdoor path is a path from A to Y that: 1. Starts with an arrow pointing *into* A (i.e., $\cdot \rightarrow A$) 2. Does not pass through any descendants of A

Example:

$L \rightarrow A \rightarrow Y$
 $L \rightarrow Y$

The path $A \leftarrow L \rightarrow Y$ is a backdoor path (it starts with an arrow into A). Therefore, L is a confounder.

Why “backdoor”?

The term “backdoor path” refers to paths that go into the back of A (against the arrow). These paths allow non-causal association to flow from A to Y via common causes.

Contrast with causal paths: Causal paths follow arrows in the forward direction ($A \rightarrow \dots \rightarrow Y$).

Confounding = open backdoor paths: If backdoor paths are open (not blocked), the association between A and Y reflects both causal and non-causal (confounded) pathways.

5.2 The Backdoor Criterion

Definition 5.2 (Backdoor Criterion). A set of variables L satisfies the **backdoor criterion** for the effect of A on Y if:

1. No variable in L is a descendant of A
2. L blocks all backdoor paths from A to Y

If L satisfies the backdoor criterion, adjusting for L eliminates confounding.

Example 5.1 (Backdoor Criterion Example). **Diagram:**

$$\begin{array}{l} U \rightarrow L \rightarrow A \rightarrow Y \\ \quad \quad L \rightarrow Y \end{array}$$

Backdoor paths from A to Y :

- $A \leftarrow L \rightarrow Y$

Does L satisfy the backdoor criterion? 1. L is not a descendant of A 2. L blocks the backdoor path $A \leftarrow L \rightarrow Y$

Yes! Adjusting for L is sufficient to identify the causal effect.

Does U satisfy the backdoor criterion?

- No. U is not on the path $A \leftarrow L \rightarrow Y$, so it doesn't block the backdoor path.

Multiple backdoor paths: In more complex diagrams, there may be many backdoor paths. The adjustment set must block **all** of them.

Minimal adjustment sets: Often, multiple sets of variables satisfy the backdoor criterion. The smallest such set is called a **minimal sufficient adjustment set**. Software like `dagitty` can find these automatically.

Over-adjustment: Including unnecessary variables in L (e.g., mediators, colliders) can introduce bias even if they're not needed for confounding control.

5.3 Selection Bias

Selection bias occurs when we condition on a collider (or its descendant) that lies on a path from A to Y .

Example:

$$A \rightarrow S \leftarrow Y$$

If we restrict analysis to individuals with $S = 1$, we induce spurious association between A and Y (collider bias).

This will be covered in detail in Chapter 8.

5.4 Measurement Bias

Measurement bias can be represented in DAGs by including nodes for both:

- True (unmeasured) variables
- Measured (error-prone) variables

Example:

$$\begin{array}{l} A_{\text{true}} \rightarrow Y \\ A_{\text{measured}} \leftarrow A_{\text{true}} \end{array}$$

If we use A_{measured} instead of A_{true} , the estimated effect will be biased (Chapter 9).

6 6.6 The Structure of Effect Modification (pp. 73-76)

Effect modification can also be represented in causal diagrams.

6.1 Effect Modification in DAGs

Effect modification by V means the effect of A on Y differs across levels of V . This can be represented by:

$V \rightarrow Y$

$A \rightarrow Y$

(with the understanding that the $A \rightarrow Y$ effect depends on V)

Some authors include an arrow $V \rightarrow A \cdot Y$ to explicitly denote interaction, but this is not standard DAG notation.

DAGs and effect modification:

Standard DAGs do not have explicit notation for effect modification or interaction. The diagram:

$V \rightarrow Y$

$A \rightarrow Y$

is compatible with:

- No effect modification (effect of A is the same for all V)
- Effect modification (effect of A differs by V)

Why? DAGs represent **qualitative** causal structure (presence/absence of effects), not **quantitative** features (magnitude or heterogeneity).

Extensions: Some frameworks (e.g., Single World Intervention Graphs, SWIGs) can represent effect modification more explicitly, but these are advanced topics beyond the scope of this chapter.

6.2 Confounding vs. Effect Modification

A variable V can be:

- A **confounder**: Opens a backdoor path ($V \rightarrow A, V \rightarrow Y$)
- An **effect modifier**: The effect of A on Y varies by V
- Both
- Neither

Example 6.1 (Confounder and Modifier in DAGs). **Scenario 1:** V is a confounder only

$V \rightarrow A \rightarrow Y$

$V \rightarrow Y$

V opens the backdoor path $A \leftarrow V \rightarrow Y$ (confounding). We must adjust for V .

Scenario 2: V is a modifier only (in a randomized trial)

$V \rightarrow Y$

$A \rightarrow Y$

(A randomized, so no $V \rightarrow A$ arrow)

V modifies the effect of A , but does not confound (no backdoor path). We should report stratum-specific effects but don't need to adjust for V to eliminate bias.

Scenario 3: V is both

$V \rightarrow A \rightarrow Y$

$V \rightarrow Y$

(and the $A \rightarrow Y$ effect varies by V)

We must adjust for V (confounding) AND report stratum-specific effects (modification).

7 Summary

This chapter introduced **causal diagrams** (DAGs) as tools for representing and reasoning about causal relationships.

Key concepts:

1. **DAGs:** Directed acyclic graphs with nodes (variables) and directed edges (causal effects)

2. **Three basic structures:**
 - **Chain** ($X \rightarrow Z \rightarrow Y$): Conditioning on Z blocks the path
 - **Fork** ($X \leftarrow Z \rightarrow Y$): Conditioning on Z blocks the path
 - **Collider** ($X \rightarrow Z \leftarrow Y$): Conditioning on Z opens the path!
3. **Collider bias:** Conditioning on a common effect of two variables induces spurious association
4. **Backdoor paths:** Non-causal paths from A to Y that begin with an arrow into A
5. **Backdoor criterion:** A set L satisfies the backdoor criterion if:
 - No member of L is a descendant of A
 - L blocks all backdoor paths from A to Y
6. **Adjustment:** If L satisfies the backdoor criterion, adjusting for L eliminates confounding

Practical use of DAGs:

1. **Draw a DAG** based on subject-matter knowledge before analyzing data
2. **Identify backdoor paths** from treatment to outcome
3. **Find an adjustment set** that satisfies the backdoor criterion
4. **Adjust for that set** using stratification, regression, or weighting
5. **Avoid adjusting** for colliders, mediators, or other variables that induce bias

Software:

- R package `dagitty`: Draw DAGs, find adjustment sets, test implications
- R package `ggdag`: Visualize DAGs
- Web app: dagitty.net

Limitations of DAGs:

- Require strong qualitative assumptions about causal structure
- Cannot represent unmeasured confounding (unless explicitly included as a node)
- Cannot represent dynamic feedback (cycles)
- Do not encode effect modification or quantitative magnitudes

Despite these limitations, DAGs are invaluable for making causal assumptions explicit and selecting appropriate adjustment strategies.

Looking ahead:

- **Chapter 7:** Detailed treatment of confounding
- **Chapter 8:** Selection bias
- **Chapters 11-15:** Methods for adjusting for confounding (IP weighting, g-formula, propensity scores)

8 References

Hernán, Miguel A, and James M Robins. 2020. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC. <https://miguelhernan.org/whatifbook>.