

# Stat 256: Causality

## SCM+DAG 2: Interventions and Identification

Chad Hazlett

UCLA

[chazlett@ucla.edu](mailto:chazlett@ucla.edu)

Spring 2026

# Where we are and where we're going

Last time: the language of SCMs, DAGs, and  $d$ -separation.

Today: how do we use that machinery to answer causal questions?

- We'll introduce the *do*-operator and interventional graphs — the formal machinery for defining causal effects.
- Then the main event: the **backdoor criterion** — a graphical tool for determining *what to condition on* to identify a causal effect from observational data.
- Finally, a brief look at what to do when backdoor fails.

# The cardinal sin of causal inference

**The mistake:** assuming that “how  $Y$  would look if we *made*  $D = d$ ” is the same as “what we see in the data where  $D = d$  already.”

*When we intervene on a variable in a model, we fix its value. We change the system. . . When we condition on a variable, we change nothing; we merely narrow our focus to the subset of cases in which the variable takes the value we are interested in. (Primer)*

In other words:

- Conditioning: groups with  $D = d$  vs.  $D = d'$  differ on things besides  $D$
- Intervening: set  $D$  externally regardless of what  $D$  would have been observed.

We have a notation for conditioning – how can we formally indicate intervention?

# The *do*-operator

Notation for interventions:

- $P(Y \mid D = d)$ : the distribution of  $Y$  when we *saw*  $D = d$ .
- $P(Y \mid do(D = d))$ : the distribution of  $Y$  if everybody in the population had  $D$  set to  $d$ .
- Perhaps better written as  $P_{do(D=d)}(Y)$

**This is our target.** A causal effect is a contrast in this quantity, e.g.:

$$P(Y \mid do(D=1)) - P(Y \mid do(D=0))$$

Or more generally, we want  $P(Y \mid do(d))$  for all values of  $d$ .

The challenge: we observe  $P(Y \mid D = d)$ , not  $P(Y \mid do(d))$ .

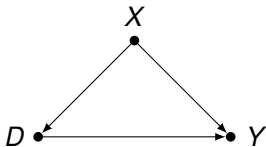
**Identification** is about bridging that gap using assumptions encoded in a DAG.

# Interventional (“mutilated”) graphs

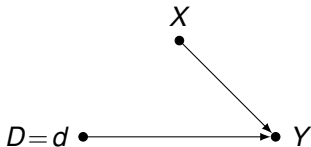
When we  $do(D = d)$ , we imagine a modified system where:

- The original links from  $D$ 's parents into  $D$  are **removed**
- $D$  is set to  $d$
- Everything else in the system is unchanged (modularity)

Running example: severity ( $X$ ), drug ( $D$ ), recovery ( $Y$ ). What happens when we fix  $D = d$ ?



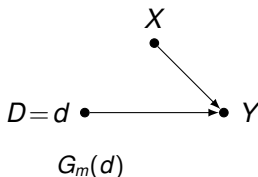
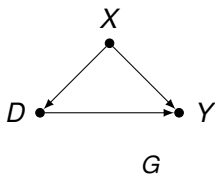
Observational:  $G$



Interventional:  $G_m(d)$

The interventional distribution  $P_m(d)$  tells us what happens under  $do(D = d)$ .  
 Identification: can we recover  $P_m(d)$  from observational data?

# From intervention to identification



$$\begin{aligned}
 P(Y=y \mid do(D=d)) &\equiv P_{m(d)}(Y=y) \quad (\text{definition: marginal of } Y \text{ in } G_m(d)) \\
 &= \sum_x P_{m(d)}(Y=y \mid X=x) P_{m(d)}(X=x) \quad (\text{LTP over } X) \\
 &= \sum_x P(Y=y \mid D=d, X=x) P(X=x) \quad (\text{invariance})
 \end{aligned}$$

*Remark on modularity/invariance:* intervention changes only  $D$ 's equation. Every other node keeps its original mechanism.  $\dots P_{m(d)} = P$  for any factor not involving  $D$ .

**Identification achieved:** only observational quantities on the RHS.

# The adjustment formula

Generalizing, when all parents of  $D$  are observed:

## Rule

**Causal effect rule / adjustment formula.** Given a graph  $G$  in which  $PA$  are the parents of  $D$ , the causal effect of  $D$  on  $Y$  is:

$$P(Y = y | do(D = d)) = \sum_z P(Y = y | D = d, PA = z) P(PA = z)$$

And we usually care about contrasts:

$$\begin{aligned} ACE &= \mathbb{E}[Y | do(D=1)] - \mathbb{E}[Y | do(D=0)] \\ &= \sum_x [P(Y=1 | D=1, X=x) - P(Y=1 | D=0, X=x)] P(X=x) \end{aligned}$$

Condition on  $X$  to “mute” it, then average over the population distribution of  $X$ .

## But we need more than this...

The adjustment formula requires observing *all parents* of the treatment.

In practice, some parents of  $D$  are often unobserved. Are we stuck?

**No.** Sometimes other observed variables can do the job.

**The question:** given a DAG (possibly with unobserved variables), *what set of observed variables should we condition on* to identify the causal effect of  $D$  on  $Y$ ?

This is the **backdoor criterion** — the central identification tool of this course.

# Frontdoor paths and backdoor paths

- A **frontdoor path** from  $D$  to  $Y$  begins with an arrow *out of*  $D$ .  
These carry the causal effect — usually we do **not** want to block these.
- A **backdoor path** from  $D$  to  $Y$  is any path that goes *into*  $D$ .  
These create spurious associations even absent any effect of  $D$  on  $Y$ .
- Both create observed associations between  $D$  and  $Y$ , but only frontdoor paths reflect causation.

**Goal:** block all backdoor paths without disrupting frontdoor paths.

# The backdoor criterion

## Definition

**Backdoor criterion:** A set of variables  $Z$  satisfies the *backdoor criterion* relative to  $(D, Y)$  if:

- 1 No node in  $Z$  is a descendant of  $D$ , and
- 2  $Z$  blocks every path between  $D$  and  $Y$  that contains an arrow *into*  $D$ .

For humans — find a set  $Z$  that:

- 1 Blocks all backdoor (spurious) paths between  $D$  and  $Y$
- 2 Leaves all directed (frontdoor/causal) paths from  $D$  to  $Y$  unblocked
- 3 Doesn't *open* new spurious paths (e.g. by conditioning on a collider or its descendants)

# What to do with a valid adjustment set

Once you find  $Z$  satisfying the backdoor criterion for  $D \rightarrow Y$ :

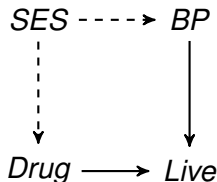
- Compute  $P(Y|do(d))$  via the adjustment formula:

$$P(Y|do(d)) = \sum_z P(Y|d, z)P(z)$$

- Or equivalently, within strata of  $Z$ , compare  $Y$  for  $D = d$  vs.  $D = d'$ , then average over  $Z$ .
- In practice: use your preferred conditioning method (regression, matching, weighting, etc.) to adjust for  $Z$ .

The DAG tells you *what to adjust for*. The choice of estimation method is a separate question.

# Example 1

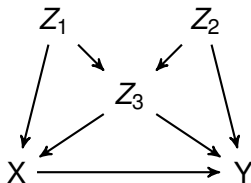


*SES* = socioeconomic status (*unobserved*), *BP* = blood pressure, *Drug* = treatment, *Live* = outcome.

- Can you adjust for parents of *Drug*? No — *SES* is unobserved.
- Backdoor paths from *Drug* to *Live*?  $Drug \leftarrow SES \rightarrow BP \rightarrow Live$
- What observed set satisfies the backdoor criterion?  $\{BP\}$
- Adjustment formula:

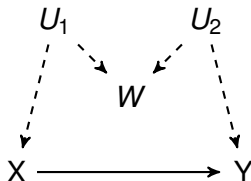
$$P(Live|do(drug)) = \sum_{bp} P(Live|drug, bp) P(bp)$$

## Example 2



- What are the backdoor paths from  $X$  to  $Y$ ?
- What observed sets block them? (There are multiple valid sets.)
- What sets would *fail*? (Think about colliders from Part 1.)

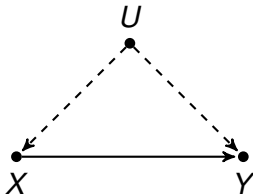
# Example 3: When conditioning hurts



- Are there any backdoor paths from  $X$  to  $Y$ ? The only candidate is  $X \leftarrow U_1 \rightarrow W \leftarrow U_2 \rightarrow Y$ , but  $W$  is a collider — so it is **already blocked**. The empty set  $\{\}$  suffices.
- What if you condition on  $W$ ? **You open** that path.  $W$  is a collider, and conditioning on it creates a spurious association between  $U_1$  and  $U_2$ .

Lesson: not conditioning on anything is sometimes the right answer — and conditioning on the wrong thing can create bias where none existed.

## Example 4: When identification fails



- Backdoor path:  $X \leftarrow U \rightarrow Y$
- Can we block it?  $U$  is unobserved. No other observed variables to condition on.
- No set of observed variables satisfies the backdoor criterion.

This is important: not every causal effect is identifiable. The graph tells you when you're stuck, not just when you're fine.

(Though there may be other ID strategies depending on additional structure.)

# Practicing the backdoor criterion

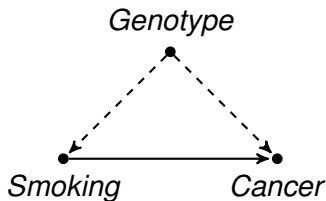
Summary of the skill:

- 1 Identify all backdoor paths from treatment to outcome  
(paths that start with an arrow *into* the treatment)
- 2 Find a set of observed variables that blocks all of them  
(using  $d$ -separation rules from Part 1)
- 3 Check you haven't included descendants of treatment  
(which could block causal paths or open new spurious ones)
- 4 Check you haven't opened new paths by conditioning on colliders

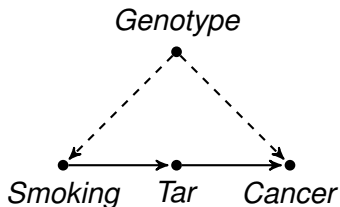
# The front-door criterion

Sometimes no observed set satisfies the backdoor criterion, yet identification is still possible.

The front-door criterion is *rarely* available, but a classic one to point out that the framework is broader than backdoor criterion.



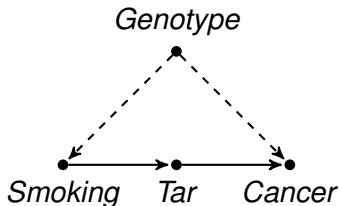
No backdoor solution



Front-door to the rescue

Any intuitions for how to solve this?

# Front-door logic



- Effect of Smoking on Tar: no backdoor paths, so  $p(t|do(s)) = p(t|s)$
- Effect of Tar on Cancer: backdoor through Genotype, but *Smoking* blocks it:

$$p(c|do(t)) = \sum_s p(c|t, s)p(s)$$

- Chain them:

$$p(c|do(s)) = \sum_t p(t|s) \sum_{s'} p(c|t, s')p(s')$$

All observational quantities — identification achieved without any backdoor

# The do-calculus: a completeness result

The backdoor and front-door criteria are the most common identification tools, but they don't exhaust all possibilities.

Pearl's **do-calculus** provides three rules that, together with probability laws, can derive *any* identifiable causal effect from a DAG.

- Rule 1: When you can ignore an observation (d-separation under intervention)
- Rule 2: When “doing” equals “seeing” (backdoor)
- Rule 3: When you can ignore an intervention (non-causation)

This is **complete**: if the do-calculus can't identify an effect from your graph, then no method can (non-parametrically).

Typically you write these rules in a very complicated way and try to apply them. We will skip that this year.

# Causal Fusion — give it a try

There is a complete algorithmic solution to find do-calc solutions when they exist. Check out:

The screenshot displays the Causal Fusion web application interface. The central canvas shows a causal diagram with three nodes:  $U_{YX}$  (unobserved confounder),  $X$  (treatment), and  $Z$  (intermediate variable), and  $Y$  (outcome). The edges are:  $U_{YX} \rightarrow X$ ,  $U_{YX} \rightarrow Y$ ,  $X \rightarrow Z$ , and  $Z \rightarrow Y$ .

The left sidebar contains a Summary panel with the following information:

- Treatment:  $X$
- Outcome:  $Y$
- Adjusted:  $P_X(Y)$
- Query:  $P_X(Y)$
- Show More Details

The Editor panel shows the graphical structure and a list of nodes and edges:

- Nodes:  $Z \rightarrow Y$ ,  $X \rightarrow Z$ ,  $U_{YX} \rightarrow Y$ ,  $U_{YX} \rightarrow X$
- Edges:  $Z \rightarrow Y$ ,  $X \rightarrow Z$ ,  $U_{YX} \rightarrow Y$ ,  $U_{YX} \rightarrow X$
- Task:  $X$
- Outcome:  $Y$
- Adjusted:  $P_X(Y)$
- Populations:  $P_X(Y)$
- Datasets

The right sidebar contains a list of analysis methods:

- Confounding Analysis
  - Admissible Sets
  - Admissibility Test
  - Instrumental Variables
  - IV Admissibility Test
- Path Analysis
  - D-Separation
  - Causal Paths
  - Confounding Paths
  - Biasing Paths
- Do-Calculus Analysis
  - e-Calculus Analysis
  - e-Inspector
  - e-Separation
- Testable Implications
  - Conditional Independencies
  - Verma Constraints
- Counterfactual Assumptions

The bottom section shows the query: "The causal effect of  $x$  on  $Y$  conditional on  $U_{YX}$  with do:  $P_X(Y)$  from  $P(V)$ ". The result is displayed as:

$$P_X(Y) = \sum_Z P(Z|X) \sum_{X'} P(Y|X', Z) P(X')$$

The result is obtained by Front-Door adjustment with an admissible set  $Z$ . The interface also includes a subgraph  $G$  and a button to load the derivation.

# Looking ahead

DAGs help us find identification opportunities (mostly by backdoor), and see mistakes or where backdoor won't be workable.

We don't yet have a notion of individual level causal effect.

The seemingly different language **potential outcomes** will (i) focus on individual units and their causal effects; (ii) provide definitions of the main causal estimands; and (iii) provide concise (but sometimes opaque) versions of identification assumptions/ requirements.

- Once we have both frameworks in hand, we'll build the formal bridge between them: when does a graphical condition (backdoor) translate into a statement about potential outcomes (conditional ignorability)?

# What's next

- Next: the potential outcomes framework — defining  $Y_i(d)$ , average effects, and the fundamental problem of causal inference
- Then: bridge between DAGs and POs
- More ID strategies and real world challenges.

