Inference & Causality

Week 4 Session 7 • October 11, 2025

Lecturer: Narges Chinichian
IU University of Applied Sciences, Berlin

Course Overview

Stay connected with the latest course materials, assignments, and announcements through our centralized course hub on Notion.

All lecture slides, readings, problem sets, and important dates are maintained in one accessible location for your convenience.

Visit Course Hub

Today's Outline

Unit 4: Do-Calculus

Front-Door and Back-Door Criteria

We learned that correlation ≠ causation, and that **interventions (do X)** tell us what happens when we *change* X on purpose.
Unit 4 now shows us how to **calculate causal effects from data**, even when we cannot actually perform experiments.

Learning Goals

01

Identify Interventions via DAG Structure

Understand how causal graphs reveal intervention opportunities and valid adjustment strategies.

02

Master Adjustment Set Conditions

Learn the formal criteria that determine when a variable set enables valid causal inference.

03

Apply Formulas to Real Examples

Practice using back-door and front-door adjustment formulas in concrete scenarios.

From Observation to Intervention

Recap from Unit 3

Observation

$P(Y \mid X)$

Correlation between variables in passive observation

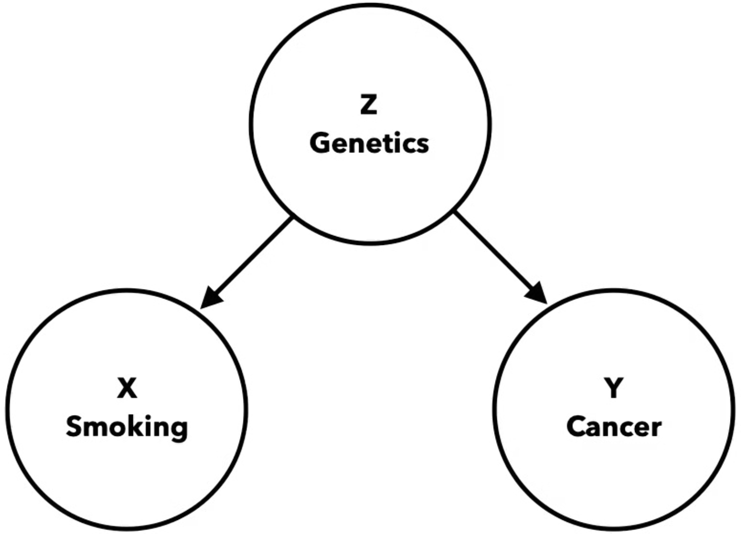
Intervention

$P(Y \mid do(X))$

Causal effect of actively forcing X to a specific value

Key Point: In observation, causal DAG stays the same and in intervention, all input arrows are blocked.

The bridge between observation and intervention requires carefully blocking confounding paths in our causal graph structure.



Blocking Paths and Confounders

Path

Any sequence of connected nodes in the graph, regardless of arrow direction. Paths can transmit associations.

Back-Door Path

A path from X to Y that *starts with an arrow pointing into X*. These create spurious associations.

Confounder

A variable that opens a back-door path between treatment and outcome, creating non-causal correlation.

> **Goal:** Find a variable set Z that blocks all back-door paths between X and Y, isolating the true causal effect.

The Back-Door Criterion

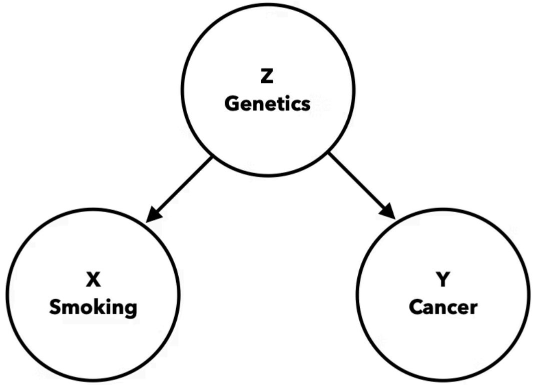
A variable set **Z** satisfies the **back-door criterion** relative to (X, Y) if two conditions hold:

1 **No Descendants**

No member of Z is a descendant of X (avoiding post-treatment bias)

2 **Complete Blocking**

Z blocks every path between X and Y that contains an arrow into X



Back-Door Solution: If genetic risk were observable, adjusting for it would block the confounding path.

Identification Formula

When these conditions are satisfied, the interventional distribution is identifiable:

$$P(Y \mid do(X)) = \sum_{z} P(Y \mid X, z)\, P(z)$$

This formula allows us to compute causal effects from purely observational data by adjusting for confounders.

Why Back-Door Adjustment Works

Simulates Randomization

Adjusting for Z mimics the balance achieved in randomized experiments

Removes Spurious Correlation

Conditioning on common causes eliminates non-causal associations

Stratification Equivalent

Mathematically equivalent to analyzing data separately within each level of Z

The adjustment process isolates the direct causal pathway from treatment to outcome by "holding constant" the confounding variables.

When Back-Door Adjustment Fails

Back-door adjustment requires identifying and measuring all confounders—but this isn't always possible. Several scenarios prevent its application:

Unobserved Confounders

When latent variable U exists between X and Y but cannot be measured or included in the analysis

Post-Treatment Bias

Incorrectly including intermediate variables (mediators) or descendants of treatment in the adjustment set

Collider Conditioning

Adjusting for variables that are effects of both treatment and outcome, which opens new spurious paths

Alternative Approach: When back-door adjustment is impossible, the **front-door criterion** provides another path to causal identification.

The Front-Door Criterion

A variable **Z** satisfies the **front-door criterion** when three conditions are met:

01

Complete Mediation

Z intercepts all directed paths from X to Y, no direct effect remains

02

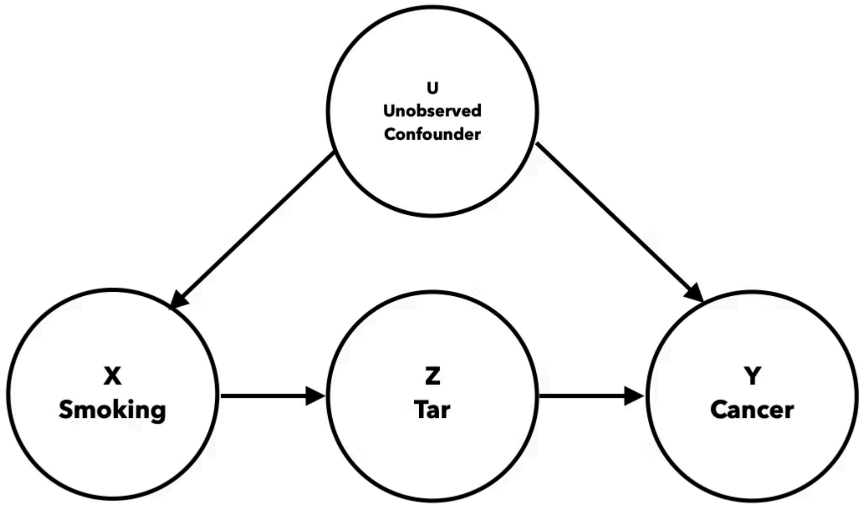
No X-Z Confounding

There are no unblocked back-door paths from X to Z

03

X Blocks Z-Y Paths

All back-door paths from Z to Y are blocked by conditioning on X



This criterion enables causal identification **even when confounders between X and Y are unobserved**. It is a powerful alternative when backdoor methods fail.

Front-Door Adjustment Formula

When variable Z satisfies all three front-door conditions, we can identify the causal effect using a two-step decomposition:

$$P(Y \mid do(X)) = \sum_z P(Z \mid X) \sum_{x'} P(Y \mid x', z) P(x')$$

Understanding the Components

First Term

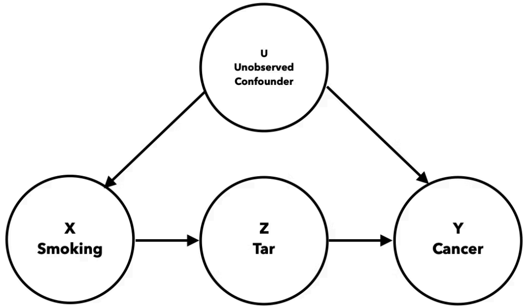
$P(Z \mid X)$

Effect of X on mediator Z—directly observable from data

Second Term

$\sum_{x'} P(Y \mid x', z) P(x')$

Effect of Z on Y, averaging over all values of X



The formula cleverly uses the mediator to "route around" unobserved confounding between X and Y.

Example: Smoking → Tar → Cancer

The Challenge

Unobserved confounder U (genetic factors) directly influences both Smoking (X) and Cancer (Y). This causal structure (U → X, X → Z → Y, U → Y) prevents standard back-door adjustment for the causal effect of X on Y. However, Tar (Z) acts as a mediator where Smoking (X) affects Tar (Z), and Tar (Z) affects Cancer (Y), fulfilling the front-door conditions to identify the causal effect.

Step-by-Step Computation

Step 1: X → Z Effect

Estimate $P(Z \mid X)$—how smoking levels (X) determine tar exposure (Z)

Step 2: Z → Y Effect

Estimate $P(Y \mid Z, X)$—how tar exposure (Z) influences cancer risk (Y), conditional on smoking (X)

Step 3: Combine

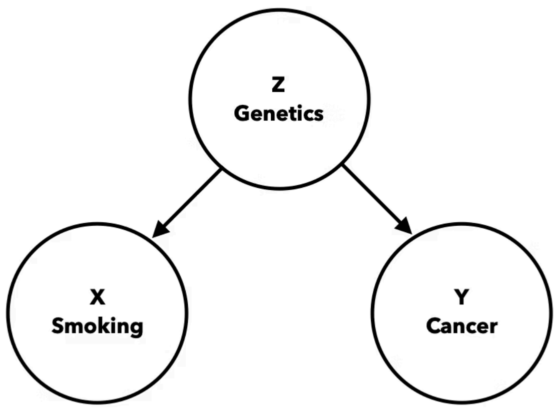
Apply the front-door formula to compute $P(Y \mid do(X))$

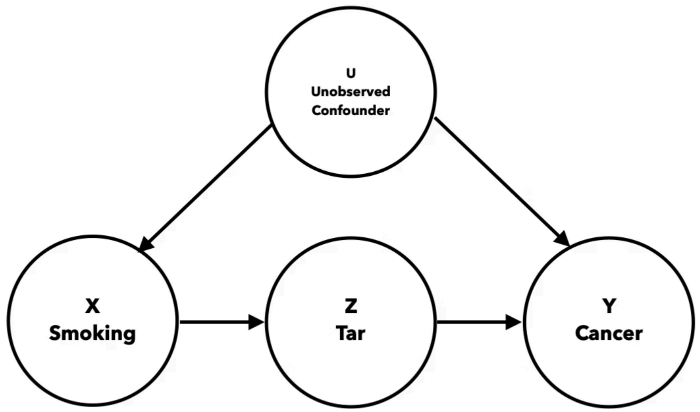
This approach successfully identifies the causal effect of smoking on cancer, even with the unmeasured genetic confounder that directly influences both smoking and cancer risk.

Comparison: Back-Door vs Front-Door

Feature	Back-Door	Front-Door
Blocks confounding	**Yes (via observed confounders)**	**Indirectly (via mediator)**
Requires all confounders observed	**Yes**	**Not necessary**
Uses mediators	**No**	**Yes**
When to use	Confounders measurable	Unmeasured confounding exists

---





Back-Door

$$\sum_z P(Y \mid X, z)P(z)$$

Front-Door

$$\sum_z P(Z \mid X) \sum_{x'} P(Y \mid x', z)P(x')$$

Practical Exercise

Group Activity (15–20 minutes)

Task 1: Draw a DAG

Create a small directed acyclic graph with variables X (treatment), Y (outcome), and Z (potential adjustment variable)

Task 2: Identify Criteria

Determine whether Z satisfies the back-door criterion, front-door criterion, both, or neither

Task 3: Derive Formula

Write out the corresponding adjustment formula based on your identification strategy

Work in small groups to discuss different scenarios. Consider how adding or removing edges changes which criterion applies.

Notebook Activities

Identifying Causal Effects

Notebook 1: Back-door Adjustment

Goal: Estimate P(Y|do(X)) by blocking confounding paths

Activities: Identify valid adjustment sets Z, apply back-door formula, simulate data verification

Learning outcome: Understanding how conditioning recovers causal effects

Notebook 2: Front-door Adjustment

Goal: Estimate P(Y|do(X)) when confounding can't be blocked directly

Activities: Work with mediated systems (e.g., Smoking → Tar → Cancer), check front-door conditions, compute front-door formula, numerical verification

Learning outcome: Recognize when mediation enables identification with unobserved confounding

Next Session: Formal Rules of Do-Calculus

In our next session, we'll delve into the formal rules of do-calculus, a complete algebraic system that provides a systematic way to determine when causal effects are identifiable from observational data. This powerful framework allows us to reason about interventions and predict their outcomes.

The Three Rules of Do-Calculus

Explore the fundamental rules that govern causal inference, enabling us to manipulate and simplify causal expressions.

Generalizing Back-Door and Front-Door

Understand how Do-Calculus unifies and extends the back-door and front-door adjustment criteria we've previously covered.

Practical Applications

Discover real-world scenarios where Do-Calculus is applied to identify and quantify causal effects in complex systems.

Let's take our unit 4 quiz.