Module 11 — Causality

Statistics

hbar.university

Potential Outcomes Framework

Position in Knowledge Graph

System: Statistics. Structural Domain: Causality (Module 11). Node 11.1.

The potential outcomes framework (Rubin causal model) is the foundational language for defining causal quantities in statistics. It formalizes what it means for a treatment to cause an effect by indexing outcomes to hypothetical interventions, making explicit the gap between what is observed and what is needed for causal claims. This node defines the estimands (ATE, ATT, CATE), the key assumptions (SUTVA, ignorability, overlap), and the fundamental problem of causal inference. It enables the graphical approach (Node 11.2), identification theory (Node 11.3), and experimental design (Node 11.4).

Formal Definition

For each unit $i \in \{1, \ldots, n\}$ and treatment value $t \in \mathcal{T}$, the **potential outcome** $Y_i(t)$ is the value of the outcome that unit i would exhibit if assigned to treatment t.

For binary treatment $T_i \in \{0, 1\}$:

- $Y_i(1)$: potential outcome under treatment.
- $Y_i(0)$: potential outcome under control.

The **individual treatment effect** (ITE) is

$$\tau_i = Y_i(1) - Y_i(0).$$

The **observed outcome** is

$$Y_i^{\text{obs}} = T_i\, Y_i(1) + (1 - T_i)\, Y_i(0) = Y_i(T_i).$$

Intuition

Causation requires comparing two states of the world: what happened under treatment versus what would have happened under control. The potential outcomes framework takes this literally—it defines both outcomes for every unit, even though only one is ever observed. The unobserved outcome is the **counterfactual**. The fundamental problem of causal inference is that we can never observe both potential outcomes for the same unit, so individual causal effects are never directly identified from data. All causal inference proceeds by making assumptions that link the observable data to the unobservable counterfactual distribution.

Structural Role

Dependencies: Builds on probability spaces (Node 1.3) and estimation theory (Node 3.1). **Enables:** DAGs (Node 11.2) provide a graphical language for the same causal relationships; identifiability (Node 11.3) gives conditions under which causal effects can be recovered; randomized experiments (Node 11.4) achieve identification by design.

The potential outcomes framework and the DAG/structural equation framework (Node 11.2) are complementary formalizations of causality. Potential outcomes define the estimands; DAGs encode the causal structure. The backdoor criterion (Node 11.3) translates between the two.

Mathematical Development

The Fundamental Problem of Causal Inference

For each unit i, only one of $Y_i(0)$ and $Y_i(1)$ is observed. The counterfactual $Y_i(1 - T_i)$ is missing. Therefore:

$$\tau_i = Y_i(1) - Y_i(0) \text{ is never identified from data for any individual } i.$$

Causal inference redirects attention from individual effects to **average effects**, which can be identified under assumptions.

Causal Estimands

Average Treatment Effect (ATE):

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

Average Treatment Effect on the Treated (ATT):

$$\tau_{\text{ATT}} = \mathbb{E}[Y(1) - Y(0) \mid T = 1].$$

Conditional Average Treatment Effect (CATE):

$$\tau(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x].$$

The ATE is the population-level average; the ATT restricts to the treated subpopulation; the CATE allows treatment effect heterogeneity across covariate values.

SUTVA: Stable Unit Treatment Value Assumption

SUTVA requires:

1. **No interference:** $Y_i(t_1, \ldots, t_n) = Y_i(t_i)$. Unit i's potential outcome depends only on its own treatment, not on the treatments of others.
2. **No hidden variations of treatment:** There is only one version of each treatment level. If $T_i = t$, then $Y_i^{\text{obs}} = Y_i(t)$ regardless of how treatment t was administered.

SUTVA is required for potential outcomes to be well-defined as individual-level quantities. It fails in settings with network effects (vaccination, social programs), where one unit's treatment affects another's outcome.

Ignorability (Unconfoundedness)

Strong ignorability given covariates X:

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

This states that, conditional on observed covariates, treatment assignment is independent of potential outcomes. It is the assumption that there are no unmeasured confounders.

Proof that ignorability identifies the ATE. Under ignorability:

$$\mathbb{E}[Y(t) \mid X] = \mathbb{E}[Y(t) \mid T = t, X] = \mathbb{E}[Y^{\text{obs}} \mid T = t, X],$$

where the first equality uses independence and the second uses consistency ($Y^{\text{obs}} = Y(T)$). Therefore:

$$\tau(x) = \mathbb{E}[Y \mid T = 1, X = x] - \mathbb{E}[Y \mid T = 0, X = x],$$

and integrating over X:

$$\tau_{\text{ATE}} = \mathbb{E}_X[\tau(X)] = \mathbb{E}_X\big[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\big].$$

Overlap (Positivity)

The **overlap** condition requires

$$0 < e(x) = \mathbb{P}(T = 1 \mid X = x) < 1 \quad \text{for all } x \text{ in the support of } X.$$

Without overlap, some covariate values have no treated (or no control) units, and $\mathbb{E}[Y \mid T = t, X = x]$ is undefined. Overlap ensures every subpopulation contains both treated and control units.

Selection Bias Decomposition

The naive estimator $\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]$ generally does not equal the ATE:

$$\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] = \tau_{\text{ATE}} + \underbrace{\mathbb{E}[Y(0) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0]}_{\text{Selection bias}}.$$

Selection bias is the difference in baseline (control) outcomes between treated and control groups. It vanishes under ignorability (or randomization) but is generically nonzero in observational studies.

Derivation. Write $\mathbb{E}[Y \mid T = 1] = \mathbb{E}[Y(1) \mid T = 1]$ and $\mathbb{E}[Y \mid T = 0] = \mathbb{E}[Y(0) \mid T = 0]$. Add and subtract $\mathbb{E}[Y(0) \mid T = 1]$:

$$\mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] = \underbrace{\mathbb{E}[Y(1) - Y(0) \mid T = 1]}_{\tau_{\text{ATT}}} + \underbrace{\mathbb{E}[Y(0) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0]}_{\text{Selection bias}}.$$

If additionally $\tau_{\text{ATT}} = \tau_{\text{ATE}}$ (homogeneous effects), the decomposition simplifies to the formula above.

Propensity Score

The **propensity score** $e(x) = \mathbb{P}(T = 1 \mid X = x)$ is a balancing score.

Theorem (Rosenbaum–Rubin, 1983). Under ignorability given X,

$$\{Y(0), Y(1)\} \perp\!\!\!\perp T \mid e(X).$$

Proof. It suffices to show $\mathbb{P}(T = 1 \mid Y(0), Y(1), e(X)) = e(X)$. By the law of iterated expectations:

$$\mathbb{P}(T = 1 \mid Y(0), Y(1), e(X)) = \mathbb{E}[\mathbb{P}(T = 1 \mid Y(0), Y(1), X) \mid Y(0), Y(1), e(X)].$$

By ignorability, $\mathbb{P}(T = 1 \mid Y(0), Y(1), X) = \mathbb{P}(T = 1 \mid X) = e(X)$. Since $e(X)$ is measurable with respect to $\sigma(e(X))$, the outer expectation gives $e(X)$.

This reduces the problem from conditioning on the high-dimensional X to conditioning on the scalar $e(X)$, enabling practical estimation (Node 11.5).

Worked Examples

Example 1: Selection Bias in a Job Training Program

A job training program is offered to unemployed workers. Outcome: earnings Y one year later. Treatment: $T = 1$ if enrolled. Observed: $\bar{Y}_1 = \$22{,}000$, $\bar{Y}_0 = \$18{,}000$.

Naive estimate: $\hat{\tau}^{\text{naive}} = \$4{,}000$.

But workers who enrolled may be more motivated (higher $Y(0)$). Suppose the true selection bias is $\mathbb{E}[Y(0) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] = \$2{,}500$. Then the true ATT is $\$4{,}000 - \$2{,}500 = \$1{,}500$.

Under ignorability given education and prior earnings X, the ATE is identified by the adjustment formula. If X does not capture motivation, ignorability fails and the naive estimate is biased.

Example 2: Computing ATE Under Ignorability

Binary treatment, binary covariate $X \in \{0, 1\}$ with $\mathbb{P}(X = 1) = 0.6$.

X	$\mathbb{E}[Y \mid T = 1, X]$	$\mathbb{E}[Y \mid T = 0, X]$	$\tau(X)$
0	5.0	3.0	2.0
1	8.0	5.5	2.5

$\tau_{\text{ATE}} = 0.4 \times 2.0 + 0.6 \times 2.5 = 0.8 + 1.5 = 2.3$.

Naive difference in means (without conditioning): suppose $\mathbb{P}(T = 1 \mid X = 1) = 0.7$, $\mathbb{P}(T = 1 \mid X = 0) = 0.3$. Then treated units are disproportionately $X = 1$.

$\mathbb{E}[Y \mid T = 1] = \frac{0.3 \times 0.4 \times 5.0 + 0.7 \times 0.6 \times 8.0}{0.3 \times 0.4 + 0.7 \times 0.6} = \frac{0.6 + 3.36}{0.12 + 0.42} = \frac{3.96}{0.54} = 7.33$.

$\mathbb{E}[Y \mid T = 0] = \frac{0.7 \times 0.4 \times 3.0 + 0.3 \times 0.6 \times 5.5}{0.7 \times 0.4 + 0.3 \times 0.6} = \frac{0.84 + 0.99}{0.28 + 0.18} = \frac{1.83}{0.46} = 3.98$.

Naive estimate: $7.33 - 3.98 = 3.35$, which overstates the true ATE of 2.3 because treated units have higher X values (confounding).

Cross-Links

AI. Counterfactual reasoning in reinforcement learning is the sequential decision analogue: the agent's potential outcomes under different action sequences are the value functions, and the fundamental problem of causal inference becomes the exploration-exploitation tradeoff.

Economics. The Roy model is a structural potential outcomes model with self-selection: agents choose treatment based on their comparative advantage, generating endogenous selection. The potential outcomes framework underpins the credibility revolution in empirical economics (Angrist, Imbens, Rubin).

Linear Algebra. The potential outcomes for n units form the matrix $\mathbf{Y} = [Y_i(t)]_{n \times |\mathcal{T}|}$, of which only one entry per row is observed. The fundamental problem is a missing data problem with a structured missingness pattern determined by the treatment assignment vector.

Quantum Mechanics. In quantum mechanics, counterfactual reasoning arises in the analysis of Bell inequalities: the potential outcomes framework applied to measurement settings formalizes the assumption of local realism, and Bell's theorem shows that no assignment of potential outcomes can reproduce quantum correlations.

Structural Compression

Invariant

Causal estimands are defined on the joint distribution of potential outcomes $(Y(0), Y(1))$; identification requires assumptions linking this unobserved joint distribution to the observed data distribution $p(Y^{\text{obs}}, T, X)$.

Mechanism

Potential outcomes index all counterfactual worlds; observed data indexes only one world per unit. Ignorability $\{Y(0), Y(1)\} \perp\!\!\!\perp T \mid X$ equates the observational conditional with the interventional distribution. Overlap ensures every covariate stratum has both treated and control units. The propensity score reduces the conditioning dimension from X to the scalar $e(X)$.

Cross-System Echo

The potential outcomes framework is the frequentist formalization of causality; in the structural equation modeling tradition (DAGs, Node 11.2), the same causal quantities are defined via interventions $do(T = t)$. In economics, the Roy model and the Heckman selection model are structural potential outcomes models with endogenous treatment choice. In AI, off-policy evaluation is the sequential analogue of the fundamental problem of causal inference.

Exercises

1. Show that $\tau_{\text{ATE}} = \tau_{\text{ATT}}$ if and only if $\mathbb{E}[Y(1) - Y(0) \mid T = 1] = \mathbb{E}[Y(1) - Y(0) \mid T = 0]$, i.e., the treatment effect is the same for treated and untreated subpopulations.
2. Prove the selection bias decomposition: $\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] = \tau_{\text{ATT}} + (\mathbb{E}[Y(0) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0])$.
3. Under ignorability and overlap, derive the inverse probability weighting (IPW) representation: $\tau_{\text{ATE}} = \mathbb{E}\left[\frac{TY}{e(X)} - \frac{(1-T)Y}{1-e(X)}\right]$.
4. Show that the propensity score is a balancing score: $\mathbb{E}[T \mid X, e(X)] = e(X)$, and use this to prove $\mathbb{E}[h(X) \mid T = 1, e(X)] = \mathbb{E}[h(X) \mid T = 0, e(X)]$ for any function h under ignorability.
5. Construct a numerical example with $n = 4$ units where the observed difference in means equals the ATE but individual treatment effects are heterogeneous, illustrating the ecological fallacy in reverse.

6. Consider a setting where SUTVA is violated (e.g., vaccination with herd immunity). Define the potential outcomes $Y_i(\mathbf{t})$ as a function of the full treatment vector $\mathbf{t} = (t_1, \ldots, t_n)$. How many potential outcomes does each unit have? Why does this make identification dramatically harder?

7. Argue, structurally, why the fundamental problem of causal inference—the impossibility of observing both potential outcomes for the same unit—is not a limitation of experimental design but a definitional feature of individual-level causation, and identify the assumptions that convert this individual-level impossibility into a population-level identification result.

Directed Acyclic Graphs

Position in Knowledge Graph

System: Statistics. Structural Domain: Causality (Module 11). Node 11.2.

Directed acyclic graphs (DAGs) provide the graphical language for encoding causal assumptions, deriving conditional independence relations, and defining interventions. Where the potential outcomes framework (Node 11.1) specifies the estimands, DAGs encode the structural relationships that make identification possible. The d-separation algorithm reads conditional independences from the graph; the do-operator formalizes interventions via graph surgery. This node enables the identifiability and backdoor criterion (Node 11.3).

Formal Definition

A **directed acyclic graph** $\mathcal{G} = (V, E)$ consists of: - A vertex set $V = \{X_1, \ldots, X_p\}$ representing random variables. - A set E of directed edges $X_i \to X_j$, with no directed cycles.

Parents: $\text{pa}(X_j) = \{X_i : X_i \to X_j \in E\}$. **Children:** $\text{ch}(X_j) = \{X_k : X_j \to X_k \in E\}$. **Descendants:** $\text{desc}(X_j)$ = all variables reachable from X_j via directed paths. **Ancestors:** $\text{an}(X_j)$ = all variables from which X_j is reachable via directed paths. **Non-descendants:** $\text{nd}(X_j) = V \setminus (\{X_j\} \cup \text{desc}(X_j))$.

Intuition

A DAG is a map of causal relationships: an arrow $X \to Y$ means X is a direct cause of Y. The absence of an arrow means no direct causal relationship. The graph structure determines which statistical associations are causal, which are confounded, and which are spurious. The power of the DAG framework is that complex causal reasoning reduces to graphical algorithms (d-separation), and interventions reduce to graph surgery (removing arrows).

Structural Role

Dependencies: Builds on conditional independence (Node 1.7) and joint distributions (Node 2.1). **Enables:** Identifiability and the backdoor criterion (Node 11.3), which uses d-separation to determine when causal effects can be estimated from observational data.

DAGs separate the statistical model (the factorization of p) from the causal model (the structural equations and interventions). Statistical dependence is readable from d-separation; causal effects require the interventional distribution, which depends on the graph structure.

Mathematical Development

Markov Factorization

A distribution P is **Markov** with respect to $\mathcal{G}$ if

$$p(x_1, \ldots, x_p) = \prod_{j=1}^{p} p(x_j \mid \text{pa}(x_j)).$$

Local Markov Property. Each variable is conditionally independent of its non-descendants given its parents:

$$X_j \perp\!\!\!\perp \mathrm{nd}(X_j) \mid \mathrm{pa}(X_j).$$

The local Markov property is equivalent to the factorization for positive distributions.

d-Separation: The Algorithm

For disjoint sets $A, B, C \subseteq V$, determine whether A and B are d-separated by C:

1. Enumerate all paths between any node in A and any node in B (paths may traverse edges in either direction).
2. A path is **blocked** by C if it contains at least one of:
   - A **chain** $X_i \to X_k \to X_j$ with $X_k \in C$.
   - A **fork** $X_i \leftarrow X_k \to X_j$ with $X_k \in C$.
   - A **collider** $X_i \to X_k \leftarrow X_j$ with $X_k \notin C$ and no descendant of X_k in C.
3. A and B are d-separated by C if **every** path between A and B is blocked.

Write $A \perp_{\mathcal{G}} B \mid C$ for d-separation.

Key subtlety: Conditioning on a collider (or its descendant) **opens** the path, creating a spurious association. This is the graphical basis of "collider bias" or "Berkson's paradox."

Global Markov Property

Theorem. If P is Markov with respect to $\mathcal{G}$, then

$$A \perp_{\mathcal{G}} B \mid C \implies A \perp\!\!\!\perp B \mid C \quad \text{under } P.$$

d-separation in the graph implies conditional independence in the distribution.

Faithfulness

A distribution P is **faithful** to $\mathcal{G}$ if the converse holds:

$$A \perp\!\!\!\perp B \mid C \quad \text{under } P \implies A \perp_{\mathcal{G}} B \mid C.$$

Faithfulness means the distribution has no "accidental" conditional independences beyond those implied by the graph. It fails on a Lebesgue-measure-zero set of parameters, so it is generically true but can fail in special cases (e.g., exact parameter cancellations).

Under both the Markov property and faithfulness, the conditional independence structure of P is exactly the set of d-separations in $\mathcal{G}$.

Causal DAGs and Structural Equations

A **causal DAG** augments the graphical structure with **structural equations** (also called structural causal models, SCMs):

$$X_j = f_j(\mathrm{pa}(X_j), U_j), \quad j = 1, \ldots, p,$$

where $U_1, \ldots, U_p$ are mutually independent exogenous noise variables. The structural equations define a **data-generating mechanism**: they specify not just the conditional distribution of X_j given its parents, but the functional relationship that would persist under interventions.

The joint distribution implied by the SCM is Markov with respect to $\mathcal{G}$ (the exogenous noise independence implies the factorization).

The do-Operator

The **intervention** $do(X_j = t)$ is defined by:

1. Replace the structural equation for X_j with $X_j = t$ (deterministic).
2. Leave all other structural equations unchanged.

Graphically: remove all edges into X_j (the **mutilated graph** $\mathcal{G}_{\overline{X_j}}$) and set $X_j = t$.

The **interventional distribution** is

$$p(x_1, \ldots, x_p \mid do(X_j = t)) = \prod_{k \neq j} p(x_k \mid \mathrm{pa}(x_k))\Big|_{x_j = t}.$$

This generally differs from the observational conditional $p(x_1, \ldots, x_p \mid X_j = t)$, because conditioning does not remove the causal influence of X_j's parents.

Observational vs. Interventional Distributions

Example. Consider $U \to T \to Y$ and $U \to Y$ (confounding).

Observational: $p(Y \mid T = t) = \int p(Y \mid T = t, U = u)\, p(U \mid T = t)\, du$. The confounder U has a different distribution for different values of T.

Interventional: $p(Y \mid do(T = t)) = \int p(Y \mid T = t, U = u)\, p(U)\, du$. Under intervention, U retains its marginal distribution because the arrow $U \to T$ is severed.

The difference between these two expressions is the essence of confounding.

Markov Equivalence

Two DAGs are **Markov equivalent** if they imply the same set of conditional independences (same d-separations). Markov equivalent DAGs have: - The same skeleton (undirected edges). - The same **v-structures** (immoralities): patterns $X_i \to X_k \leftarrow X_j$ where X_i and X_j are not adjacent.

The set of all DAGs Markov equivalent to $\mathcal{G}$ is the **Markov equivalence class**, represented by a **completed partially directed acyclic graph** (CPDAG). Observational data alone can identify the Markov equivalence class but not the DAG itself (without additional assumptions or interventional data).

Worked Examples

Example 1: d-Separation in a Confounding Structure

DAG: $X \leftarrow U \to Y$, $X \to Y$. Variables: confounder U, treatment X, outcome Y.

Is $X \perp_{\mathcal{G}} Y \mid \emptyset$**?** Path 1: $X \to Y$ (direct, no collider, not blocked). Answer: No.

Is $X \perp_{\mathcal{G}} Y \mid U$**?** Path 1: $X \to Y$ (chain with no middle node in $C = \{U\}$, so not blocked). Answer: No. Conditioning on U blocks the backdoor path $X \leftarrow U \to Y$ but does not block the direct path $X \to Y$.

Example 2: Collider Bias

DAG: $X \to C \leftarrow Y$. No direct path between X and Y.

Is $X \perp_{\mathcal{G}} Y \mid \emptyset$**?** The only path is $X \to C \leftarrow Y$, which contains a collider at C. Since $C \notin \emptyset$, the path is blocked. Answer: Yes, $X \perp_{\mathcal{G}} Y$.

Is $X \perp_{\mathcal{G}} Y \mid C$**?** Now $C \in \{C\}$, so the collider is conditioned on, which **opens** the path. Answer: No, $X \not\perp_{\mathcal{G}} Y \mid C$.

This is Berkson's paradox: conditioning on a common effect creates a spurious association between its causes. For instance, if X = talent and Y = luck, and C = success (which requires one or both), then among successful people, talent and luck appear negatively correlated.

Cross-Links

AI. DAGs are the foundation of Bayesian networks. The d-separation algorithm underlies message-passing (belief propagation) for efficient inference in graphical models. Causal discovery algorithms (PC, GES, FCI) learn DAG structure from data using conditional independence tests.

Economics. Structural equation models in econometrics are causal DAGs with specific functional forms (typically linear). The distinction between structural and reduced-form equations corresponds to the distinction between causal and observational distributions.

Linear Algebra. The adjacency matrix A of a DAG (with $A_{ij} = 1$ if $X_i \to X_j$) is strictly upper-triangular (after topological sorting), hence nilpotent. The reachability matrix $(I - A)^{-1} = \sum_{k=0}^{p-1} A^k$ encodes all ancestor-descendant relationships.

Quantum Mechanics. Quantum causal models extend DAGs to quantum channels, where the edges represent quantum operations rather than classical conditional distributions. The no-signaling principle in quantum mechanics is the analogue of d-separation: spacelike-separated parties are d-separated by their common past.

Structural Compression

Invariant

d-separation in $\mathcal{G}$ implies conditional independence in any distribution Markov with respect to $\mathcal{G}$; under faithfulness, the implication is an equivalence. The interventional distribution is the observational distribution of the mutilated graph.

Mechanism

The Markov factorization decomposes the joint distribution into local conditional distributions, one per variable. Interventions sever parent-child links by replacing a structural equation with a constant, altering the factorization and producing the interventional distribution. d-separation is an algorithmic procedure for reading off conditional independences from the factorization structure.

Cross-System Echo

DAGs are the graphical language of Bayesian networks in AI; d-separation underlies message-passing algorithms (belief propagation) for inference in graphical models. In econometrics, the DAG framework formalizes the concept of exogeneity and the conditions under which regression estimates have causal interpretations. In quantum mechanics, causal structure is formalized by quantum causal models extending DAGs to quantum channels.

Exercises

1. For the DAG $X_1 \to X_3 \leftarrow X_2$, $X_3 \to X_4$, determine all conditional independence relations implied by d-separation (i.e., enumerate all valid $A \perp_{\mathcal{G}} B \mid C$ statements).
2. Show that the Markov factorization $p(x_1, \ldots, x_p) = \prod_j p(x_j \mid \mathrm{pa}(x_j))$ implies the local Markov property $X_j \perp\!\!\!\perp \mathrm{nd}(X_j) \mid \mathrm{pa}(X_j)$.

3. Construct two DAGs on three variables that are Markov equivalent (same skeleton and v-structures) and one that is not. Identify the distinguishing conditional independence.

4. For the DAG $Z \to T \to Y$, $Z \to Y$, compute the interventional distribution $p(Y \mid do(T = t))$ using the truncated factorization and show it differs from $p(Y \mid T = t)$.

5. Prove that conditioning on a collider opens the path: for $X \to C \leftarrow Y$ with $X \perp\!\!\!\perp Y$ marginally, show that $X \not\perp\!\!\!\perp Y \mid C$ when the structural equations are linear with Gaussian noise.

6. Show that the adjacency matrix of a DAG is nilpotent (i.e., $A^p = 0$), and use this to prove that the factorization $p(x) = \prod_j p(x_j \mid \mathrm{pa}(x_j))$ defines a valid joint distribution (no circular dependencies).

7. Argue, structurally, why observational data alone can identify the Markov equivalence class but not the specific DAG, and explain what additional information (interventions, temporal ordering, or functional form assumptions) is needed to distinguish between Markov-equivalent DAGs.

Identifiability and the Backdoor Criterion

Position in Knowledge Graph

System: Statistics. Structural Domain: Causality (Module 11). Node 11.3.

Identifiability is the question that precedes all estimation: can the causal effect be expressed as a functional of the observed data distribution? This node develops the backdoor criterion (the most commonly applicable graphical condition), the front-door criterion (for unmeasured confounding settings), and the do-calculus (the complete set of rules for converting interventional distributions to observational ones). It connects the potential outcomes framework (Node 11.1) to the DAG framework (Node 11.2) and enables both experimental (Node 11.4) and observational (Node 11.5) estimation.

Formal Definition

Let $\mathcal{G}$ be a causal DAG over (T, Y, X), where T is treatment, Y is outcome, and X is a vector of observed covariates. The target is

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

A causal quantity $Q = Q(p(\cdot \mid do(\cdot)))$ is **identified** from the observational distribution $p(Y, T, X)$ if it can be uniquely computed from $p(Y, T, X)$ for all structural causal models compatible with $\mathcal{G}$.

Non-identifiability arises when two SCMs generate the same observational distribution but imply different values of Q.

Intuition

The causal effect of T on Y is the change in Y when we physically set T to different values (an intervention). But in observational data, T was not set by the researcher—it was determined by a potentially confounded process. Identifiability asks whether the confounding can be removed by conditioning on observed variables. The backdoor criterion gives a graphical test: if we can block all "backdoor paths" (non-causal paths that create spurious association) by conditioning on observed covariates, the causal effect equals the covariate-adjusted observational contrast.

Structural Role

Dependencies: Requires the potential outcomes framework (Node 11.1) for the estimands and DAGs (Node 11.2) for the graphical language. **Enables:** Experimental design (Node 11.4) as the case where identifiability holds by design; observational estimation methods (Node 11.5) which require identification as a precondition.

Identifiability analysis determines whether causal effects can be estimated at all, prior to any estimation procedure. It is the logical prerequisite for all of Nodes 11.4–11.6.

Mathematical Development

Backdoor Paths

A **backdoor path** from T to Y is any path that begins with an arrow into T (i.e., the first edge on the path is $\cdot \to T$). Backdoor paths represent non-causal associations between T and Y—they transmit confounding.

The Backdoor Criterion

Definition. A set of variables X satisfies the **backdoor criterion** relative to (T, Y) in $\mathcal{G}$ if:

1. No variable in X is a descendant of T.
2. X d-separates T from Y in the graph $\mathcal{G}_{\underline{T}}$ (the graph with all arrows out of T removed), or equivalently, X blocks every backdoor path from T to Y.

Theorem: Backdoor Adjustment Formula

Theorem (Pearl, 1993). If X satisfies the backdoor criterion relative to (T, Y), then

$$p(Y = y \mid do(T = t)) = \int p(Y = y \mid T = t, X = x)\, p(X = x)\, dx.$$

Proof sketch. In the mutilated graph $\mathcal{G}_{\overline{T}}$ (arrows into T removed, representing $do(T = t)$):

$$p(y \mid do(t)) = \int p(y \mid t, x, \mathcal{G}_{\overline{T}})\, p(x \mid \mathcal{G}_{\overline{T}})\, dx.$$

By condition 1, X is not a descendant of T, so removing arrows into T does not affect the distribution of X: $p(x \mid \mathcal{G}_{\overline{T}}) = p(x)$.

By condition 2, X blocks all backdoor paths, so in $\mathcal{G}_{\overline{T}}$, $Y \perp\!\!\!\perp$ (removed parents of T) $\mid T, X$. This means the conditional $p(y \mid t, x)$ is the same in the original and mutilated graphs: $p(y \mid t, x, \mathcal{G}_{\overline{T}}) = p(y \mid t, x)$.

Combining: $p(y \mid do(t)) = \int p(y \mid t, x)\, p(x)\, dx$.

The **ATE** is therefore:

$$\tau = \mathbb{E}_X \big[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\big].$$

This is the **g-formula** or **adjustment formula**.

Connection to Ignorability

The backdoor criterion in the DAG framework is equivalent to ignorability in the potential outcomes framework:

$$X \text{ satisfies backdoor} \iff \{Y(0), Y(1)\} \perp\!\!\!\perp T \mid X.$$

The DAG makes the structural reason explicit: X blocks all non-causal paths. The potential outcomes framework states the consequence: treatment is as-good-as-random conditional on X.

The Front-Door Criterion

When no measured set satisfies the backdoor criterion (unmeasured confounding between T and Y), identification may still be possible via a **mediator** M.

Definition. A set M satisfies the **front-door criterion** relative to (T, Y) if:

1. T blocks all backdoor paths from M to Y: there are no unblocked backdoor paths from M to Y that do not go through T.
2. M intercepts all directed paths from T to Y: every causal path from T to Y goes through M.
3. There are no backdoor paths from T to M: the effect of T on M is identified without adjustment.

Front-Door Adjustment Formula:

$$p(Y = y \mid do(T = t)) = \sum_{m} p(M = m \mid T = t) \sum_{t'} p(Y = y \mid M = m, T = t')\, p(T = t').$$

Derivation. Apply the do-calculus in two steps:

Step 1: $p(y \mid do(t)) = \sum_m p(y \mid do(t), do(m))\, p(m \mid do(t))$. Since all $T \to Y$ paths go through M, $p(y \mid do(t), do(m)) = p(y \mid do(m))$.

Step 2: $p(m \mid do(t)) = p(m \mid t)$ (no backdoor paths from T to M).

Step 3: $p(y \mid do(m)) = \sum_{t'} p(y \mid m, t')\, p(t')$ (applying the backdoor adjustment with T blocking backdoor paths from M to Y).

Combining: $p(y \mid do(t)) = \sum_m p(m \mid t) \sum_{t'} p(y \mid m, t')\, p(t')$.

Pearl's do-Calculus

The **do-calculus** consists of three rules for manipulating expressions involving $do(\,\cdot\,)$. Let X, Y, Z, W be disjoint sets of variables in a causal DAG $\mathcal{G}$.

Rule 1 (Insertion/deletion of observations):

$$p(y \mid do(x), z, w) = p(y \mid do(x), w) \quad \text{if } Y \perp_{\mathcal{G}_{\overline{X}}} Z \mid W.$$

Rule 2 (Action/observation exchange):

$$p(y \mid do(x), do(z), w) = p(y \mid do(x), z, w) \quad \text{if } Y \perp_{\mathcal{G}_{\overline{X}\underline{Z}}} Z \mid W.$$

Rule 3 (Insertion/deletion of actions):

$$p(y \mid do(x), do(z), w) = p(y \mid do(x), w) \quad \text{if } Y \perp_{\mathcal{G}_{\overline{X}\,\overline{Z(S)}}} Z \mid W,$$

where $Z(S)$ is the set of nodes in Z that are not ancestors of any node in W in $\mathcal{G}_{\overline{X}}$.

Completeness Theorem (Huang–Valtorta, Shpitser–Pearl, 2006). An interventional distribution $p(y \mid do(x))$ is identifiable from observational data if and only if it can be reduced to an observational expression using the three rules of the do-calculus.

Worked Examples

Example 1: Backdoor Adjustment

DAG: $U \to T$, $U \to Y$, $T \to Y$, $X \to U$, $X \to T$. Observed: T, Y, X. Unobserved: U.

Does X satisfy the backdoor criterion for (T, Y)?

Backdoor paths from T to Y: $T \leftarrow U \to Y$ and $T \leftarrow X \to U \to Y$. The path $T \leftarrow U \to Y$: is it blocked by X? The fork $T \leftarrow U \to Y$ requires $U \in X$ to block, but U is unobserved. The path $T \leftarrow X \to U \to Y$: the fork at X is blocked by conditioning on X.

So X blocks one backdoor path but not $T \leftarrow U \to Y$. The backdoor criterion is **not satisfied** by X alone. We cannot identify the causal effect of T on Y by adjusting for X only; the unmeasured confounder U creates a non-blockable backdoor path.

Example 2: Front-Door Criterion Application

DAG: $U \to T$, $U \to Y$, $T \to M \to Y$. All of T, M, Y observed; U unobserved. All directed paths from T to Y go through M. No backdoor paths from T to M (the only candidate $T \leftarrow U \to \cdots \to M$ does not exist since $U \not\to M$). Backdoor paths from M to Y: $M \leftarrow T \leftarrow U \to Y$, blocked by T.

The front-door criterion is satisfied by M. The causal effect is:

$$\mathbb{E}[Y \mid do(T = t)] = \sum_m P(M = m \mid T = t) \sum_{t'} \mathbb{E}[Y \mid M = m, T = t'] \, P(T = t').$$

Suppose $T, M \in \{0, 1\}$, $P(M = 1 \mid T = 1) = 0.8$, $P(M = 1 \mid T = 0) = 0.2$, $P(T = 1) = 0.5$, and the conditional means of Y are: $\mathbb{E}[Y \mid M = 1, T = 1] = 10$, $\mathbb{E}[Y \mid M = 1, T = 0] = 8$, $\mathbb{E}[Y \mid M = 0, T = 1] = 4$, $\mathbb{E}[Y \mid M = 0, T = 0] = 3$.

$\mathbb{E}[Y \mid do(T = 1)] = 0.8[0.5 \cdot 10 + 0.5 \cdot 8] + 0.2[0.5 \cdot 4 + 0.5 \cdot 3] = 0.8 \cdot 9 + 0.2 \cdot 3.5 = 7.2 + 0.7 = 7.9.$

$\mathbb{E}[Y \mid do(T = 0)] = 0.2 \cdot 9 + 0.8 \cdot 3.5 = 1.8 + 2.8 = 4.6.$

Causal effect: $\tau = 7.9 - 4.6 = 3.3$.

Cross-Links

AI. Causal inference in reinforcement learning requires identifying the effect of policies (interventions on action sequences) from offline data. The do-calculus generalizes to sequential settings where the "treatment" is a sequence of actions, and identification requires blocking backdoor paths at each time step.

Economics. The g-formula (adjustment formula) is the semiparametric identification result underlying the potential outcomes approach in economics. Instrumental variables (Node 11.6) handle cases where the backdoor criterion fails due to unmeasured confounders.

Linear Algebra. The adjustment formula $\mathbb{E}[Y \mid do(T = t)] = \sum_x p(y \mid t, x)p(x)$ is a marginalization operation: a matrix-vector product when X is discrete, with the transition matrix $p(y \mid t, x)$ and the marginal distribution $p(x)$ as the vector.

Quantum Mechanics. In quantum causal models, the do-operator corresponds to a quantum intervention (a completely positive trace-preserving map applied to a quantum system), and the analogue of the backdoor criterion determines when quantum causal effects are identifiable from observed quantum correlations.

Structural Compression

Invariant

A causal effect is identified if and only if all non-causal (backdoor) paths between treatment and outcome can be blocked by observed variables; the adjustment formula then expresses the causal effect as a weighted average of the observational conditional.

Mechanism

The backdoor criterion uses d-separation in the mutilated graph to ensure that $p(y \mid t, x) = p(y \mid do(t), x)$. Marginalizing over X with respect to $p(x)$ (not $p(x \mid t)$) removes the confounding. The front-door criterion chains two identification steps through a mediator. The do-calculus provides the complete algorithmic procedure for testing identifiability in arbitrary DAGs.

Cross-System Echo

The g-formula is the causal analogue of the law of total probability. In AI, offline policy evaluation requires identifying the effect of a target policy from data collected under a behavior policy—the sequential

generalization of backdoor adjustment via importance weighting. In economics, the Frisch–Waugh theorem for instrumental variables is the linear-model specialization of identification under confounding.

Exercises

1. For the DAG $Z \to T \to Y$, $Z \to Y$, verify that Z satisfies the backdoor criterion for (T, Y) and write out the adjustment formula.
2. Prove the backdoor adjustment formula by showing that in the mutilated graph $\mathcal{G}_{\overline{T}}$, $p(y \mid t, x, \mathcal{G}_{\overline{T}}) = p(y \mid t, x)$ and $p(x \mid \mathcal{G}_{\overline{T}}) = p(x)$ when X satisfies the backdoor criterion.
3. Construct a DAG where X contains a descendant of T, and show that conditioning on X biases the estimate of the causal effect (an instance of "bad control").
4. Apply the front-door criterion to the smoking–tar–cancer DAG (smoking $\to$ tar deposits $\to$ cancer, with an unmeasured common cause of smoking and cancer) and derive the front-door adjustment formula.
5. Using the three rules of the do-calculus, reduce $p(y \mid do(t))$ to an observational expression for the DAG $T \to M \to Y$, $U \to T$, $U \to Y$, verifying the front-door formula.
6. Show that for the DAG $T \to Y \leftarrow U \to T$ with U unobserved and no other variables, the causal effect $\mathbb{E}[Y \mid do(T = t)]$ is not identified. Explain why the do-calculus fails.
7. Argue, structurally, why completeness of the do-calculus means that identifiability is a purely graphical question—it depends on the DAG structure and which variables are observed, not on the functional forms or parameter values of the structural equations.

Randomized Experiments

Position in Knowledge Graph

System: Statistics. Structural Domain: Causality (Module 11). Node 11.4.

Randomized experiments are the gold standard for causal identification: they make ignorability hold by design, eliminating confounding without requiring knowledge of the causal structure. This node develops the theory of completely randomized experiments, including the Neyman framework for ATE estimation, Fisher's exact randomization test, and covariate adjustment via ANCOVA. It operationalizes the potential outcomes framework (Node 11.1) under the strongest possible identification conditions, and sets the benchmark against which observational methods (Node 11.5) are evaluated.

Formal Definition

In a **completely randomized experiment (CRE)**, n units are assigned to treatment ($T_i = 1$) or control ($T_i = 0$) by a known random mechanism, with n_1 treated and $n_0 = n - n_1$ control units. The assignment satisfies:

$$\{Y_i(0), Y_i(1)\}_{i=1}^n \perp\!\!\!\perp (T_1, \ldots, T_n),$$

and the assignment mechanism is fully specified: $\mathbb{P}(\mathbf{T} = \mathbf{t}) = \binom{n}{n_1}^{-1}$ for all $\mathbf{t}$ with $\sum t_i = n_1$.

Intuition

Randomization physically severs all arrows into the treatment variable in the causal DAG. No matter how complex the confounding structure, random assignment ensures that treated and control groups are comparable on average—both on observed and unobserved characteristics. The beauty of randomization is that identification follows from the design, not from modeling assumptions. The price is feasibility: many causal questions cannot be answered by experiments (ethics, cost, logistics).

Structural Role

Dependencies: Builds on the potential outcomes framework (Node 11.1) for the estimand definitions and hypothesis testing (Node 5.1) for the inferential framework. **Enables:** Observational methods (Node 11.5) by providing the ideal benchmark against which observational designs are compared.

Randomization removes all backdoor paths by design (Node 11.3). Fisher's test provides exact finite-sample inference; Neyman's framework provides asymptotically valid confidence intervals. ANCOVA leverages pre-treatment covariates for precision gain without risking bias.

Mathematical Development

Identification Under Randomization

Under CRE and SUTVA:

$$\mathbb{E}[Y \mid T = 1] = \mathbb{E}[Y(1) \mid T = 1] = \mathbb{E}[Y(1)],$$

where the last equality uses independence. Similarly $\mathbb{E}[Y \mid T = 0] = \mathbb{E}[Y(0)]$. Therefore:

$$\tau_{\text{ATE}} = \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0].$$

This requires no modeling, no covariate adjustment, and no distributional assumptions. It follows entirely from the randomization mechanism.

Neyman's Repeated Sampling Framework

The **difference-in-means** estimator is

$$\hat{\tau} = \bar{Y}_1 - \bar{Y}_0 = \frac{1}{n_1} \sum_{i:T_i=1} Y_i - \frac{1}{n_0} \sum_{i:T_i=0} Y_i.$$

Theorem (Neyman, 1923). Under CRE:

$$\mathbb{E}[\hat{\tau}] = \tau_{\text{ATE}}.$$

Proof. $\mathbb{E}[\bar{Y}_1] = \mathbb{E}\Big[\frac{1}{n_1} \sum_{i:T_i=1} Y_i(1)\Big]$. Under random assignment, each unit i has probability n_1/n of being treated, and $\mathbb{E}[\sum_{i:T_i=1} Y_i(1)] = \frac{n_1}{n} \sum_{i=1}^{n} Y_i(1)$, so $\mathbb{E}[\bar{Y}_1] = \frac{1}{n} \sum_i Y_i(1) = \bar{Y}(1)$. Similarly $\mathbb{E}[\bar{Y}_0] = \bar{Y}(0)$. Therefore $\mathbb{E}[\hat{\tau}] = \bar{Y}(1) - \bar{Y}(0) = \tau_{\text{ATE}}$.

Variance of $\hat{\tau}$:

$$\text{Var}(\hat{\tau}) = \frac{S_1^2}{n_1} + \frac{S_0^2}{n_0} - \frac{S_\tau^2}{n},$$

where $S_t^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i(t) - \bar{Y}(t))^2$ and $S_\tau^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\tau_i - \bar{\tau})^2$ is the variance of individual treatment effects.

The term S_τ^2/n involves both potential outcomes and is **not estimable** from data (the fundamental problem). The **conservative variance estimator** is

$$\hat{V} = \frac{s_1^2}{n_1} + \frac{s_0^2}{n_0}, \quad s_t^2 = \frac{1}{n_t - 1} \sum_{i:T_i=t} (Y_i - \bar{Y}_t)^2,$$

which overestimates $\text{Var}(\hat{\tau})$ by $S_\tau^2/n \geq 0$. Under constant treatment effects ($\tau_i = \tau$ for all i), $S_\tau^2 = 0$ and the estimator is exact.

Fisher's Randomization Test

Sharp null hypothesis:

$$H_0^F : Y_i(1) = Y_i(0) \quad \text{for all } i = 1, \ldots, n.$$

Under H_0^F, both potential outcomes equal Y_i^{obs}, so the complete data table is known. For any test statistic $T(\mathbf{T}, \mathbf{Y}^{\text{obs}})$, the exact null distribution is obtained by enumerating all $\binom{n}{n_1}$ possible treatment assignments:

$$p\text{-value} = \frac{|\{\mathbf{t} : T(\mathbf{t}, \mathbf{Y}^{\text{obs}}) \geq T_{\text{obs}}\}|}{\binom{n}{n_1}}.$$

Properties: - Exact for any sample size (no asymptotic approximation). - No distributional assumptions on Y. - Valid for any test statistic (difference in means, rank statistics, Kolmogorov–Smirnov, etc.). - Tests the sharp null (no effect for any unit), which is stronger than Neyman's null ($\tau_{\text{ATE}} = 0$).

Computational note. For large n, exact enumeration is infeasible. Monte Carlo approximation: sample B random permutations and compute the proportion exceeding T_{obs}. This gives a randomization p-value with precision $O(1/\sqrt{B})$.

Covariate Adjustment: ANCOVA

Even under randomization, pre-treatment covariates X can improve precision.

Lin's ANCOVA estimator. Fit the regression:

$$Y_i = \alpha + \tau T_i + \gamma^\top X_i + \delta^\top (T_i \cdot (X_i - \bar{X})) + \varepsilon_i,$$

where the interaction $T_i \cdot (X_i - \bar{X})$ allows different slopes in treated and control groups. The coefficient $\hat{\tau}^{\text{Lin}}$ on T_i is the ANCOVA estimator.

Theorem (Lin, 2013). Under randomization, $\hat{\tau}^{\text{Lin}}$ is:

1. Consistent for τ_{ATE} regardless of whether the linear model is correctly specified.
2. Asymptotically at least as efficient as the unadjusted $\hat{\tau}$.
3. Strictly more efficient when X is predictive of Y.

Why adjustment does not bias. Under randomization, $T \perp\!\!\!\perp X$, so including X as a covariate does not introduce confounding. The regression coefficient on T estimates the same causal effect with or without covariates; the improvement is purely in variance.

Precision gain. If R^2 is the fraction of outcome variance explained by X, the ANCOVA variance is approximately $(1 - R^2) \cdot \text{Var}(\hat{\tau})$. For $R^2 = 0.5$, the effective sample size doubles.

Stratified (Block) Randomization

In **stratified randomization**, units are grouped into strata based on X, and randomization occurs within each stratum. This guarantees exact covariate balance and further reduces variance.

The stratified estimator is

$$\hat{\tau}^{\text{strat}} = \sum_s \frac{n_s}{n} \hat{\tau}_s,$$

where $\hat{\tau}_s = \bar{Y}_{1,s} - \bar{Y}_{0,s}$ is the within-stratum difference in means and n_s is the stratum size. This is guaranteed to have $\text{Var}(\hat{\tau}^{\text{strat}}) \leq \text{Var}(\hat{\tau})$.

Worked Examples

Example 1: Neyman Inference for a Drug Trial

$n = 200$ patients randomized equally: $n_1 = n_0 = 100$. Outcome: blood pressure reduction (mmHg).

Observed: $\bar{Y}_1 = 12.3$, $\bar{Y}_0 = 8.7$, $s_1^2 = 25.0$, $s_0^2 = 20.0$.

$\hat{\tau} = 12.3 - 8.7 = 3.6$ mmHg. $\hat{V} = 25/100 + 20/100 = 0.45$. SE $= \sqrt{0.45} = 0.671$.

95% CI: $3.6 \pm 1.96 \times 0.671 = (2.28, 4.92)$. The CI is conservative (it overestimates the true variance by omitting S_τ^2/n).

Test $H_0 : \tau = 0$: $z = 3.6/0.671 = 5.37$, $p < 0.001$. Strong evidence of a treatment effect.

Example 2: Fisher Randomization Test

$n = 8$, $n_1 = 4$. Observed outcomes: treated = {7, 9, 6, 8}, control = {3, 5, 4, 2}.

$T_{\text{obs}} = \bar{Y}_1 - \bar{Y}_0 = 7.5 - 3.5 = 4.0$.

Under H_0^F, all 8 outcomes are fixed. There are $\binom{8}{4} = 70$ possible assignments.

For each permutation, compute $\bar{Y}_1 - \bar{Y}_0$. Count the number of permutations yielding ≥ 4.0.

The maximum possible difference (assigning {7, 8, 9, 6} vs. {2, 3, 4, 5}) is $7.5 - 3.5 = 4.0$, but other permutations achieving this: {9, 8, 7, 6} vs. {5, 4, 3, 2} gives 4.0. After enumeration, suppose 2 out of 70 permutations yield ≥ 4.0.

p-value $= 2/70 = 0.029$. Reject H_0^F at the 5% level.

Cross-Links

AI. A/B testing in product development is randomized experimentation with the ATE as the estimand. Multi-armed bandit algorithms extend the experimental framework to sequential settings where treatment assignment adapts based on interim results, trading off exploration and exploitation.

Economics. Randomized controlled trials (RCTs) are the basis for the credibility revolution in program evaluation (Angrist, Duflo, Banerjee). Field experiments in development economics randomize at the village or school level (cluster randomization), requiring cluster-robust variance estimators.

Linear Algebra. The design matrix of a CRE has a specific structure: T is a binary vector with exactly n_1 ones. The hat matrix for the unadjusted estimator is $H = \text{diag}(T/n_1 + (1-T)/n_0)$, and the ANCOVA hat matrix is the orthogonal projection onto the column space of $[T, X, T \cdot X]$.

Quantum Mechanics. Bell tests are randomized experiments in physics: measurement settings (treatments) are randomly assigned to entangled particles, and the outcome statistics test whether local hidden variable models (the causal DAG with unmeasured common cause) can explain the observed correlations. Randomization of measurement settings is essential to close the "freedom of choice" loophole.

Structural Compression

Invariant

Randomization renders treatment independent of all potential outcomes; the observed difference in means is an unbiased estimator of the ATE requiring no model for the outcome and no knowledge of the confounding structure.

Mechanism

Physical randomization removes all backdoor paths between T and Y in the causal DAG. The known assignment mechanism enables exact permutation inference (Fisher) and conservative variance estimation (Neyman). Covariate adjustment reduces variance without introducing bias because $T \perp\!\!\!\perp X$ by design.

Cross-System Echo

Fisher's randomization test is the prototype of permutation tests in nonparametric statistics. In AI, A/B testing is the standard experimental paradigm for causal evaluation of algorithmic changes. In economics, RCTs underpin the credibility revolution. In physics, Bell tests use randomized measurement settings to test fundamental causal structure of quantum mechanics.

Exercises

1. Prove that the difference-in-means estimator $\hat{\tau} = \bar{Y}_1 - \bar{Y}_0$ is unbiased for τ_{ATE} under completely randomized assignment. State explicitly where independence of treatment and potential outcomes is used.

2. Derive the Neyman variance formula $\text{Var}(\hat{\tau}) = S_1^2/n_1 + S_0^2/n_0 - S_\tau^2/n$. Show that omitting S_τ^2/n yields a conservative estimator.

3. For $n = 6$, $n_1 = 3$, outcomes $= \{10, 4, 7, 3, 8, 5\}$ (first 3 treated), compute the Fisher exact p-value for the difference-in-means test statistic under the sharp null. Enumerate all $\binom{6}{3} = 20$ permutations.

4. Show that under constant treatment effects ($\tau_i = \tau$ for all i), the Neyman variance formula simplifies to $\text{Var}(\hat{\tau}) = (S_0^2/n)(1/n_1 + 1/n_0)$, and the conservative estimator is exact.

5. Prove Lin's result: under randomization, the ANCOVA estimator with treatment-covariate interactions is consistent for τ_{ATE} even when the linear model is misspecified. (Hint: use $T \perp\!\!\!\perp X$ and the law of iterated expectations.)

6. Compare the variance of the unadjusted estimator, the ANCOVA estimator, and the stratified estimator for an experiment with $n = 100$, two strata of equal size, and $R^2 = 0.4$. Quantify the efficiency gain from each adjustment.

7. Argue, structurally, why randomization is the only design that achieves identification of the ATE without assumptions about the functional form of the outcome model or the set of confounders, and explain why observational methods (Node 11.5) can never fully replicate this property.

Observational Studies and Treatment Effect Estimation

Position in Knowledge Graph

System: Statistics. Structural Domain: Causality (Module 11). Node 11.5.

When randomization is infeasible, observational studies must rely on identification assumptions to estimate causal effects. This node develops three major classes of methods: (1) propensity score methods under ignorability (IPW, matching), (2) instrumental variables when ignorability fails, and (3) quasi-experimental designs (difference-in-differences, regression discontinuity) that exploit specific structural features of the data-generating process. Each method identifies a specific estimand for a specific subpopulation under distinct assumptions. The doubly robust estimator unifies propensity score and outcome modeling approaches.

Formal Definition

In observational studies, treatment T_i may depend on covariates X_i and unobserved factors. Under **strong ignorability** $\{Y(0), Y(1)\} \perp\!\!\!\perp T \mid X$ and **overlap** $0 < e(x) < 1$, where

$$e(x) = \mathbb{P}(T = 1 \mid X = x)$$

is the **propensity score**, the ATE is identified via the backdoor formula (Node 11.3).

When ignorability fails (unmeasured confounding), additional structure is needed: instruments, panel variation, or threshold rules.

Intuition

Observational studies face the core challenge that treated and control groups differ systematically. Propensity score methods attempt to recreate the conditions of a randomized experiment by reweighting or matching observations so that covariate distributions are balanced. Instrumental variables circumvent confounding by finding an external source of variation that affects treatment but not the outcome directly. Regression discontinuity exploits a known threshold rule to create local randomization. Each approach trades off generality (what estimand, what subpopulation) against the strength of required assumptions.

Structural Role

Dependencies: Requires identification theory (Node 11.3) as a logical prerequisite, randomized experiments (Node 11.4) as the benchmark, large-sample theory (Node 7.5), and regression (Node 8.6). **Enables:** Policy evaluation (Node 11.6), which extends these methods to decision rules and sequential settings.

All methods in this node require assumptions that are fundamentally untestable from observational data alone. The choice between methods depends on which assumptions are most credible in the specific application.

Mathematical Development

Propensity Score Methods

Theorem (Rosenbaum–Rubin, 1983). Under ignorability given X, $\{Y(0), Y(1)\} \perp\!\!\!\perp T \mid e(X)$.

This reduces conditioning from the potentially high-dimensional X to the scalar $e(X)$, enabling practical estimation.

Inverse Probability Weighting (IPW)

The **Horvitz–Thompson IPW estimator** of the ATE is

$$\hat{\tau}^{\mathrm{IPW}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{T_i Y_i}{\hat{e}(X_i)} - \frac{(1-T_i)Y_i}{1-\hat{e}(X_i)}\right).$$

Derivation. Under ignorability:

$$\mathbb{E}\left[\frac{TY}{e(X)}\right] = \mathbb{E}\left[\frac{TY(1)}{e(X)}\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{T}{e(X)} \mid X\right] Y(1)\right] = \mathbb{E}[Y(1)],$$

since $\mathbb{E}[T/e(X) \mid X] = e(X)/e(X) = 1$. Similarly for $\mathbb{E}[(1-T)Y/(1-e(X))] = \mathbb{E}[Y(0)]$.

Hajek (normalized) estimator: $\hat{\tau}^{\mathrm{H}} = \frac{\sum_i T_i Y_i/\hat{e}(X_i)}{\sum_i T_i/\hat{e}(X_i)} - \frac{\sum_i (1-T_i)Y_i/(1-\hat{e}(X_i))}{\sum_i (1-T_i)/(1-\hat{e}(X_i))}$. This is more stable (lower variance) than the Horvitz–Thompson version.

Sensitivity to extreme propensity scores. When $e(x)$ is near 0 or 1, the weights $1/e(x)$ or $1/(1-e(x))$ explode, inflating variance. **Trimming** (excluding units with $e(x) < \epsilon$ or $e(x) > 1-\epsilon$) or **truncation** (capping weights) mitigates this at the cost of changing the estimand.

Outcome Regression

The **outcome regression (OR) estimator** models $\mu_t(x) = \mathbb{E}[Y \mid T = t, X = x]$ and computes

$$\hat{\tau}^{\mathrm{OR}} = \frac{1}{n}\sum_{i=1}^{n}[\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)].$$

This is consistent under correct specification of $\mu_t(x)$ and does not require modeling the propensity score. However, it is sensitive to extrapolation in regions where treatment and control groups do not overlap in covariate space.

Doubly Robust (DR) Estimation

The **augmented IPW (AIPW)** estimator combines both approaches:

$$\hat{\tau}^{\mathrm{DR}} = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i) + \frac{T_i(Y_i - \hat{\mu}_1(X_i))}{\hat{e}(X_i)} - \frac{(1-T_i)(Y_i - \hat{\mu}_0(X_i))}{1-\hat{e}(X_i)}\right].$$

Double robustness. $\hat{\tau}^{\mathrm{DR}}$ is consistent if either $\hat{e}$ or $\hat{\mu}_t$ is correctly specified (but not necessarily both).

Proof sketch. Write $\hat{\tau}^{\mathrm{DR}} = \hat{\tau}^{\mathrm{OR}} + \text{correction}$. If $\hat{\mu}_t$ is correct, the residuals $Y_i - \hat{\mu}_t(X_i)$ have conditional mean zero and the correction vanishes. If $\hat{e}$ is correct, the IPW component identifies the ATE independently. The correction term is the product of estimation errors in both models, which is $o_p(1/\sqrt{n})$ when either model is correct.

Semiparametric efficiency. The DR estimator achieves the semiparametric efficiency bound for the ATE under the nonparametric model. Its influence function is

$$\phi(Y, T, X) = \frac{T(Y - \mu_1(X))}{e(X)} - \frac{(1-T)(Y - \mu_0(X))}{1-e(X)} + \mu_1(X) - \mu_0(X) - \tau,$$

and the efficiency bound is $\mathrm{Var}(\phi)/n$.

Matching

Nearest-neighbor matching constructs a matched control for each treated unit:

$$\hat{\tau}^{\text{match}} = \frac{1}{n_1} \sum_{i:T_i=1} \big(Y_i - Y_{j(i)}\big), \quad j(i) = \arg\min_{j:T_j=0} \|X_i - X_j\|.$$

This estimates the ATT. Matching on the propensity score $e(X)$ is valid under ignorability and avoids the curse of dimensionality in X-space.

Bias of matching. With a fixed number of matches M, the bias is $O(n^{-1/d})$ in d dimensions. Bias correction (Abadie–Imbens, 2011) adjusts for the remaining difference between matched and target covariate values.

Instrumental Variables (IV)

When ignorability fails, an **instrument** Z provides identification if:

1. Z affects T (**relevance**).
2. Z affects Y only through T (**exclusion restriction**).
3. Z is independent of confounders (**independence**).

Wald estimator (binary Z, binary T):

$$\hat{\tau}^{\text{IV}} = \frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{T}_{Z=1} - \bar{T}_{Z=0}} = \frac{\text{Reduced form}}{\text{First stage}}.$$

Two-stage least squares (2SLS). Stage 1: regress T on Z and covariates to get $\hat{T}$. Stage 2: regress Y on $\hat{T}$ and covariates. The coefficient on $\hat{T}$ is $\hat{\tau}^{\text{2SLS}}$.

Under monotonicity ($T_i(1) \geq T_i(0)$ for all i), IV identifies the **LATE** (Local Average Treatment Effect) for compliers:

$$\tau^{\text{LATE}} = \mathbb{E}[Y(1) - Y(0) \mid T(1) > T(0)].$$

Difference-in-Differences (DiD)

With panel data (units observed before and after treatment):

$$\hat{\tau}^{\text{DiD}} = (\bar{Y}_{1,\text{post}} - \bar{Y}_{1,\text{pre}}) - (\bar{Y}_{0,\text{post}} - \bar{Y}_{0,\text{pre}}).$$

Parallel trends assumption: absent treatment, treated and control units would have followed the same time trend: $\mathbb{E}[Y_i(0)_{\text{post}} - Y_i(0)_{\text{pre}} \mid T = 1] = \mathbb{E}[Y_i(0)_{\text{post}} - Y_i(0)_{\text{pre}} \mid T = 0]$.

Under parallel trends, DiD identifies the ATT. The assumption is untestable but can be assessed by checking pre-treatment trend similarity.

Regression Discontinuity (RD)

Treatment is assigned by a running variable R_i crossing a threshold c: $T_i = \mathbf{1}(R_i \geq c)$ (sharp RD).

$$\tau^{\text{RD}} = \lim_{r \downarrow c} \mathbb{E}[Y \mid R = r] - \lim_{r \uparrow c} \mathbb{E}[Y \mid R = r].$$

This identifies the ATE at the threshold $R = c$, under continuity of potential outcome means.

Fuzzy RD. When T is not a deterministic function of R but exhibits a discontinuity in $e(R)$ at c, the fuzzy RD estimand is

$$\tau^{\text{FRD}} = \frac{\lim_{r \downarrow c} \mathbb{E}[Y \mid R = r] - \lim_{r \uparrow c} \mathbb{E}[Y \mid R = r]}{\lim_{r \downarrow c} \mathbb{E}[T \mid R = r] - \lim_{r \uparrow c} \mathbb{E}[T \mid R = r]},$$

which is a local Wald (IV) estimator using the threshold as the instrument.

Worked Examples

Example 1: IPW Estimation

Binary treatment, 3 covariate strata. Data:

Stratum	n	n_1	$\bar{Y}_1$	$\bar{Y}_0$	e
A	100	70	15	10	0.7
B	200	100	12	9	0.5
C	100	20	18	14	0.2

IPW estimate of $\mathbb{E}[Y(1)]$: Weight treated observations by $1/e$. Within each stratum, $\sum T_i Y_i / e = n_1 \bar{Y}_1 / e$. Normalized:

$\hat{\mathbb{E}}[Y(1)] = \frac{70 \cdot 15/0.7 + 100 \cdot 12/0.5 + 20 \cdot 18/0.2}{70/0.7 + 100/0.5 + 20/0.2} = \frac{1500+2400+1800}{100+200+100} = \frac{5700}{400} = 14.25.$

$\hat{\mathbb{E}}[Y(0)] = \frac{30 \cdot 10/0.3 + 100 \cdot 9/0.5 + 80 \cdot 14/0.8}{30/0.3 + 100/0.5 + 80/0.8} = \frac{1000+1800+1400}{100+200+100} = \frac{4200}{400} = 10.5.$

$\hat{\tau}^{\text{IPW}} = 14.25 - 10.5 = 3.75.$

Naive (unadjusted): $\hat{\tau}^{\text{naive}} = \frac{70 \cdot 15 + 100 \cdot 12 + 20 \cdot 18}{190} - \frac{30 \cdot 10 + 100 \cdot 9 + 80 \cdot 14}{210} = 13.05 - 11.33 = 1.72$. The naive estimate is biased because treatment is more likely in stratum A (high e) which has lower outcome.

Example 2: Regression Discontinuity

Test score threshold for a scholarship: $c = 75$. Running variable: test score R. Treatment: scholarship ($T = \mathbf{1}(R \geq 75)$). Outcome: college GPA.

Fit local linear regressions on each side of the cutoff using a bandwidth $h = 5$:

Left of cutoff ($R \in [70, 75)$): $\hat{\mathbb{E}}[Y \mid R = 75^-] = 2.80$. Right of cutoff ($R \in [75, 80]$): $\hat{\mathbb{E}}[Y \mid R = 75^+] = 3.15$.

$\hat{\tau}^{\text{RD}} = 3.15 - 2.80 = 0.35$ GPA points.

This is the causal effect of the scholarship for students at the threshold. Validity requires that students cannot precisely manipulate their test scores to cross the threshold (no sorting), so that the density of R is continuous at c. A McCrary density test checks this.

Cross-Links

AI. Off-policy evaluation in reinforcement learning uses IPW and doubly robust estimators to estimate the value of a target policy from data collected under a different behavior policy. This is the sequential generalization of observational causal inference, where the "treatment" is a sequence of actions.

Economics. IV estimation is the workhorse of empirical economics (Angrist–Imbens Nobel, 2021). DiD is standard in policy evaluation (Card–Krueger minimum wage study). RD is used in education (test score thresholds), healthcare (age cutoffs for eligibility), and political science (close elections).

Linear Algebra. 2SLS can be written as a projection: $\hat{\beta}^{\text{2SLS}} = (X^\top P_Z X)^{-1} X^\top P_Z Y$, where $P_Z = Z(Z^\top Z)^{-1} Z^\top$ is the projection onto the instrument space. The first stage is the projection of T onto the column space of Z.

Quantum Mechanics. In quantum causal inference, unmeasured confounders correspond to entangled quantum states shared between parties. Instrumental variable-type arguments (using random measurement settings as instruments) are used to bound causal effects in the presence of quantum confounding, with Bell inequalities as the key identification constraints.

Structural Compression

Invariant

Under ignorability and overlap, the ATE is identified and the DR estimator achieves the semiparametric efficiency bound. When ignorability fails, IV identifies the LATE for compliers, DiD identifies the ATT under parallel trends, and RD identifies the threshold ATE under continuity.

Mechanism

IPW reweights the observed distribution to remove confounding; outcome regression models the conditional expectation surface; the DR estimator combines both and is consistent under misspecification of either nuisance model. IV exploits exogenous variation; DiD exploits time-series variation; RD exploits threshold rules. Each design replaces ignorability with a different structural assumption.

Cross-System Echo

Doubly robust estimation is a special case of semiparametric efficiency theory (Bickel–Klaassen–Ritov–Wellner). In AI, off-policy evaluation uses these estimators for distribution shift correction. In economics, the hierarchy IV > DiD > RD > selection-on-observables reflects decreasing reliance on untestable assumptions, and the credibility revolution is built on matching research designs to credible identification strategies.

Exercises

1. Derive the IPW estimator by showing that $\mathbb{E}[TY/e(X)] = \mathbb{E}[Y(1)]$ under ignorability and overlap. State where each assumption is used.
2. Show that the DR estimator is consistent when the propensity score model is correctly specified but the outcome model is misspecified. Then show the converse.
3. For the Wald estimator $\hat{\tau}^{\text{IV}} = (\bar{Y}_{Z=1} - \bar{Y}_{Z=0})/(\bar{T}_{Z=1} - \bar{T}_{Z=0})$, derive the asymptotic distribution using the delta method and compute the standard error.
4. State the parallel trends assumption for DiD formally using potential outcomes notation. Construct a numerical example where parallel trends holds and DiD recovers the ATT, and another where it fails.
5. Show that the sharp RD estimand $\tau^{\text{RD}} = \lim_{r \downarrow c} \mathbb{E}[Y(1) \mid R = r] - \lim_{r \uparrow c} \mathbb{E}[Y(0) \mid R = r]$ equals the ATE at the threshold under continuity of $\mathbb{E}[Y(t) \mid R = r]$ at $r = c$.
6. Compare the assumptions and estimands of IPW, IV, DiD, and RD in a table. For each method, give one example where it is the appropriate design and explain why the alternatives are not credible.
7. Argue, structurally, why no observational method can replicate the unconditional guarantees of a randomized experiment, and identify the specific untestable assumption that each method introduces as the price for working without randomization.

Policy Evaluation and Structural Identification

Position in Knowledge Graph

System: Statistics. Structural Domain: Causality (Module 11). Node 11.6.

This capstone node extends causal inference from estimating treatment effects to evaluating decision rules (policies). It develops inverse propensity weighting for policy evaluation, the doubly robust AIPW estimator, and the semiparametric efficiency bound for the ATE. The connection to structural equation models links causal inference to mechanism design and reinforcement learning. It synthesizes the entire causality module: potential outcomes (Node 11.1), DAGs (Node 11.2), identification (Node 11.3), experiments (Node 11.4), and observational methods (Node 11.5) into a unified framework for evaluating interventions in complex systems.

Formal Definition

A **policy** (or treatment rule) is a mapping $\pi : \mathcal{X} \to \{0, 1\}$ (deterministic) or $\pi : \mathcal{X} \to [0, 1]$ (stochastic) that assigns treatment based on observed covariates.

The **value** of policy π is

$$V(\pi) = \mathbb{E}[Y(\pi(X))] = \mathbb{E}[\pi(X)\, Y(1) + (1 - \pi(X))\, Y(0)].$$

The **optimal policy** maximizes the value:

$$\pi^* = \arg\max_{\pi} V(\pi).$$

Under ignorability and overlap, the value is identified:

$$V(\pi) = \mathbb{E}[\mathbb{E}[Y \mid T = 1, X]\, \pi(X) + \mathbb{E}[Y \mid T = 0, X]\, (1 - \pi(X))]\,.$$

Intuition

Treatment effect estimation asks: what is the average effect of treatment? Policy evaluation asks a more refined question: who should be treated? The optimal policy treats unit i if and only if $\tau(X_i) = \mathbb{E}[Y(1) - Y(0) \mid X = X_i] > 0$. This connects causal inference to statistical decision theory: the policy is a decision rule, and the value function is the expected utility. The challenge is estimating the value of a proposed policy from data collected under a different (possibly non-randomized) treatment assignment.

Structural Role

Dependencies: Builds on observational treatment estimation (Node 11.5) for the estimators, regression (Node 8.3) for outcome modeling, and decision theory (Node 6.3) for the policy optimization framework. **Enables:** This is the terminal node of Module 11 and the statistics course, connecting to reinforcement learning in AI and mechanism design in economics.

Policy evaluation is the bridge between causal inference and decision-making. It transforms estimates of causal effects into actionable rules, closing the loop from data to intervention.

Mathematical Development

Inverse Propensity Weighted Policy Evaluation

Given data $\{(X_i, T_i, Y_i)\}_{i=1}^n$ collected under a **behavior policy** $b(x) = \mathbb{P}(T = 1 \mid X = x)$, the value of a target policy π is estimated by reweighting:

$$\hat{V}^{\text{IPW}}(\pi) = \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbf{1}(T_i = \pi(X_i))}{T_i\, b(X_i) + (1 - T_i)(1 - b(X_i))}\, Y_i.$$

For a stochastic policy:

$$\hat{V}^{\text{IPW}}(\pi) = \frac{1}{n} \sum_{i=1}^{n} \frac{\pi(X_i)}{b(X_i)}\, T_i\, Y_i + \frac{1 - \pi(X_i)}{1 - b(X_i)}\, (1 - T_i)\, Y_i.$$

This is unbiased for $V(\pi)$ when $b(x)$ is known, and consistent when $b(x)$ is estimated consistently. The variance is high when π and b differ substantially (large importance weights).

Outcome Regression Policy Evaluation

Directly model the outcome:

$$\hat{V}^{\text{OR}}(\pi) = \frac{1}{n} \sum_{i=1}^{n} \left[\pi(X_i)\, \hat{\mu}_1(X_i) + (1 - \pi(X_i))\, \hat{\mu}_0(X_i)\right].$$

Consistent under correct specification of $\mu_t(x)$. No importance weights, so lower variance, but sensitive to model misspecification.

Doubly Robust Policy Evaluation (AIPW)

The **augmented IPW (AIPW)** estimator for policy value is

$$\hat{V}^{\text{DR}}(\pi) = \frac{1}{n} \sum_{i=1}^{n} \left[\pi(X_i) \left(\hat{\mu}_1(X_i) + \frac{T_i(Y_i - \hat{\mu}_1(X_i))}{\hat{b}(X_i)}\right) + (1 - \pi(X_i)) \left(\hat{\mu}_0(X_i) + \frac{(1 - T_i)(Y_i - \hat{\mu}_0(X_i))}{1 - \hat{b}(X_i)}\right)\right].$$

Properties: 1. Doubly robust: consistent if either $\hat{b}$ or $\hat{\mu}_t$ is correctly specified. 2. Achieves the semiparametric efficiency bound when both are correctly specified. 3. Allows the use of flexible/nonparametric estimators for nuisance functions while maintaining $\sqrt{n}$-convergence for the policy value (via cross-fitting).

Semiparametric Efficiency Bound for the ATE

The ATE $\tau = V(\pi \equiv 1) - V(\pi \equiv 0) = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$ is a special case of policy evaluation. The **efficient influence function** for τ under the nonparametric model is

$$\phi(Y, T, X; \tau) = \frac{T(Y - \mu_1(X))}{e(X)} - \frac{(1 - T)(Y - \mu_0(X))}{1 - e(X)} + \mu_1(X) - \mu_0(X) - \tau.$$

The **semiparametric efficiency bound** is

$$\text{Var}(\phi) = \mathbb{E}\left[\frac{\text{Var}(Y \mid T = 1, X)}{e(X)} + \frac{\text{Var}(Y \mid T = 0, X)}{1 - e(X)} + (\tau(X) - \tau)^2\right].$$

Derivation. The influence function is derived from the tangent space of the nonparametric model. The nuisance tangent space is spanned by functions of (T, X); the efficient influence function is the projection of the pathwise derivative of τ onto the orthocomplement of this space in $L^2(P)$.

The three terms in the efficiency bound have clear interpretations: - $\text{Var}(Y \mid T = 1, X)/e(X)$: noise in treated outcomes, amplified when few units are treated. - $\text{Var}(Y \mid T = 0, X)/(1 - e(X))$: noise in control outcomes. - $(\tau(X) - \tau)^2$: treatment effect heterogeneity.

Cross-Fitting and Machine Learning

When $\hat{\mu}_t$ and $\hat{b}$ are estimated by flexible methods (random forests, neural networks), the AIPW estimator requires **cross-fitting** (sample splitting) to avoid Donsker conditions:

1. Split data into K folds.
2. For each fold k, estimate nuisance functions $\hat{\mu}_t^{(-k)}, \hat{b}^{(-k)}$ on the complement.
3. Evaluate the AIPW estimand on fold k using the cross-fitted nuisance estimates.
4. Average across folds.

This **DML (Double/Debiased Machine Learning)** procedure (Chernozhukov et al., 2018) achieves $\sqrt{n}$-consistent, asymptotically normal estimation of τ even when nuisance functions converge at slower rates (e.g., $n^{-1/4}$).

Optimal Policy Learning

The optimal deterministic policy is

$$\pi^*(x) = \mathbf{1}(\tau(x) > 0).$$

Estimation approaches:

1. **Indirect method:** Estimate $\hat{\tau}(x)$ (e.g., via CATE estimation using causal forests or metalearners), then set $\hat{\pi}(x) = \mathbf{1}(\hat{\tau}(x) > 0)$.
2. **Direct method (policy search):** Maximize the estimated policy value over a class of policies Π:

$$\hat{\pi} = \arg\max_{\pi \in \Pi} \hat{V}^{\text{DR}}(\pi).$$

The policy regret $V(\pi^*) - V(\hat{\pi})$ converges at rate $O(n^{-1/2})$ for finite-dimensional policy classes and $O(n^{-1/(2+d)})$ for d-dimensional policy classes under smoothness.

Structural Equations and Policy Counterfactuals

In the **structural equation model (SEM)**:

$$Y = f(T, X, U), \quad T = g(Z, X, V),$$

where U, V are unobserved. A policy π replaces the treatment equation with $T = \pi(X)$. The policy counterfactual is

$$Y^\pi = f(\pi(X), X, U),$$

and the policy value is $V(\pi) = \mathbb{E}[f(\pi(X), X, U)]$.

When the SEM is identified (via instruments, panel structure, or functional form), policy evaluation reduces to computing expectations under the structural functions. This connects to mechanism design: the structural equations describe how the system responds to interventions, and the policy is the designed intervention.

Connection to Reinforcement Learning

In sequential settings, the policy maps state histories to actions: $\pi : \mathcal{H}_t \to \mathcal{A}$. The value function is

$$V(\pi) = \mathbb{E}_\pi \left[\sum_{t=0}^{T} \gamma^t R_t \right],$$

where R_t is the reward at time t and γ is a discount factor. **Off-policy evaluation** estimates $V(\pi)$ from data collected under a behavior policy b, using:

- **Importance sampling:** $\hat{V}^{\text{IS}}(\pi) = \frac{1}{n} \sum_{i=1}^{n} \prod_{t=0}^{T} \frac{\pi(A_{it}|H_{it})}{b(A_{it}|H_{it})} \sum_t \gamma^t R_{it}$. Variance grows exponentially with horizon T.
- **Doubly robust (sequential):** Analogous to the AIPW estimator, using fitted Q-functions as the outcome model and importance ratios as the propensity weights.

This is the sequential generalization of the single-stage causal inference of Nodes 11.1–11.5.

Worked Examples

Example 1: Policy Evaluation via AIPW

Data: $n = 500$, binary treatment. Behavior policy: $b(x) = \text{logit}^{-1}(0.5 + x_1)$. Estimated outcome models: $\hat{\mu}_1(x) = 3 + 2x_1$, $\hat{\mu}_0(x) = 3 + 0.5x_1$.

Target policy: treat if $x_1 > 0$, i.e., $\pi(x) = \mathbf{1}(x_1 > 0)$.

For a specific observation with $X_1 = 0.8$, $T = 1$, $Y = 5.2$:

$\hat{b}(X) = \text{logit}^{-1}(0.5 + 0.8) = \text{logit}^{-1}(1.3) = 0.786$. $\hat{\mu}_1(X) = 3 + 1.6 = 4.6$, $\hat{\mu}_0(X) = 3 + 0.4 = 3.4$. $\pi(X) = 1$ (since $x_1 > 0$).

AIPW contribution: $1 \cdot (4.6 + \frac{1 \cdot (5.2 - 4.6)}{0.786}) + 0 \cdot (\cdots) = 4.6 + 0.763 = 5.363$.

Averaging such terms over all 500 observations gives $\hat{V}^{\text{DR}}(\pi)$.

Example 2: Optimal Policy via CATE Estimation

Estimated CATE: $\hat{\tau}(x) = 1.5 - 3x_1 + 2x_2$. The optimal policy treats when $\hat{\tau}(x) > 0$, i.e., $1.5 - 3x_1 + 2x_2 > 0$, or $x_2 > 1.5x_1 - 0.75$.

For patient profiles: - $(x_1, x_2) = (0.2, 0.5)$: $\hat{\tau} = 1.5 - 0.6 + 1.0 = 1.9 > 0$. Treat. - $(x_1, x_2) = (1.0, 0.0)$: $\hat{\tau} = 1.5 - 3.0 + 0.0 = -1.5 < 0$. Do not treat. - $(x_1, x_2) = (0.5, 0.3)$: $\hat{\tau} = 1.5 - 1.5 + 0.6 = 0.6 > 0$. Treat.

The policy $\hat{\pi}$ personalizes treatment based on predicted benefit, targeting treatment to those who gain most.

Cross-Links

AI. Offline reinforcement learning is the sequential generalization of policy evaluation. Fitted Q-evaluation, importance sampling, and doubly robust methods for sequential decisions all extend the single-stage methods of this node. Contextual bandits are the intermediate case between single-stage and full sequential problems.

Economics. Mechanism design asks: given agents' strategic responses (the structural equations), what policy maximizes social welfare? This is the game-theoretic generalization of optimal policy learning. The welfare analysis of tax policies, subsidies, and regulations uses structural models to evaluate counterfactual policies.

Linear Algebra. The semiparametric efficiency bound involves the inverse of the propensity score as a weighting matrix. In the linear model, the efficient estimator for the ATE reduces to a weighted least squares

problem with weights $w_i = 1/[e(X_i)(1 - e(X_i))]$, and the efficiency bound is the trace of the inverse of the weighted Fisher information matrix.

Quantum Mechanics. Quantum control theory optimizes a policy (a sequence of pulses or measurements) to drive a quantum system to a target state. The value function is the fidelity of the final state, and the "treatment effect" is the change in fidelity from applying vs. not applying a control pulse. Doubly robust estimation has analogues in quantum process tomography with adaptive measurement designs.

Structural Compression

Invariant

Each design (IPW, AIPW, structural models) provides a consistent estimator of the policy value $V(\pi)$ under distinct assumptions; the AIPW estimator achieves the semiparametric efficiency bound and is doubly robust.

Mechanism

IPW reweights observations by the ratio of target-to-behavior policy probabilities; the AIPW estimator adds an outcome-regression correction that reduces variance and provides double robustness. The efficient influence function characterizes the information-theoretic lower bound for estimating the policy value. Cross-fitting enables the use of flexible machine learning estimators for nuisance functions while preserving $\sqrt{n}$-inference for the policy value.

Cross-System Echo

Policy evaluation is the bridge between causal inference and decision theory. In AI, off-policy evaluation in RL is the sequential generalization, with importance sampling replacing IPW and fitted Q-functions replacing outcome regression. In economics, structural estimation of policy counterfactuals is the mechanism-based approach, where the "policy" is a change in the rules of a game and the "value" is social welfare. The doubly robust principle—combine a model of the world with a correction for model error—is universal across these domains.

Exercises

1. Derive the IPW estimator for the policy value $V(\pi)$ starting from $V(\pi) = \mathbb{E}[\pi(X)Y(1)+(1-\pi(X))Y(0)]$ and using ignorability to replace potential outcomes with observables.
2. Show that the AIPW estimator for policy value is doubly robust: consistent if either the propensity score or the outcome model is correctly specified.
3. Derive the efficient influence function for the ATE $\tau = \mathbb{E}[Y(1) - Y(0)]$ by projecting the pathwise derivative onto the orthocomplement of the nuisance tangent space. Verify that the AIPW estimator is the sample analogue.
4. For a binary covariate $X \in \{0, 1\}$ with $\mathbb{P}(X = 1) = 0.6$, $\tau(0) = -1$, $\tau(1) = 3$, compute the value of the optimal policy $\pi^*(x) = \mathbf{1}(\tau(x) > 0)$ and compare it to the value of treating everyone and treating no one.
5. Explain the DML (cross-fitting) procedure for estimating the ATE with machine-learning nuisance estimators. Why is sample splitting necessary? What goes wrong without it?
6. Compare the variance of the IPW, OR, and AIPW estimators for the ATE in a setting where the propensity score is extreme ($e(x)$ near 0 or 1 for some x). Which estimator is most robust to this practical challenge?

7. Argue, structurally, why policy evaluation subsumes treatment effect estimation as a special case, and explain how the connection to reinforcement learning generalizes the single-decision framework of causal inference to sequential decision problems with delayed rewards.