Causal Data Science with Directed Acyclic Graphs (DAGs)

Section 4: Confounding Bias and Surrogate Experiments

Dr. Paul Hünermund

Online Course at Udemy.com

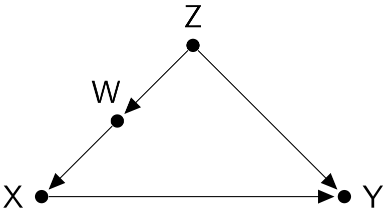
Course Outline

Confounding Bias

- How do we know whether a particular causal effect of an intervention we're interested in is identified (i.e., can be computed) from observational data?
  - We want to know $P(Y|do(X))$ but we cannot manipulate X ourselves
  - We have to transform $P(Y|do(X))$ into a "do-free" expression that we can estimate
- On this problem, graphical models of causation demonstrate their full power
- First, we'll see easily applicable graphical criteria, such as the backdoor and frontdoor criterion, which will help us to decide whether a causal effect is identified
- Second, we'll get to know powerful algorithms, generalizing these graphical criteria, which automate the identification task for us

Confounding Bias (II)

- Problem of confounding: Paths between treatment X and outcome Y that are not emitted by X and which create an association between X and Y that is not causal



- In this example there is one confounding path: $X \leftarrow W \leftarrow Z \rightarrow Y$
- The other path between X and Y $(X \rightarrow Y)$ is emitted by X and therefore causal

Confounding Bias (III)

- So confounding paths create spurious correlations between treatment and outcome. But remember the d-separation criterion

Definition: *d*-separation (Pearl et al., 2016, p. 46)

A path p is blocked by a set of nodes Z if and only if

1. p contains a chain of nodes $A \rightarrow B \rightarrow C$ or a fork $A \leftarrow B \rightarrow C$ such that the middle node B is in Z (i.e., B is conditioned on), or
2. p contains a collider $A \rightarrow B \leftarrow C$ such that the collision node B is not in Z, and no descendant of B is in Z

- We can block biasing paths by conditioning on intermediate variables on these paths that are not colliders or descendants of colliders

Backdoor Adjustment

Definition: The Backdoor Criterion (Pearl et al., 2016, p. 61)

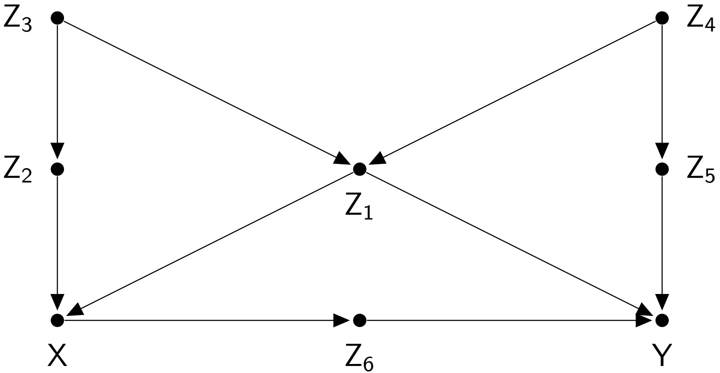
Given an ordered pair of of variables (X, Y) in a directed acyclic graph G, a set of variables Z satisfies the backdoor criterion relative to (X, Y) if no node in Z is a descendant of X, and Z blocks every path between X and Y that contains an arrow into X.

- If a set of variables Z satisfies the backdoor criterion fo X and Y, then the causal effect is given by

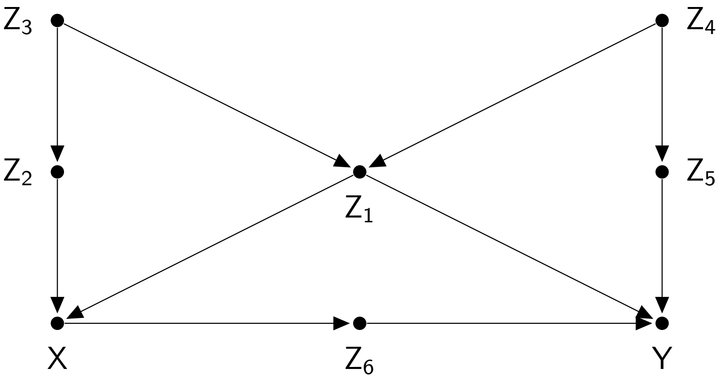
$$P(Y = y | do(X = x)) = \sum_z P(Y = y | X = x, Z = z) P(Z = z)$$

- I.e., condition on the values of Z and average over their joint distribution

Backdoor Adjustment (II)



Backdoor Adjustment (II)

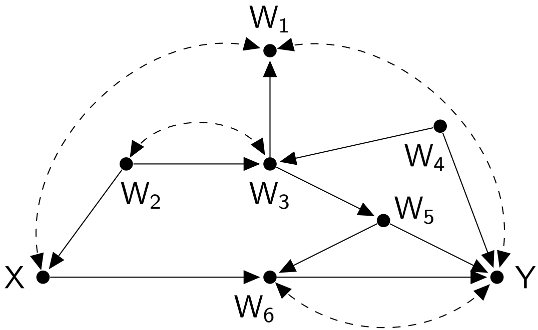


- Minimum sufficient adjustment sets: $\{Z_1, Z_2\}$, $\{Z_1, Z_3\}$, $\{Z_1, Z_4\}$, $\{Z_1, Z_5\}$

Backdoor Adjustment (III)

- Note that it's neither necessary nor sufficient to condition on all variables in the model
  - Sometimes it's even harmful to condition on a (collider) variable, if it opens up a new spurious path
  - In other words, simply putting all variables contained in your data set as controls in your regression won't do!
- In the previous example also, e.g., $\{Z_1, Z_2, Z_3, Z_4\}$ satisfies the backdoor criterion, but we chose to restrict our attention to the smallest sets possible
  - We can choose the adjustment set that is most convenient for us
  - That way we can economize on data collection efforts
  - Also, we save degrees of freedom in small samples
  - The fact that we should find similar causal effects for all admissible adjustment sets has testable implications for model diagnostics

Backdoor Adjustment in Large Graphs



- Backdoor-admissible adjustment sets:

$$Z = \{\{W_2\}, \{W_2, W_3\}, \{W_2, W_4\}, \{W_3, W_4\}, \{W_2, W_3, W_4\}, \{W_2, W_5\}, \\ \{W_2, W_3, W_5\}, \{W_4, W_5\}, \{W_2, W_4, W_5\}, \{W_3, W_4, W_5\}, \{W_2, W_3, W_4, W_5\}\}$$

Estimation

- Once we have found an admissible adjustment set, we can estimate the causal effect by matching, inverse probability weighting, or linear regression (if you're willing to assume linearity)
- The backdoor criterion is thus a convenient way of justifying the *unconfoundedness* assumption from the treatment effects literature (Imbens, 2004)
  - Unconfoundedness (ignorability): treatment has to be independent of potential outcomes conditional on a set of control variables: $(Y^1, Y^0) \perp\!\!\!\perp X|Z$
  - If we have a backdoor admissible adjustment set, we can apply every conventional estimator that is based on unconfoundedness from the treament effects literature

Example: Inverse Probability Weighting

- Let's demonstrate how we can transform the backdoor adjustment formula into an IPW expression

$$
\begin{aligned}
P(Y=y|do(X=x)) &= \sum_z P(Y=y|X=x,Z=z)P(Z=z) \\
&= \sum_z \frac{P(Y=y|X=x,Z=z)P(X=x|Z=z)P(Z=z)}{P(X=x|Z=z)} \\
&= \sum_z \frac{P(Y=y,X=x,Z=z)}{P(X=x|Z=z)}
\end{aligned}
$$

- I.e., we weight the joint distribution of the data by the "propensity score" $P(X=x|Z=z)$

Frontdoor Adjustment

- In the first graph, all admissible adjustment sets contained Z_1.
- What if we don't observe Z_1? In this example we can apply another graphical identification criterion

Definition: The Frontdoor Criterion (Pearl et al., 2016, p. 69)

A set of variables Z is said to satisfy the frontdoor criterion relative to an ordered pair of variables (X, Y) if

1. Z intercepts all directed paths from X to Y
2. There is no unblocked path from X to Z
3. All backdoor paths from Z to Y are blocked by X

Frontdoor Adjustment (II)

Theorem: Frontdoor Adjustment (Pearl et al., 2016, p. 69)

If Z satisfies the frontdoor criterion relative to (X, Y) and if $P(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by the formula

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

- In the graph above (with Z_1 unobserved), Z_6 fulfills the frontdoor criterion
- Frontdoor adjustment essentially boils down to a sequential application of the backdoor criterion

Do-Calculus

- The frontdoor and backdoor criteria might not be always applicable
- In these cases we can apply *do-calculus*
    - Do-calculus is a powerful symbolic machinery that provides a set of inference rules by which sentences involving interventions can be transformed into other sentences (Pearl, 2009; Pearl et al., 2016)
- Do-calculus can be used to solve the confounding problem
    - Apply the rules of *do*-calculus repeatedly until a do-expression is translated into an equivalent expression involving only standard probabilities of observed quantities

Do-Calculus (II)

Theorem: Rules of Do-Calculus (Pearl, 2009, p. 85)

Let G be the directed acyclic graph associated with a [structural] causal model [...], and let $P(\cdot)$ stand for the probability distribution induced by that model. For any disjoint subset of variables X, Y, Z, and W, we have the following rules.

Rule 1 (Insertion/deletion of observations):

$$P(y|do(x), z, w) = P(y|do(x), w) \quad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}.$$

Rule 2 (Action/observation exchange): Illustration

$$P(y|do(x), do(z), w) = P(y|do(x), z, w) \quad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}.$$

Do-Calculus (III))

Rule 3 (Insertion/deletion of actions):

$$P(y|do(x), do(z), w) = P(y|do(x), w) \quad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\,\overline{Z(W)}}},$$

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

- $G_{\overline{X}}$ denotes the graph obtained by deleting from G all arrows pointing to nodes in X
- $G_{\underline{X}}$ denotes the graph obtained by deleting from G all arrows emerging from nodes in X

Example

Completeness of Do-Calculus

- Do-calculus is shown to be *complete*, meaning that if a causal effect is identifiable there exists a sequence of steps applying the rules of do-calculus that transforms the causal effect formula into an expression that includes only observable quantities (Shpitser and Pearl, 2006; Huang and Valtorta, 2006)
  - Put differently, if do-calculus fails, the causal effect is guaranteed to be unidentifiable
- Completeness proofs are notoriously difficult and showing this for the case of do-calculus was a major breakthrough in the literature (Pearl and Mackenzie, 2018)

Automatizing the Identification Task

- We know that do-calculus is complete, but the theorem is only procedural and does not tell us which series of steps leads to the desired solution
- Shpitser and Pearl (2006), building on work by Tian and Pearl (2002), propose an algorithm that automates this task
  - The algorithm takes a description of a DAG as input an returns an expression for a queried causal effect, if it exists
  - Since the algorithm is based on *do*-calculus and therefore complete, if it doesn't return a causal effect expression involving only observable quantities, no such expression exists
  - If you've ever seen how complicated proofs of identification can become in, e.g., structural models in economics, you will immediately appreciate how user-friendly this approach is

Causal Inference by Surrogate Experiments

- Idea behind instrumental variable estimation: identify the causal effect of X on Y by experimental variation in a third variable Z ($=$ instrument)
    - Example: Encouragement designs in development economics (Duflo and Saez, 2003)
- Z-identifiability generalizes this idea for DAGs
    - If $do(x)$ is not possible, can we use do-calculus to reduce $P(y|do(x))$ to an expression that only contains $do(z)$?
    - We can then conduct an experiment in which we randomize Z and identify $P(y|do(x))$ in the joint distribution of the data stemming from that controlled experiment
- Bareinboim and Pearl (2012) generalize the algorithm by Shpitser and Pearl (2006) to check for z-identifiability

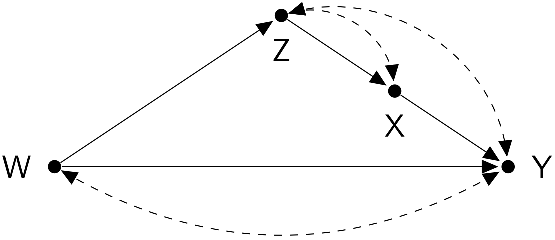
Z-identification

Theorem 3 in Bareinboim and Pearl (2012)

Let X, Y, Z be disjoint sets of variables and let G be the causal diagram. The causal effect $Q = P(y|do(x))$ is $zI\mathcal{D}$ [z-identifiable] in G if and only if one of the following conditions hold:

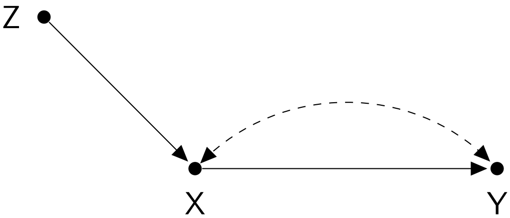
a. Q is identifiable in G; or,
b. There exists $Z' \subseteq Z$ such that the following conditions hold,
   (i) X intercepts all directed paths from Z' to Y and
   (ii) Q is identifiable in $G_{\overline{Z'}}$.

Z-identification (II)



- Here $Q = P(y|do(x))$ is not identifiable by backdoor adjustment because Z is a collider on path $X \leftarrow\!\!\dashrightarrow Z \leftarrow\!\!\dashrightarrow Y$
- But Q is z-identifiable as $\sum_w P(y|do(z), w, x)P(w|do(z))$
- Note that this holds for any value z, so in principle Z doesn't need to vary at all (in reality, we need several levels of Z to ensure enough variation though)

Relation to instrumental variable estimation



- This is the canonical IV setup with endogenous X ($X \leftarrow\!\!-\!-\!-\!\!\rightarrow Y$), Z is both relevant ($Z \rightarrow X$) and excludable ($Z \not\leftarrow\!\!-\!-\!\!\rightarrow Y$)
- But effect of X on Y is *not* z-identifiable because condition (ii) in Theorem 3 of Bareinboim and Pearl (2012) is violated

Relation to instrumental variable estimation (II)

- The fact that the IV-estimator is not non-parametrically identified is well-known in econometrics (Manski, 1990; Balke and Pearl, 1995)
  - E.g., in studies with imperfect compliance only those probands that expect the highest return from a job training program take part, which introduces an unobserved confounder between treatment X and outcome Y. In this case, non-parametric identification is compromised
- To solve this problem, we can either try to put bounds on the causal effect (i.e., resort to "partial identification", Pearl 2009, ch. 8.2.4)
- Or we can introduce functional form assumptions such as *monotonicity* (Imbens and Angrist, 1994) or linearity

References I

Balke, A. and Pearl, J. (1995). Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, 92:1171–1176.

Bareinboim, E. and Pearl, J. (2012). Causal inference by surrogate experiments: z-identifiability. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence*, pages 113–120.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3):815–842.

Huang, Y. and Valtorta, M. (2006). Pearl's calculus of interventions is complete. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence (UAI2006)*.

References II

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, 86(1):4–29.

Imbens, G. W. and Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475.

Manski, C. F. (1990). Nonparametric bounds on treatment effects. *American Economic Review, Papers and Proceedings*, 80:319–323.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, United States, NY, 2nd edition.

Pearl, J., Glymour, M., and Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. John Wiley & Sons Ltd, West Sussex, United Kingdom.

References III

Pearl, J. and Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York.

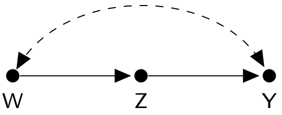
Shpitser, I. and Pearl, J. (2006). Identification of Joint Interventional Distributions in Recursive Semi-Markovian Causal Models. In *Twenty-First National Conference on Artificial Intelligence*.

Tian, J. and Pearl, J. (2002). A general identification condition for causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, pages 567–573, Menlo Park, CA. AAAI Press/The MIT Press.

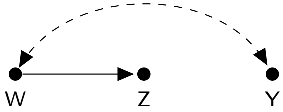
Do-Calculus Rule 2

$$P(y|do(x), do(z), w) = P(y|do(x), z, w) \qquad \text{if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}\underline{Z}}}$$

G



$G_{\underline{Z}}$



Assume we are interested in the query $P(y|do(z), w)$. The graph that results from deleting all arrows emitted by Z in G is denoted as $G_{\underline{Z}}$.

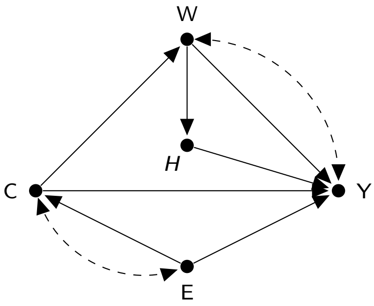
In $G_{\underline{Z}}$, W blocks the only backdoor path between Z and Y: $Z \leftarrow W \dashleftarrow\!\dashrightarrow Y$.

Thus, by d-separation $(Y \perp\!\!\!\perp Z|W)_{G_{\underline{Z}}}$ and therefore the second rule of do-calculus applies.

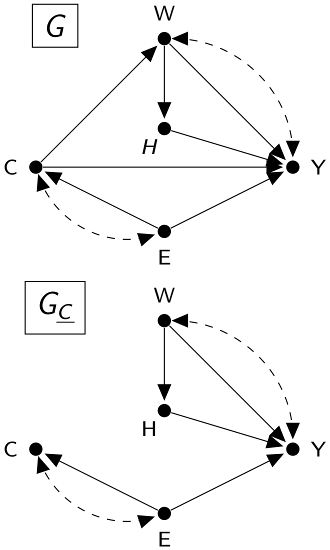
Consequently, we can get rid of the do-operator by setting $P(y|do(z), w) = P(y|z, w)$. The latter expression is estimable from observational data.

Example: Applying Do-Calculus

- Take the stylized example of the college wage premium
    - C: college degree
    - Y: earnings
    - W: occupation
    - H: work-related health
    - E: other socio-economic factors
- Task: Transform $P(y|do(c))$ into a do-free expression by using the rules of do-calculus



Example: Applying Do-Calculus (II)



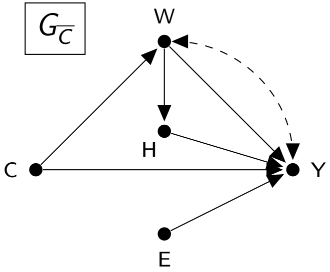
There are two backdoor paths in G, which can both be blocked by E. Conditioning and summing over all values of E yields

$$P(y|do(c)) = \sum_e P(y|do(c), e)P(e|do(c)).$$

By rule 2 of do-calculus

$$P(y|do(c), e) = P(y|c, e), \quad \text{since } (Y \perp\!\!\!\perp C|E)_{G_{\underline{C}}}.$$

Example: Applying Do-Calculus (III)



By rule 3 of do-calculus

$$P(e|do(c)) = P(e), \quad \text{since } (E \perp\!\!\!\perp C)_{G_{\overline{C}}}.$$

It follows that

$$P(y|do(c)) = \sum_e P(y|c,e)P(e).$$

The right-hand-side expression is do-free and can therefore be estimated from observational data.

Back