Methods for Causal Inference
Lecture 11: Front-Door Criterion

Ava Khamseh
School of Informatics
2023-2024
The adjustment formula

T: Drug usage
X: Sex
Y: Recovery

To know how effective the drugs is in the population, compare the hypothetical interventions by which

(i) the drug is administered uniformly to the entire population do(T=1) vs
(ii) complement, i.e., everyone is prevented from taking the drug do(T=0)

Aim: Estimate the difference (Average Causal Effect ACE, aka ATE)

\[ p(Y = 1|do(T = 1)) - p(Y = 1|do(T = 0)) \]
The Backdoor Criterion

Under what conditions does a causal model permit computing the causal effect of one variable on another, from data obtained from passive observations, with no intervention? i.e., Under what conditions is the structure of a causal graph sufficient of computing a causal effect from a given data set? Identifiability

Backdoor Criterion: Given an ordered pair of variables (T,Y) in a DAG G, a set of variables X satisfies the backdoor criterion relative to (T,Y) if:

(i) no node in X is a descendent of T
(ii) X block every path between T and Y that contains an arrow into T

If X satisfies the backdoor criterion then the causal effect of T on Y is given by:

\[ p(Y = y | do(T = t)) = \sum_x p(Y = y | T = t, X = x) p(X = x) \]
The Backdoor Criterion

Under what conditions does a causal model permit computing the causal effect of one variable on another, from data obtained from passive observations, with no intervention? i.e., Under what conditions is the structure of a causal graph sufficient of computing a causal effect from a given data set? **Identifiability**

**Backdoor Criterion:** Given an ordered pair of variables (T,Y) in a DAG G, a set of variables X satisfies the backdoor criterion relative to (T,Y) if:

(i) no node in X is a descendent of T
(ii) X block every path between T and Y that contains an arrow into T

If X satisfies the backdoor criterion then the causal effect of T on Y is given by:

\[
p(Y = y|do(T = t)) = \sum_x p(Y = y|T = t, X = x)p(X = x)
\]

In other words, condition on a set of nodes X such that:

(i) We block all spurious paths between T and Y
(ii) We leave all direct paths from T to Y unperturbed
(iii) We create no new spurious paths (do not unblock any new paths)
Recall ...

- Backdoor does not exhaust all ways of estimating causal effects from a graph.
- Front-door criterion can still be used for patterns that do not satisfy the backdoor criterion.
- Example: Smoking and lung cancer (1970), industry argued to prevent antismoking regulation by suggesting that the correlation could be explained by a carcinogenic genotype that induces a craving for nicotine.
- Recall sensitivity analysis.
- Recall instrumental variable approach.
Instrumental Variable assumptions

- **SUTVA**: Potential outcomes for each individual $i$ are unrelated to the treatment status of other individuals:

\[ Y^{(i)}(Z, T) = Y^{(i)}(Z^{(i)}, T^{(i)}) , \ |Z| = |T| = N \text{ individuals} \]

- Non-zero average/relevant: Treatment assignment $Z$ associated with the treatment

\[ \mathbb{E} \left[ \left( T^{(i)}|z = 1 \right) - \left( T^{(i)}|z = 0 \right) \right] \]

- Treatment assignment $Z$ is random (Z and Y do not share a cause).

\[ \left( Y^{(i)}|z = 1, t \right) = \left( Y^{(i)}|z = 0, t \right) \]

- **Exclusion Restriction**: Any effect of $Z$ on $Y$ is via an effect of $Z$ on $T$, i.e., $Z$ should not affect $Y$ when $T$ is held constant

- **Monotonicity** (increasing encouragement “dose” increases probability of treatment, no defiers):

\[ \left( T^{(i)}|z = 1 \right) \geq \left( T^{(i)}|z = 0 \right) \]
Overview of the course

- **Lecture 1**: Introduction & Motivation, why do we care about causality? Why deriving causality from observational data is non-trivial.
- **Lecture 2**: Recap of probability theory, variables, events, conditional probabilities, independence, law of total probability, Bayes’ rule
- **Lecture 3**: Recap of regression, multiple regression, graphs, SCM
- **Lecture 4-20**:

  - **Causality**
    - **Causal Effect Estimation**
      - Obsv confounders
        - Regression Adjustment
        - Propensity score
      - Unobsv confounders
        - IV
        - Front-door criterion
    - **Casual Discovery**
      - Constraint-based
      - Score-based
      - FCMs

Rubin, Pearl
Pearl’s Front-Door Criterion: An Example

- Fig (a): The graph does not satisfy the backdoor, since the quantity we need to condition on to block the path, i.e. the genotype, is unobserved.
- Fig (b): Additional measurement available: tar deposits in patients’ lungs.
- Fig (b) still does not satisfy the backdoor criterion but we can determine the causal effect:

\[ p(Y = y | do(X = x)) \]

**Figure 3.10** A graphical model representing the relationships between smoking (X) and lung cancer (Y), with unobserved confounder (U) and a mediating variable Z.
**Pearl’s Front-Door Criterion: A crafted example**

**Interpretation 1: Tobacco industry**

**Beneficial effect of smoking:**
15% of smokers have developed lung cancer vs 90.25% of non-smokers within tar and non-tar subgroups, smokers have a much lower percentage of cancer than non-smokers (numbers in the table are engineered to illustrate the point that observations are not to be trusted)

<table>
<thead>
<tr>
<th>Table 3.1</th>
<th>A hypothetical data set of randomly selected samples showing the percentage of cancer cases for smokers and nonsmokers in each tar category (numbers in thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tar 400</td>
<td>No tar 400</td>
</tr>
<tr>
<td>Smokers</td>
<td>Nonsmokers</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>No cancer</td>
<td>380</td>
</tr>
<tr>
<td>(85%)</td>
<td>(5%)</td>
</tr>
<tr>
<td>Cancer</td>
<td>323</td>
</tr>
<tr>
<td>(15%)</td>
<td>(95%)</td>
</tr>
<tr>
<td>57</td>
<td>19</td>
</tr>
</tbody>
</table>
Pearl’s Front-Door Criterion: A crafted example

Interpretation 2: Anti-smoking lobbyists

Smoking **increases** the risk of lung cancer

If one chooses to smoke, then one’s chances of building tar deposits are 95% (380/400) vs 5% (20/400) for the non-smokers.

To evaluate effect of tar, look at **smokers and non-smokers separately**. Tar has harmful effects in both groups: in smokers it increases risk of cancer from 10% to 15% and in non-smokers 90% to 95%. Therefore: Smoking -> tar -> cancer.

Regardless of any natural craving, avoid harmful tar by not smoking.

---

**Table 3.2** Reorganization of the data set of Table 3.1 showing the percentage of cancer cases in each smoking-tar category (numbers in thousands)

<table>
<thead>
<tr>
<th></th>
<th>Smokers 400</th>
<th>Nonsmokers 400</th>
<th>All subjects 800</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tar</td>
<td>No tar</td>
<td>Tar</td>
</tr>
<tr>
<td>No cancer</td>
<td>380</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>323</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(85%)</td>
<td>(90%)</td>
<td>(5%)</td>
</tr>
<tr>
<td>Cancer</td>
<td>57</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>(15%)</td>
<td>(10%)</td>
<td>(95%)</td>
</tr>
</tbody>
</table>
Pearl’s Front-Door Criterion

\[ X \rightarrow Z \text{ is identifiable, since no back path from } X \text{ and } Z: \quad X \leftarrow U \rightarrow Y \leftarrow Z \]

\[ p(Z = z|do(X = x)) = p(Z = z|X = x) \quad \star \]

\[ Z \rightarrow Y \text{ is identifiable, since backdoor from } Z \text{ to } Y: \quad Z \leftarrow X \leftarrow U \rightarrow Y \]

is blocked by conditioning on \( X \):

\[ p(Y = y|do(Z = z)) = \sum_{x} p(Y = y|Z = z, X = x)p(X = x) \quad \star \star \]
Pearl’s Front-Door Criterion

Letting \( z \) be the value \( Z \) takes when setting \( X=x \), from the graph, we have:

\[
p(Y|\text{do}(X = x)) = p(Y|\text{do}(X = x), Z) = p(Y|\text{do}(Z = z))
\]

Then summing over all states \( z \) of \( Z \):

\[
p(Y = y|\text{do}(X = x)) = \sum_{z} p(Y = y, z|\text{do}(X = x))
\]

**Total prob rule**

**Product rule:**

\[
\sum_{z} p(Y = y, z|\text{do}(X = x))p(z|\text{do}(X = x)) = \sum_{z} p(Y = y|\text{do}(Z = z))p(z|\text{do}(X = x))
\]
Pearl’s Front-Door Criterion

\[ p(Z = z \mid do(X = x)) = p(Z = z \mid X = x) \quad \ast \]

\[ p(Y = y \mid do(Z = z)) = \sum_{x'} p(Y = y \mid Z = z, X = x') p(X = x') \quad \ast\ast \]

\[ p(Y = y \mid do(X = x)) = \sum_{z} p(Y = y \mid do(Z = z)) p(Z = z \mid do(X = x)) \]

Using \(\ast\) and \(\ast\ast\) summing over all states \(z\) of \(Z\):

\[ p(Y = y \mid do(X = x)) = \sum_{z} \sum_{x'} p(Y = y \mid Z = z, X = x') p(X = x') p(Z = z \mid X = x) \]
Pearl’s Front-Door Criterion: Which group is right?

\[ p(Y = y|do(X = x)) = \sum_z \sum_{x'} p(Y = y|Z = z, X = x')p(X = x')p(Z = z|X = x) \]

\[ p(Y = 1|do(X = 1)) = p(Y = 1|z = 0, x' = 0)p(x' = 0)p(z = 0|x = 1) \]
\[ + p(Y = 1|z = 0, x' = 1)p(x' = 1)p(z = 0|x = 1) \]
\[ + p(Y = 1|z = 1, x' = 0)p(x' = 0)p(z = 1|x = 1) \]
\[ + p(Y = 1|z = 1, x' = 1)p(x' = 1)p(z = 1|x = 1) \]
\[ = 0.5475 \]

\[ p(Y = 1|do(X = 0)) = 0.5025 \]

**Average Causal Effect ACE:** \[ p(Y = 1|do(X = 1)) - p(Y = 1|do(X = 0)) = 0.045 \]

4.5% increase

<table>
<thead>
<tr>
<th>No cancer</th>
<th>Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smokers 400</td>
<td>Nonsmokers 400</td>
</tr>
<tr>
<td>Tar</td>
<td>No tar</td>
</tr>
<tr>
<td>380</td>
<td>20</td>
</tr>
<tr>
<td>323</td>
<td>18</td>
</tr>
<tr>
<td>(85%)</td>
<td>(90%)</td>
</tr>
<tr>
<td>No cancer</td>
<td>Cancer</td>
</tr>
<tr>
<td>Smokers 400</td>
<td>Nonsmokers 400</td>
</tr>
<tr>
<td>Tar</td>
<td>No tar</td>
</tr>
<tr>
<td>57</td>
<td>2</td>
</tr>
<tr>
<td>(15%)</td>
<td>(10%)</td>
</tr>
</tbody>
</table>
Pearl’s Front-Door Criterion: Which group is right?

\[ p(Y = y | do(X = x)) = \sum_{z} \sum_{x'} p(Y = y | Z = z, X = x') p(X = x') p(Z = z | X = x) \]

\[ p(Y = 1 | do(X = 1)) = p(Y = 1 | z = 0, x' = 0) p(x' = 0) p(z = 0 | x = 1) + p(Y = 1 | z = 0, x' = 1) p(x' = 1) p(z = 0 | x = 1) + p(Y = 1 | z = 1, x' = 0) p(x' = 0) p(z = 1 | x = 1) + p(Y = 1 | z = 1, x' = 1) p(x' = 1) p(z = 1 | x = 1) \]

\[ = 0.5475 \]

\[ p(Y = 1 | do(X = 0)) = 0.5025 \]

**Average Causal Effect (ACE):**

\[ p(Y = 1 | do(X = 1)) - p(Y = 1 | do(X = 0)) = 0.045 \]

4.5% increase

**Table 3.2** Reorganization of the data set of Table 3.1 showing the percentage of cancer cases in each smoking-tar category (numbers in thousands)

<table>
<thead>
<tr>
<th></th>
<th>Smokers</th>
<th>Nonsmokers</th>
<th>All subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tar</td>
<td>No Tar</td>
<td>Tar</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>No cancer</td>
<td>380</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>(85%)</td>
<td>(90%)</td>
<td>(5%)</td>
</tr>
<tr>
<td>Cancer</td>
<td>57</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>(15%)</td>
<td>(10%)</td>
<td>(95%)</td>
</tr>
</tbody>
</table>

Pearl, Causal Inference in Statistics (2016)
Pearl’s Front-Door Criterion: Which group is right?

\[ p(Y = y|do(X = x)) = \sum_z \sum_{x'} p(Y = y|Z = z, X = x')p(X = x')p(Z = z|X = x) \]

\[ p(Y = 1|do(X = 1)) = p(Y = 1|z = 0, x' = 0)p(x' = 0)p(z = 0|x = 1) \]
\[ + p(Y = 1|z = 0, x' = 1)p(x' = 1)p(z = 0|x = 1) \]
\[ + p(Y = 1|z = 1, x' = 0)p(x' = 0)p(z = 1|x = 1) \]
\[ + p(Y = 1|z = 1, x' = 1)p(x' = 1)p(z = 1|x = 1) \]
\[ = 0.5475 \]

\[ p(Y = 1|do(X = 0)) = 0.5025 \]

Average Causal Effect ACE:
\[ p(Y = 1|do(X = 1)) - p(Y = 1|do(X = 0)) = 0.045 \]

4.5% increase

<table>
<thead>
<tr>
<th></th>
<th>Smokers 400</th>
<th>Nonsmokers 400</th>
<th>All subjects 800</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tar</td>
<td>No tar</td>
<td>Tar</td>
</tr>
<tr>
<td>No cancer</td>
<td>380</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>(85%) (90%)</td>
<td>(5%) (10%)</td>
<td>(81%) (19%)</td>
</tr>
<tr>
<td>Cancer</td>
<td>57</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>(15%) (10%)</td>
<td>(95%) (90%)</td>
<td>(19%) (81%)</td>
</tr>
</tbody>
</table>
Pearl’s Front-Door Criterion: Which group is right?

\[ p(Y = y | do(X = x)) = \sum_{z} \sum_{x'} p(Y = y | Z = z, X = x') p(X = x') p(Z = z | X = x) \]

\[
p(Y = 1 | do(X = 1)) = p(Y = 1 | z = 0, x' = 0) p(x' = 0) p(z = 0 | x = 1) + p(Y = 1 | z = 0, x' = 1) p(x' = 1) p(z = 0 | x = 1) + p(Y = 1 | z = 1, x' = 0) p(x' = 0) p(z = 1 | x = 1) + p(Y = 1 | z = 1, x' = 1) p(x' = 1) p(z = 1 | x = 1) = 0.5475
\]

\[
p(Y = 1 | do(X = 0)) = 0.5025
\]

**Average Causal Effect ACE:**

\[ p(Y = 1 | do(X = 1)) - p(Y = 1 | do(X = 0)) = 0.045 \]

4.5% increase
Pearl’s Front-Door Criterion: Which group is right?

\[
p(Y = y|do(X = x)) = \sum_z \sum_{x'} p(Y = y|Z = z, X = x')p(X = x')p(Z = z|X = x)
\]

\[
p(Y = 1|do(X = 1)) = p(Y = 1|z = 0, x' = 0)p(x' = 0)p(z = 0|x = 1)
+ p(Y = 1|z = 0, x' = 1)p(x' = 1)p(z = 0|x = 1)
+ p(Y = 1|z = 1, x' = 0)p(x' = 0)p(z = 1|x = 1)
+ p(Y = 1|z = 1, x' = 1)p(x' = 1)p(z = 1|x = 1)
= 0.5475
\]

\[
p(Y = 1|do(X = 0)) = 0.5025
\]

**Average Causal Effect (ACE):**

\[
p(Y = 1|do(X = 1)) - p(Y = 1|do(X = 0)) = 0.045
\]

4.5% increase
Pearl’s Front-Door Criterion: Which group is right?

\[ p(Y = y|\text{do}(X = x)) = \sum_z \sum_{x'} p(Y = y|Z = z, X = x') p(X = x') p(Z = z|X = x) \]

\[ p(Y = 1|\text{do}(X = 1)) = p(Y = 1|z = 0, x' = 0) p(x' = 0) p(z = 0|x = 1) + p(Y = 1|z = 0, x' = 1) p(x' = 1) p(z = 0|x = 1) + p(Y = 1|z = 1, x' = 0) p(x' = 0) p(z = 1|x = 1) + p(Y = 1|z = 1, x' = 1) p(x' = 1) p(z = 1|x = 1) = 0.5475 \]

\[ p(Y = 1|\text{do}(X = 0)) = 0.5025 \]

**Average Causal Effect ACE:**

\[ p(Y = 1|\text{do}(X = 1)) - p(Y = 1|\text{do}(X = 0)) = 0.045 \]

4.5% increase
Pearl’s Front-Door Criterion: Which group is right?

\[
p(Y = y|do(X = x)) = \sum_{z} \sum_{x'} p(Y = y|Z = z, X = x') p(X = x') p(Z = z|X = x)
\]

\[
p(Y = 1|do(X = 1)) = p(Y = 1|z = 0, x' = 0) p(x' = 0) p(z = 0|x = 1) + p(Y = 1|z = 0, x' = 1) p(x' = 1) p(z = 0|x = 1) + p(Y = 1|z = 1, x' = 0) p(x' = 0) p(z = 1|x = 1) + p(Y = 1|z = 1, x' = 1) p(x' = 1) p(z = 1|x = 1)
\]

\[
\begin{align*}
&= 342/380 + 2/20 + 19/20 + 57/380 \\
&= 0.5475
\end{align*}
\]

\[
p(Y = 1|do(X = 0)) = 0.5025
\]

Average Causal Effect ACE: \[p(Y = 1|do(X = 1)) - p(Y = 1|do(X = 0)) = 0.045\]

4.5% increase

Table 3.2 Reorganization of the data set of Table 3.1 showing the percentage of cancer cases in each smoking-tar category (numbers in thousands)

<table>
<thead>
<tr>
<th></th>
<th>Smokers 400</th>
<th>Nonsmokers 400</th>
<th>All subjects 800</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tar</td>
<td>No tar</td>
<td>Tar</td>
</tr>
<tr>
<td>No cancer</td>
<td>380</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>323</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(85%)</td>
<td>(90%)</td>
<td>(5%)</td>
</tr>
<tr>
<td>Cancer</td>
<td>57</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>(15%)</td>
<td>(10%)</td>
<td>(95%)</td>
</tr>
</tbody>
</table>
Pearl’s Front-Door Adjustment

Front-door criterion: A set of variables Z is said to satisfy the front-door criterion relative to (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

Front-door adjustment: If Z satisfied the front-door 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:

\[ p(y|do(x)) = \sum_z p(z|x) \sum_{x'} p(y|x', z)p(x') \]
Pearl’s Do Calculus

Do-calculus: Contains, as subsets:
- Backdoor criterion
- Front-door criterion

Allows analysis of more intricate structure beyond back- and front-door

Uncovers all causal effects that can be identified from a given causal graph

Power of causal graphs is not just representation but towards discovery of causal information
Causal Inference

- **Model** a causal inference problem with assumptions manifest in Causal Graphical Models [Pearl]

- **Identify** an expression for the causal effect under these assumptions ("causal estimand"), [Pearl]

- **Estimate** the expression using statistical methods such as matching or instrumental variables, [Rubin’s Potential Outcomes]

- **Verify** the validity of the estimate using a variety of robustness checks.
Causal Inference: Packages and simulations

Simple DoWhy tutorials on my GitHub ‘Causality in Biomedicine’

DoWhy tutorials:
https://www.pywhy.org/dowhy/v0.9.1/index.html
https://github.com/py-why/dowhy

CausalGraphicalModels Tutorials:
https://github.com/ijmbarr/causalgraphicalmodels


Also see ML extensions to DoWhy, e.g. EconML:
https://github.com/microsoft/EconML
Methods for Causal Inference
Lecture 11: Front-Door Criterion

Ava Khamseh
School of Informatics
2023-2024