Problem Set 1 Solutions

6.S091: Causality

IAP 2023

Problem 1: Interventions and Adjustment

Preliminaries

(a) $\mathbb{P}_{\mathcal{X}} = \mathsf{Ber}(U; 0.5) \times \mathsf{Ber}(A; U/4) \times \mathsf{Ber}(M; 1/2 + A/10) \times \mathsf{Ber}(Y; M/2 + U/4)$

(b)

- $\mathbb{P}_{\mathcal{X}}(Y = 1) \approx 0.381$
- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid M = 0, A = 0) \approx 0.107$
- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid M = 0, A = 1) = 0.25$

Interventional

(c) $\mathsf{Ber}(U; 0.5) \times \delta_1(A) \times \mathsf{Ber}(M; 0.6) \times \mathsf{Ber}(Y; U/4 + M/2)$

(d)

- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid \mathtt{do}(A = 1)) = 0.425$
- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid \mathtt{do}(A = 0)) = 0.375$

Backdoor Adjustment

(e) The probabilities $\mathbb{P}_{\mathcal{X}}(Y \mid A = a, U = u)$ for $(a, u) \in \{0, 1\} \times \{0, 1\}$ are:

- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = 0, U = 0) = 0.25$
- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = 0, U = 1) = 0.5$
- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = 1, U = 0) = 0.3$
- $\mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = 1, U = 1) = 0.55$

$$\begin{aligned}
\mathbb{P}_{\mathcal{X}}(Y = 1 \mid \mathtt{do}(A = 1)) &= \sum_{u \in \{0,1\}} \mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = 1, U = u)\mathbb{P}_{\mathcal{X}}(U = u) \\
&= 0.3 \cdot 0.5 + 0.55 \cdot 0.5 = 0.425
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}_{\mathcal{X}}(Y = 1 \mid \mathtt{do}(A = 0)) &= \sum_{u \in \{0,1\}} \mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = 0, U = u)\mathbb{P}_{\mathcal{X}}(U = u) \\
&= 0.25 \cdot 0.5 + 0.5 \cdot 0.5 = 0.375
\end{aligned}$$

Note: as correctly pointed out by many students, $\mathbb{P}_{\mathcal{X}}(Y=1 \mid A=1, U=0)$ is technically undefined, since it involves conditioning on an event of probability zero. In practice, this means that, even from infinite data, one could not estimate $\mathbb{P}_{\mathcal{X}}(Y=1 \mid A=1, U=0)$. This phenomenon is a violation of the ***overlap*** *assumption required for adjustment. However, for the context of this problem, one may compute $\mathbb{P}_{\mathcal{X}}(Y=1 \mid A=1, U=0)$ without dividing by $\mathbb{P}_{\mathcal{X}}(A=1, U=0)$. In particular:*

$$\begin{aligned}
\mathbb{P}_{\mathcal{X}}(Y=1 \mid A=1, U=0) &= \sum_{m \in \{0,1\}} \mathbb{P}_{\mathcal{X}}(Y=1 \mid A=1, U=0, M=m)\mathbb{P}_{\mathcal{X}}(M=m \mid A=1, U=0) \\
&= \sum_{m \in \{0,1\}} \mathbb{P}_{\mathcal{X}}(Y=1 \mid U=0, M=m)\mathbb{P}_{\mathcal{X}}(M=m \mid A=1) \\
&= 0 \cdot 0.4 + 0.5 \cdot 0.6 \\
&= 0.3
\end{aligned}$$

Frontdoor Adjustment

(f)

$$\begin{aligned}
\mathbb{P}_{\mathcal{X}}(Y=1 \mid \texttt{do}(A=1)) &= \sum_{m \in \{0,1\}} \mathbb{P}_{\mathcal{X}}(M=m \mid A=1) \sum_{a' \in \{0,1\}} \left(\mathbb{P}_{\mathcal{X}}(Y \mid M=m, A=a')\mathbb{P}_{\mathcal{X}}(A=a')\right) \\
&= 0.038 + 0.013 + 0.319 + 0.056 = 0.425
\end{aligned}$$

using the terms:

- $\mathbb{P}_{\mathcal{X}}(M=0 \mid A=1)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=0, A=0)\mathbb{P}_{\mathcal{X}}(A=0)\right) \approx 0.038$
- $\mathbb{P}_{\mathcal{X}}(M=0 \mid A=1)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=0, A=1)\mathbb{P}_{\mathcal{X}}(A=1)\right) \approx 0.013$
- $\mathbb{P}_{\mathcal{X}}(M=1 \mid A=1)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=1, A=0)\mathbb{P}_{\mathcal{X}}(A=0)\right) \approx 0.319$
- $\mathbb{P}_{\mathcal{X}}(M=1 \mid A=1)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=1, A=1)\mathbb{P}_{\mathcal{X}}(A=1)\right) \approx 0.056$

$$\begin{aligned}
\mathbb{P}_{\mathcal{X}}(Y=1 \mid \texttt{do}(A=0)) &= \sum_{m \in \{0,1\}} \mathbb{P}_{\mathcal{X}}(M=m \mid A=01) \sum_{a' \in \{0,1\}} \left(\mathbb{P}_{\mathcal{X}}(Y \mid M=m, A=a')\mathbb{P}_{\mathcal{X}}(A=a')\right) \\
&= 0.047 + 0.016 + 0.266 + 0.047 = 0.375
\end{aligned}$$

using the terms:

- $\mathbb{P}_{\mathcal{X}}(M=0 \mid A=0)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=0, A=0)\mathbb{P}_{\mathcal{X}}(A=0)\right) \approx 0.047$
- $\mathbb{P}_{\mathcal{X}}(M=0 \mid A=0)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=0, A=1)\mathbb{P}_{\mathcal{X}}(A=1)\right) \approx 0.016$
- $\mathbb{P}_{\mathcal{X}}(M=1 \mid A=0)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=1, A=0)\mathbb{P}_{\mathcal{X}}(A=0)\right) \approx 0.266$
- $\mathbb{P}_{\mathcal{X}}(M=1 \mid A=0)\left(\mathbb{P}_{\mathcal{X}}(Y \mid M=1, A=1)\mathbb{P}_{\mathcal{X}}(A=1)\right) \approx 0.047$

Problem 2: Moral separation implies d-separation

In this problem, we prove the converse of Theorem 2.2 from Lecture 2. In particular, we wish to show that $\mathcal{I}^m_{\perp\!\!\!\perp}(\mathcal{G}) \subseteq \mathcal{I}_{\perp\!\!\!\perp}(\mathcal{G})$. Suppose there is a d-connecting path γ from $\mathbf{A}$ and $\mathbf{B}$ in $\mathcal{G}$ given $\mathbf{S}$.

(a) Show that all nodes in γ are in $\mathbf{V} = \overline{\text{an}}_{\mathcal{G}}(\mathbf{A} \cup \mathbf{B} \cup \mathbf{S})$.

Consider node x_k in γ. There are two cases to consider:

- x_k is a collider. In this case, in order for γ to be unblocked, $\overline{\text{de}}_{\mathcal{G}}(x_k) \cap \mathbf{S} \neq \varnothing$; i.e., x_k or one of its descendants must be in $\mathbf{S}$. This implies that $x_k \in \overline{\text{an}}_{\mathcal{G}}(\mathbf{S})$.
- x_k is not a collider. In this case, x_k must have an outgoing arrow along the path γ. WLOG, assume that if we traverse γ in this direction the final node is in $\mathbf{B}$. There are two possible cases as we traverse this path:
  - There are not colliders. In this case, all arrows must be pointing in the same direction (towards the ending node in $\mathbf{B}$), and thus $x_k \in \overline{\text{an}}_{\mathcal{G}}(\mathbf{B})$.
  - There is at least one collider. Consider the first collider encountered when traversing γ in the direction of the outgoing edge from x_k. Let's call this x_j. As described above, $x_j \in \overline{\text{an}}_{\mathcal{G}}(\mathbf{S})$. All edges from x_k to x_j must be in the same direction (since we did not encounter a collider); hence, x_k is an ancestor of x_j. This further implies that $x_k \in \overline{\text{an}}_{\mathcal{G}}(\mathbf{S})$.

  Since we have shown that an arbitrary node $x_k \in \mathbf{V}$, we conclude that all nodes in γ must be in $\mathbf{V}$.

(b) Show that there is a path in $\overline{\mathcal{G}[\mathbf{V}]}$ from $\mathbf{A}$ to $\mathbf{B}$ which does not pass through $\mathbf{S}$.

First note that all nodes in γ are in $\overline{\mathcal{G}[\mathbf{V}]}$, since by part (a), all nodes in γ are in $\mathbf{V}$. We construct a new path γ' as follows. For each node $x_k \in \gamma$:

- If $x_k \notin \mathbf{S}$, add x_k to γ'.
- If $x_k \in \mathbf{S}$, x_k must be a collider on the path γ (otherwise γ would not be d-connecting in $\mathcal{G}$). This means that x_{k-1} and x_{k+1} must have an edge in $\overline{\mathcal{G}[\mathbf{V}]}$ (since we moralized the parents of x_k). Do not add x_k to γ' and continue to $x_{k+1} \in \gamma$.

Note that the first node and last node in γ are clearly not colliders, so they are also in γ'. Thus, γ' is a path from $\mathbf{A}$ to $\mathbf{B}$ that does not pass through $\mathbf{S}$.

Problem 3: Simpson's paradox

(a)

$$\mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = N) \approx 0.78$$
$$\mathbb{P}_{\mathcal{X}}(Y = 1 \mid A = O) \approx 0.83$$

(b)

$$\mathbb{P}_{\mathcal{X}}(Y = 1 \mid \texttt{do}(A = N)) \approx 0.83$$
$$\mathbb{P}_{\mathcal{X}}(Y = 1 \mid \texttt{do}(A = O)) \approx 0.78$$

Problem 4: Instrumental Variables

β_{wa}	$\hat{\beta}_{wa}$	$\hat{\beta}_{wy}$	$\hat{\beta}_{wa}/\hat{\beta}_{wy}$
0.05	0.326467	2.490757	7.629419
0.50	0.729087	5.542779	7.602360
5.00	4.527291	33.847992	7.476433