Why Causal Inference?

Caleb Leedy[1]

March 22, 2023

[1] Thanks to Dr. Kim for sharing some of his notes.

Why is causal inference important?

- Often the questions that we ask are important in their respective disciplines.
- Causal inference tries to get at *why* something happened.

What is causal inference?

- Causal inference is the study and identification of treatment effects.
- We want to know how an intervention changed an outcome.

Frameworks for Causal Inference

- Potential Outcomes
- Directed Acyclic Graphs (DAGs)

Background and Notation

Potential Outcomes

- Outcome variable Y
- Treatment assignment $A \in \{0, 1\}$
- Covariates X
- The *potential outcomes* of Y for individual i are

$$Y_i(0) \text{ and } Y_i(1).$$

Potential Outcomes

- Often the goal is to understand: $Y(1) - Y(0)$.
- For each individual i, we only observe *either* $Y_i(0)$ *or* $Y_i(1)$.

Assumptions

- Stable unit treatment value assumption (SUTVA):

$$Y_i = Y_i(1)A_i + Y_i(0)(1 - A_i).$$

- No Unmeasured Confounders (NUC):

$$Y(1), Y(0) \perp A \mid X$$

- Positivity: For X with $p(X) > 0$,

$$0 < \Pr(A = 1 \mid X) < 1.$$

Goals:

- Average treatment effect (ATE):

$$E[Y(1) - Y(0)]$$

- Average treatment effect of the treated (ATT):

$$E[Y(1) - Y(0) \mid A = 1]$$

- Conditional average treatment effect (CATE):

$$E[Y(1) - Y(0) \mid X]$$

- Individual treatment effect (ITE):

$$Y_i(1) - Y_i(0)$$

Average Treatment Effect

- Consider the ATE denoted by

$$\tau = E[Y(1) - Y(0)].$$

Outcome Model

- From our assumptions, $E[Y(a)] = E[E[Y \mid X, A = a]]$.
- We can construct a model, $m_a(X) = E[Y \mid X, A = a]$ and estimate τ using
$$\hat{\tau}_{\text{OR}} = \hat{m}_1(X) - \hat{m}_0(X).$$
- If the model for $\hat{m}_a(X)$ is correctly specified, $\hat{\tau}_{\text{OR}}$ is consistent.

Propensity Score

- Instead of modeling the outcome $E[Y \mid X]$ we can model the response probability $\pi(X) = \Pr(A = 1 \mid X)$.

Inverse Propensity Score Weighting

- If $\pi(X)$ is known then we can estimate τ with $\hat{\tau}_{\text{IPW}} = \hat{\mu}_1 - \hat{\mu}_0$.
- We estimate $E[Y(1)]$ with

$$\hat{\mu}_1 = n^{-1} \sum_{i=1}^{n} \frac{A_i Y_i}{\pi(X_i)}.$$

- We can estimate $E[Y(0)]$ with

$$\hat{\mu}_0 = n^{-1} \sum_{i=1}^{n} \frac{(1 - A_i) Y_i}{1 - \pi(X_i)}.$$

- This is similar to a Horvitz-Thompson estimator.

Result

- The IPW estimator is consistent.

$$
\begin{aligned}
E\left[\frac{AY}{\pi(X)}\right] &= E\left[\frac{AY(1)}{\pi(X)}\right] \\
&= E\left[E\left[\frac{AY(1)}{\pi(X)}\right] \mid Y(1), X\right] \\
&= E\left[E\left[\frac{AY(1)}{\pi(X)}\right] \mid X\right] \\
&= E\left[\frac{\pi(X)Y(1)}{\pi(X)}\right] \\
&= E[Y(1)].
\end{aligned}
$$

Double Robust Estimation

- We can combine our outcome model and response model together to get a doubly robust estimator: $\hat{\tau}_{\text{DR}} = \hat{\mu}_{1,\text{DR}} - \hat{\mu}_{0,\text{DR}}$ where

$$\hat{\mu}_{1,\text{DR}} = n^{-1} \sum_{i=1}^{n} \left(m_1(x_i) + \frac{A_i}{\hat{\pi}(X_i)} (Y_i - m_1(x_i)) \right)$$

and

$$\hat{\mu}_{0,\text{DR}} = n^{-1} \sum_{i=1}^{n} \left(m_0(x_i) + \frac{1 - A_i}{1 - \hat{\pi}(X_i)} (Y_i - m_0(x_i)) \right).$$

Result

- $\hat{\tau}_{\text{DR}}$ is consistent if either the outcome or the response model is true.
- If the outcome model is correctly specified ($m_1(x) = E[Y(1) \mid X]$), then

$$
\begin{aligned}
E[\hat{\mu}_{1,\text{DR}}] &= n^{-1} \sum_{i=1}^{n} \left(E[m_1(x_i)] + E\left[\frac{A_i}{\hat{\pi}(X_i)} E[Y_i - m_1(x_i) \mid X] \right] \right) \\
&= n^{-1} \sum_{i=1}^{n} E[m_1(x_i)] \\
&= E[Y(1)].
\end{aligned}
$$

Result

- If the response model is correctly specified ($\hat{\pi}(X) = \pi(X)$), then

$$
\begin{aligned}
E[\hat{\mu}_{1,\mathrm{DR}}] &= n^{-1} \sum_{i=1}^{n} E\left[\left(1 - \frac{A_i}{\pi(X_i)}\right) m_1(x_i) + \frac{A_i}{\pi(X_i)} Y_i\right] \\
&= n^{-1} \sum_{i=1}^{n} E\left[E\left[\left(1 - \frac{A_i}{\pi(X_i)}\right) m_1(x_i) + \frac{A_i}{\pi(X_i)} Y_i \mid X\right]\right] \\
&= n^{-1} \sum_{i=1}^{n} E\left[\left(1 - \frac{\pi(X_i)}{\pi(X_i)}\right) m_1(x_i) + \frac{\pi(X_i)}{\pi(X_i)} Y_i(1)\right] \\
&= E[Y(1)].
\end{aligned}
$$

Next Week

- Estimation of ATE
- Instrumental variables (IVs)