

# Causal inference

Matching, Weighting, and More

---

Brad Hackinen



## Propensity Scores

---

## Propensity Scores

**Propensity Scores** A *propensity score* is the probability that the unit of observation receives treatment, given its observed covariates.

$$PS(X_i) = P[D_i = 1 | X_i]$$

## Estimating Propensity Scores

Instead of estimating the outcome as a function of treatment and covariates, we estimate treatment as a function of covariates.

For example, it is common to use *Logistic Regression* to estimate propensity scores:

$$\widehat{PS}(X) = \text{Logistic}(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots)$$

## **What can we do with Propensity Scores?**

Propensity scores are useful for:

1. Checking for common support between treated and untreated groups
2. Controlling for observable covariates

## Propensity Scores and Common Support

Recall that causal inference can fail badly outside of the common support of the treated and untreated groups because we are forced to make extrapolations about counterfactual outcomes with little or no data.

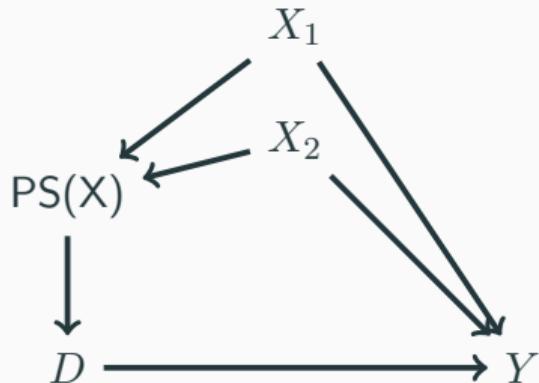
- When we have multiple covariates, it can be hard to visualize common support.
- Propensity scores provide a one-dimensional summary of covariates that we can use to check for common support.
- We just need to check if there are observations with extremely high or low propensity scores

## Controlling for Covariates with Propensity Scores

An important but non-obvious fact about causal inference is that we can control for all confounders simultaneously if we can control for the probability of treatment.

**Intuition:** Confounding bias is only possible because the confounder alters the probability of treatment. Thus, comparing units with the same probability of treatment removes the possibility of confounding bias.

## Propensity Score as a DAG



**Q:** What does the Back-Door criterion imply about controlling for  $PS(X)$ ?

## **Causal Inference with Propensity Scores**

Once we have estimated propensity scores, we can use them to estimate average treatment effects in two ways:

1. Propensity Score Matching
2. Inverse Propensity Score Weighting

## Propensity Score Matching

Propensity scores can be used to match treated and untreated units with similar values of the propensity score.

- The simplest approach is *nearest neighbor matching*, where each unit is matched to the unit from the other treatment group with the closest propensity score.
- This is a flexible way of controlling for the propensity score without making parametric assumptions about the relationship between the propensity score and outcomes (e.g., linearity).

## Inverse Propensity Score Weighting (IPW)

Inverse Propensity Score Weighting (IPW) uses propensity scores to construct weights for each observation such that:

1. The total weight on treated and untreated units in each region of the propensity score is the same.
2. The total weight in each region matches the target distribution we want to estimate.

If the propensity scores have a small number of discrete values, this is equivalent to post-stratification by propensity score.

## Intuition via Post-Stratification

Recall the formula for estimating the ATE with post-stratification:

$$\widehat{ATE} = \underbrace{\frac{1}{\sum_{s=1}^S N_s} \sum_{s=1}^S N_s \left( \underbrace{\frac{1}{N_{T_s}} \sum_{i \in T_s} Y_i - \frac{1}{N_{U_s}} \sum_{i \in U_s} Y_i}_{\text{SDO for strata } s} \right)}_{\text{Weighted average of strata SDO estimates}}$$

- Treated observations are multiplied by  $N_s/N_{T_s}$ 
  - (This is the inverse probability that a unit in strata  $s$  is treated)
- Untreated observations are multiplied by  $N_s/N_{U_s}$ 
  - (This is the inverse probability that a unit in strata  $s$  is untreated)

Similar relationships exist if we are estimating the ATT or ATU.

## Choosing an Estimand

Suppose  $\hat{\pi}_i$  is the predicted propensity score for unit  $i$ . Then we can estimate the ATE, ATT, or ATU by weighting observations as follows:

---

Estimand	Treated Weight	Control Weight	Explanation
ATE	$\frac{1}{\hat{\pi}_i}$	$\frac{1}{1-\hat{\pi}_i}$	Inverse Propensity Score
ATT	1	$\frac{\hat{\pi}_i}{1-\hat{\pi}_i}$	Re-Weight by $\hat{\pi}_i$
ATU	$\frac{1-\hat{\pi}_i}{\hat{\pi}_i}$	1	Re-Weight by $1 - \hat{\pi}_i$

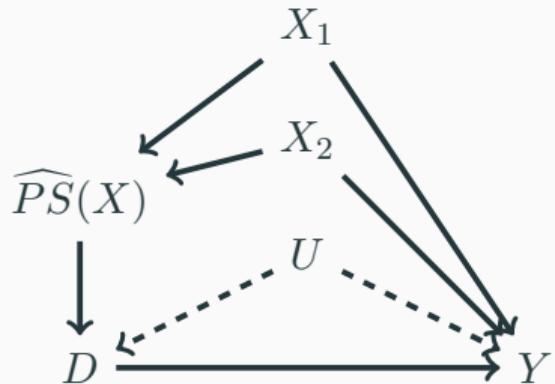
---

## Propensity Scores are not Magic

Some descriptions of propensity score methods make them sound like they magically estimate causal effects. However, if you have to *estimate* the propensity score, they are simply another way to control for observed variables.

*Controlling for the estimated propensity score is equivalent to controlling for the set of covariates you used to estimate the propensity score.*

## Estimated Propensity Score as a DAG



**Q:** What does the Back-Door criterion imply about controlling for  $\widehat{PS}(X)$ ?

## Doubly Robust Estimators

Doubly robust estimators combine propensity score weighting with outcome imputation to estimate treatment effects:

1. Estimate an outcome model to predict counterfactual outcomes
2. Estimate a propensity score model to predict treatment assignment, and use it to compute appropriate weights
3. Compute the estimate as a weighted average of the difference between observed and predicted outcomes

The result is an estimator that is consistent if *either* the propensity score model *or* the outcome model is correctly specified.

## R Demo

PropensityScoresForCI\_blank.Rmd

Goals:

- Estimate propensity scores with logistic regression
- Use propensity scores to check for common support
- Use propensity scores to estimate treatment effects with weighting