Model-based matching for causal inference in observational studies

Kellie Ottoboni
with Philip B. Stark

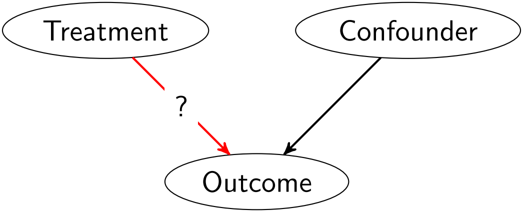
Department of Statistics, UC Berkeley
Berkeley Institute for Data Science

March 15, 2016

Outline

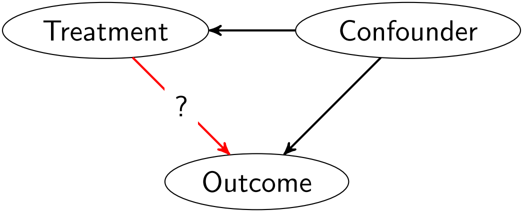
Observational Studies vs Experiments

- **Problem:** Estimate the causal effect of a treatment on outcome of interest
- In randomized experiments, treatment is assigned to individuals at random.
- In observational studies, the way individuals select into treatment groups is unknown.



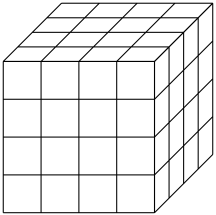
Observational Studies vs Experiments

- **Problem:** Estimate the causal effect of a treatment on outcome of interest
- In randomized experiments, treatment is assigned to individuals at random.
- In observational studies, the way individuals select into treatment groups is unknown.



Matching

- **Ideal:** group individuals by important confounders to estimate subgroup treatment effects and then average over subgroups
- **Reality:** many covariates, perhaps continuous, make it difficult to stratify

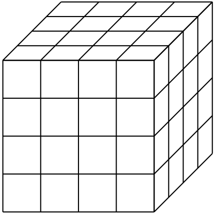


Matching

- If d covariates are split into k bins, we have d^k groups.
- To guarantee that we have at least one treated and one control in each group with 95% probability, we need

$$n \geq \frac{2\log(1-(0.95)^{1/k^{d+1}})}{\log(\frac{k^d-1}{k^d})}$$

- If $d = 5$ and $k = 2$, $n \geq 225$.
- If $d = 10$ and $k = 2$, $n \geq 10,844$.



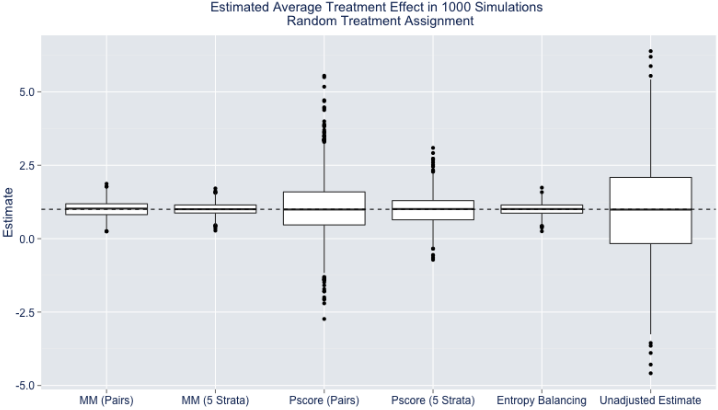
Matching

- **Solution:** use a one-dimensional score to match or group individuals
- Stratify according to the "best" prediction of the response, $\hat{Y}$, based on all covariates except for the treatment

Outline

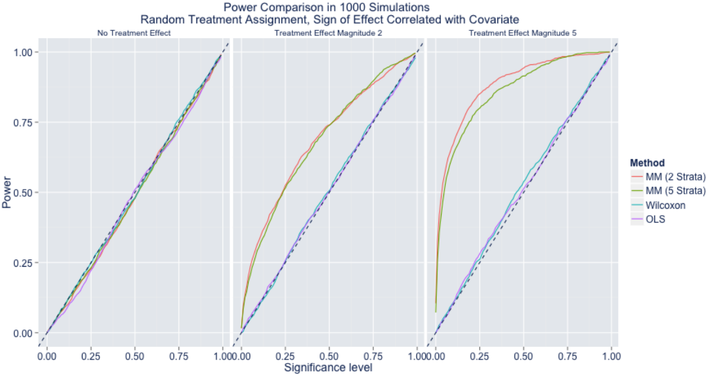
Model-based Matching

- Under standard assumptions (conditional independence of treatment and potential outcomes given X), the **average treatment effect** is nonparametrically identified
- Estimate it using the difference in average residuals, $Y - \hat{Y}$, between treated and controls



Model-based Matching

- Use stratified permutation test to test the **strong null hypothesis** of no treatment effect whatsoever
- Stratifying on $\hat{Y}$ allows us to detect non-constant and non-linear treatment effects



Outline

Outline

Future Directions

- Do different test statistics give greater power when the treatment effect is nonlinear?
- What is the optimal way to stratify?
- How to quantify uncertainty – standard errors and confidence intervals?