Causal Inference and ML Experiments

Rayid Ghani and Kit Rodolfa

Carnegie Mellon University

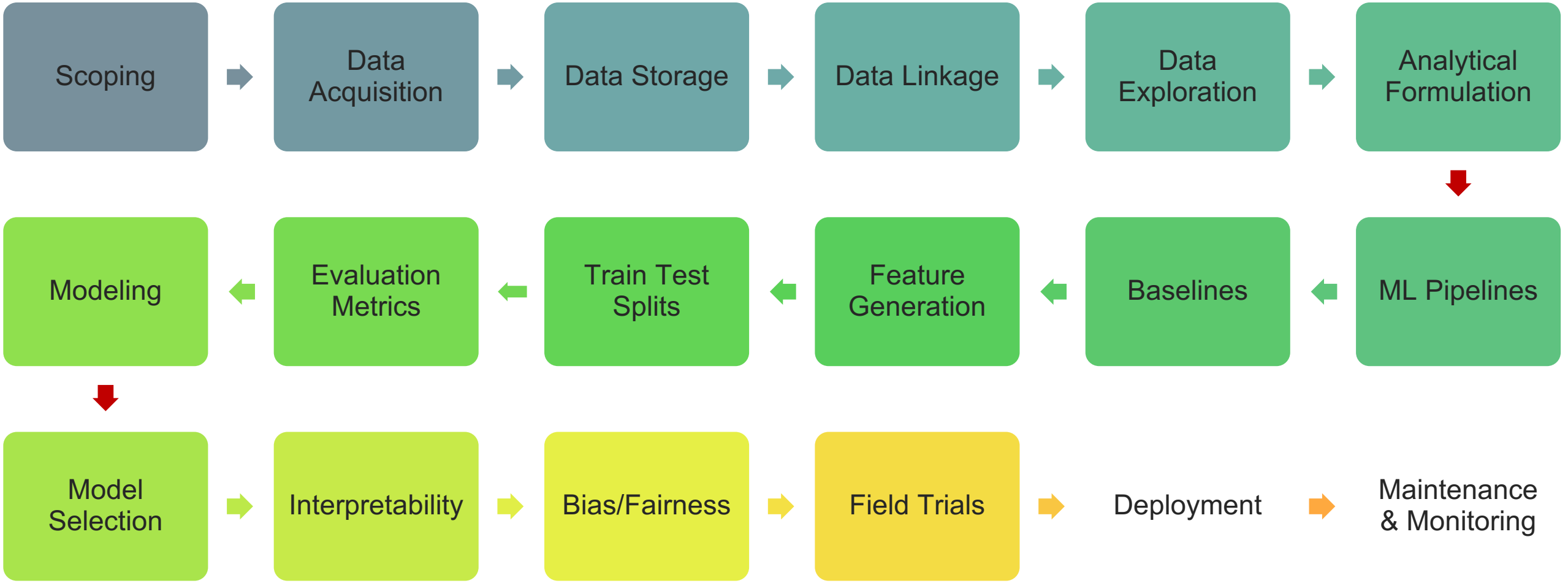
ML
MACHINE LEARNING
DEPARTMENT

Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

Things to cover today

- Reminders for list of things to do for your project
- Causal Inference 001
- With Observational Data
  - Methods
  - Examples
- Experiments
  - Standard Randomized Controlled Trials
    - Trial Design
    - Examples
  - Validating ML models/systems
    - Trial Design
    - Examples

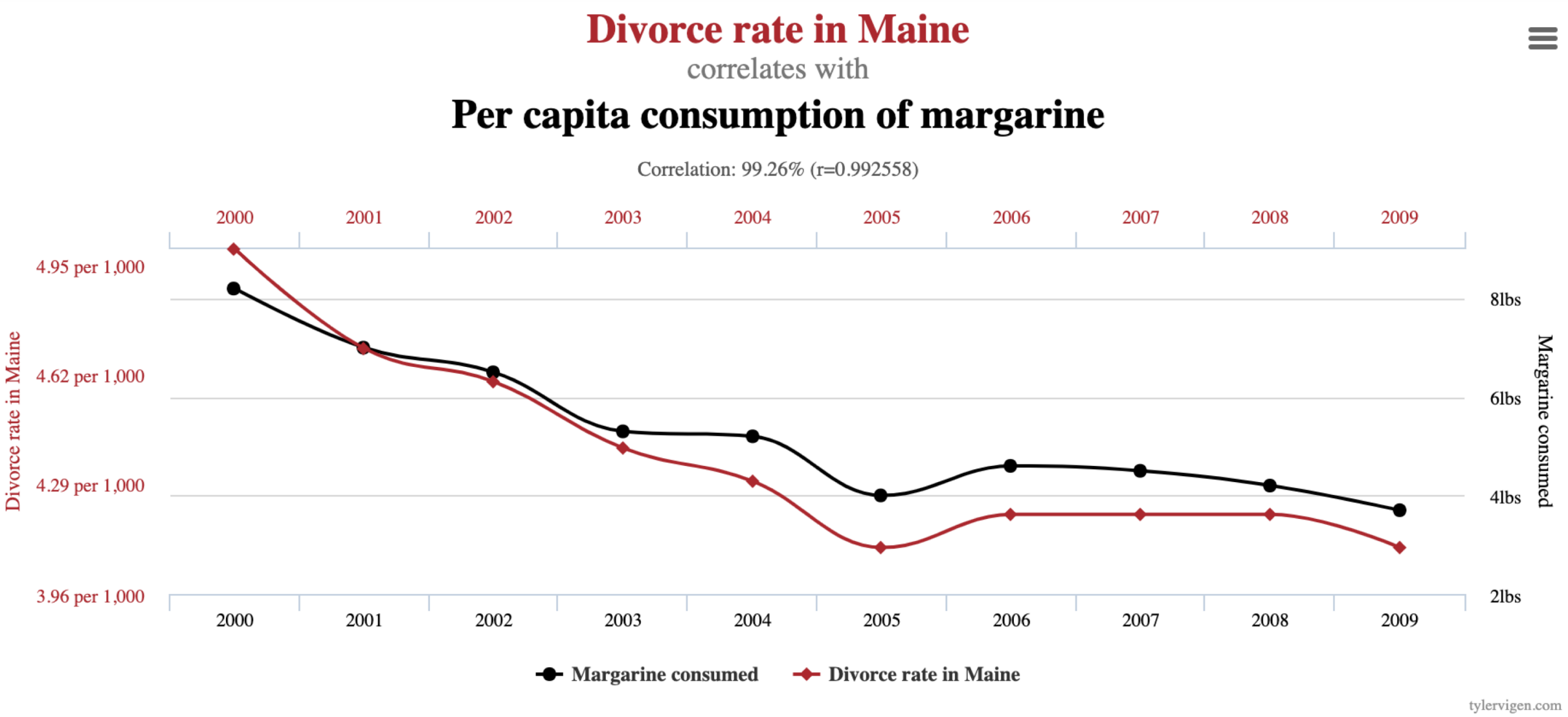
Class Recap


Scoping
Data Acquisition
Data Storage
Data Linkage
Data Exploration
Analytical Formulation
ML Pipelines
Baselines
Feature Generation
Train Test Splits
Evaluation Metrics
Modeling
Model Selection
Interpretability
Bias/Fairness
Field Trials
Deployment
Maintenance & Monitoring

Project checklist

- Train Test Splits – make sure you're using all of them
- Features – make sure you have relevant features
  - Imputation
  - scaling
- Models
  - Make sure you have an appropriately large grid and model types
- Model Selection
  - Initial metric (sanity check with PR-k curve), and selection metric over time
- Interpretability
  - Feature importances, Cross-Tabs
- Bias
  - Protected Group and Bias Metric

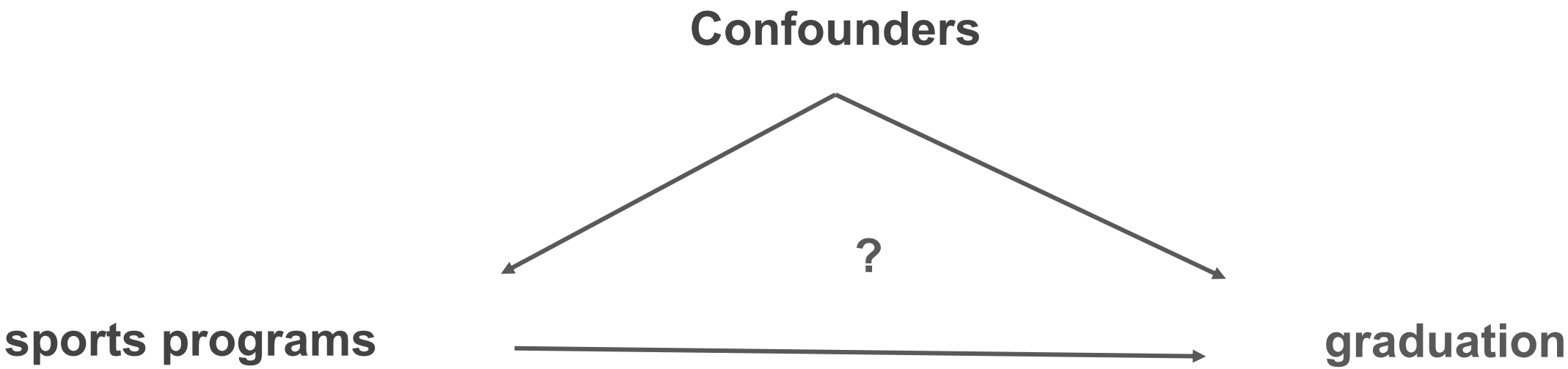
Goals of Causal Inference



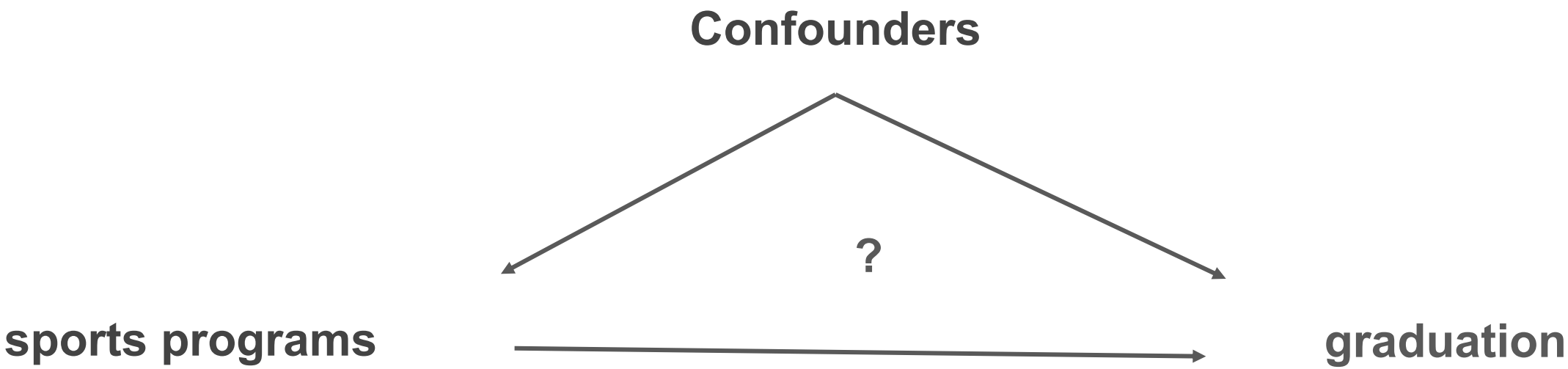
Data sources: National Vital Statistics Reports and U.S. Department of Agriculture

Do sports programs affect school graduation rates?

Do sports programs affect school graduation rates?



Do sports programs affect school graduation rates?

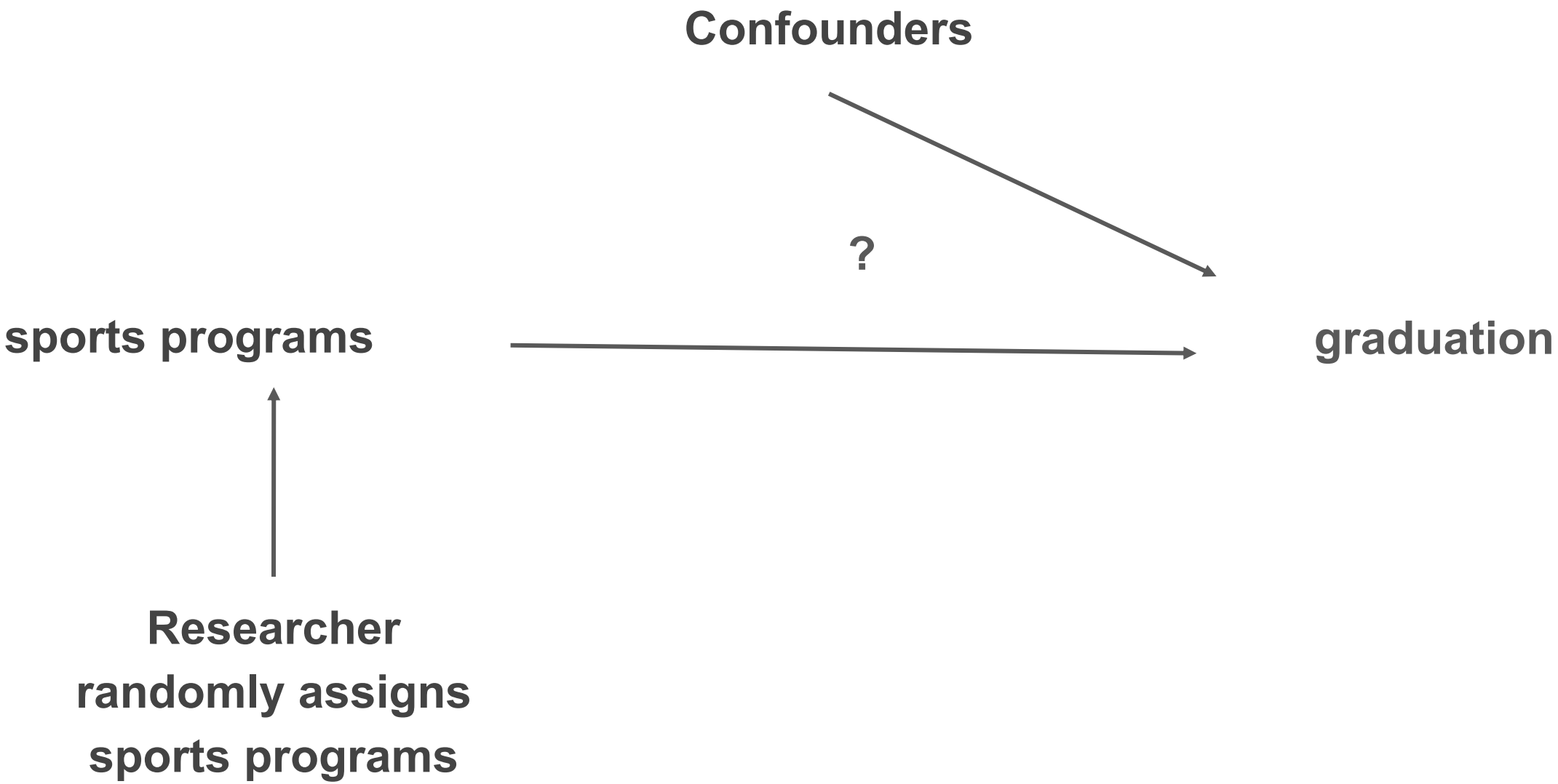


- Examples?

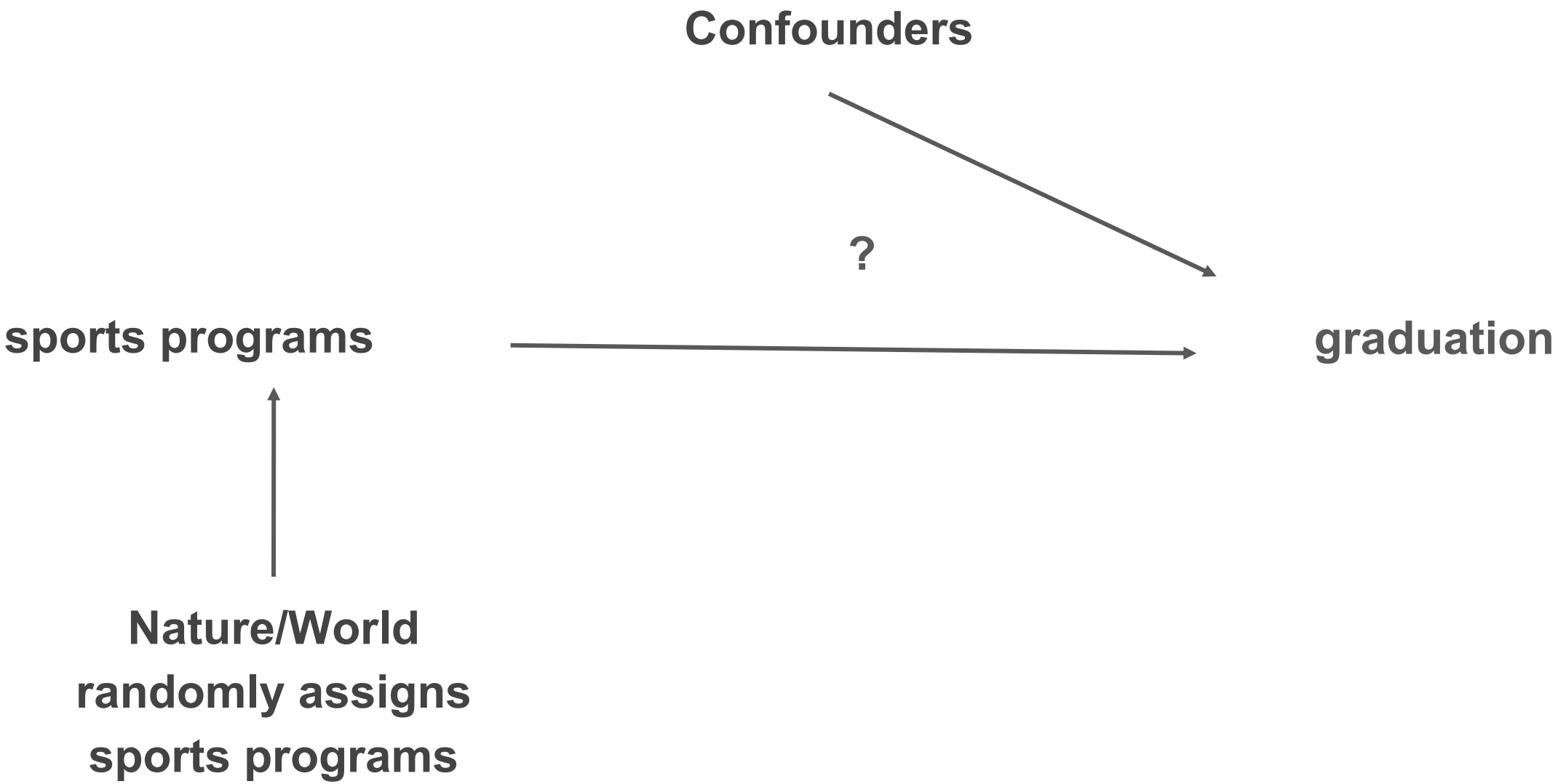
Causal Research Design

1. Experiments
2. Quasi-experiments
3. Observational studies

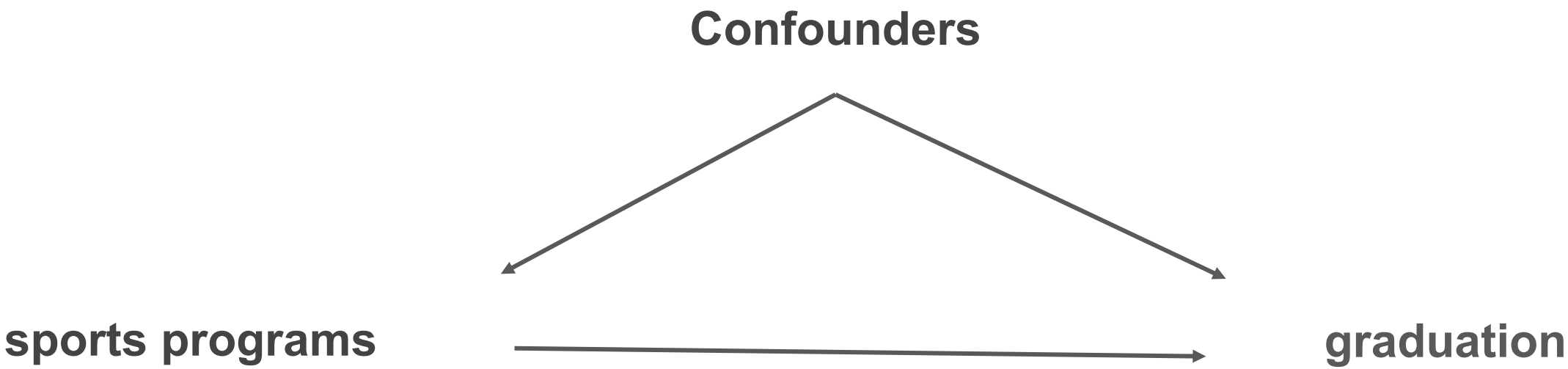
Experiment


Confounders
?
sports programs
graduation
Researcher randomly assigns sports programs

Quasi-Experimental



Observational Study



Include confounders in regression

Methods for causal inference

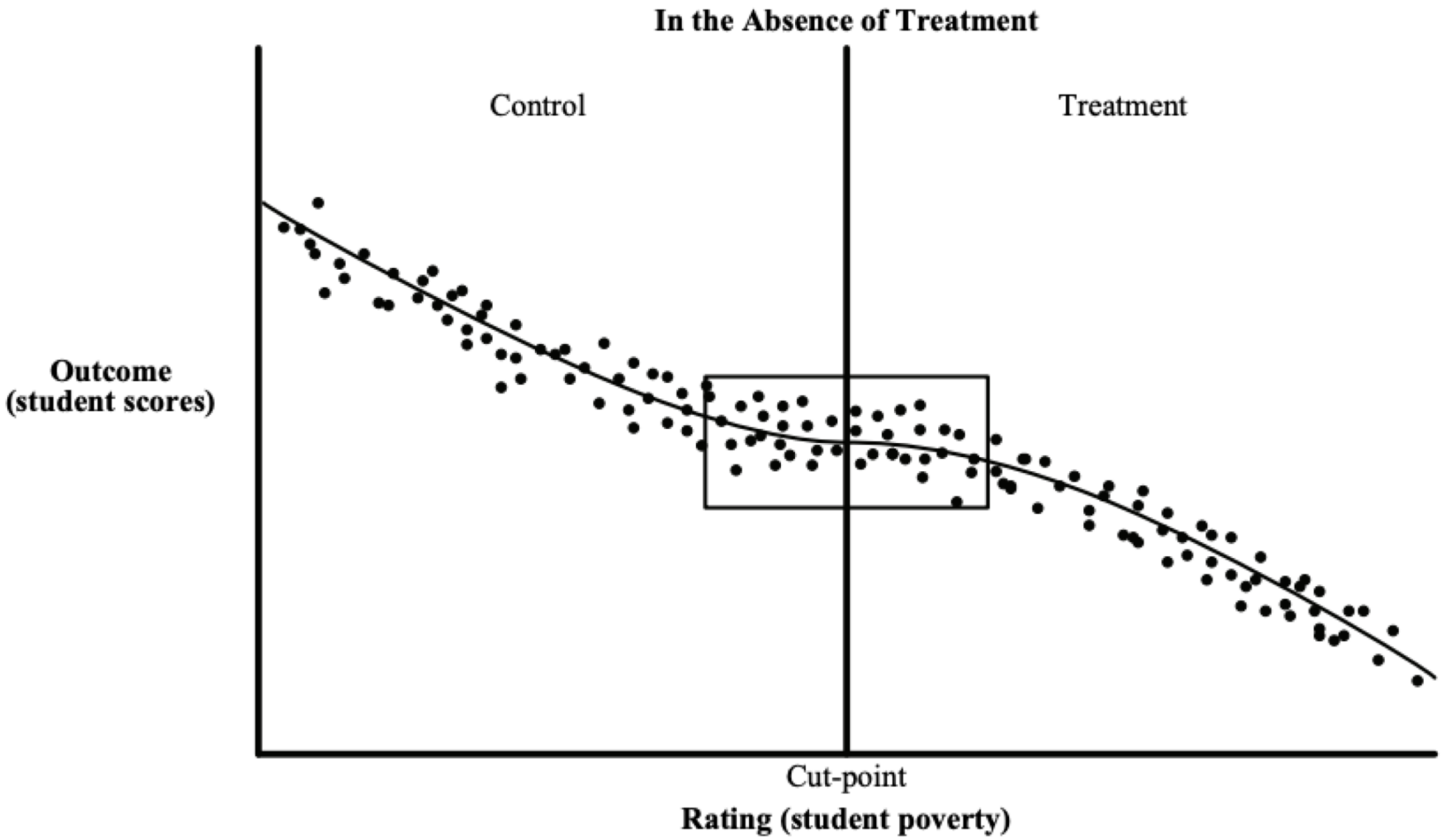
- Using Observational Data
- Using Experiments

Causal inference with observational data

- Matching
- Difference-in-Difference (DID)
- Regression Discontinuity Design (RDD)
- Instrumental Variables

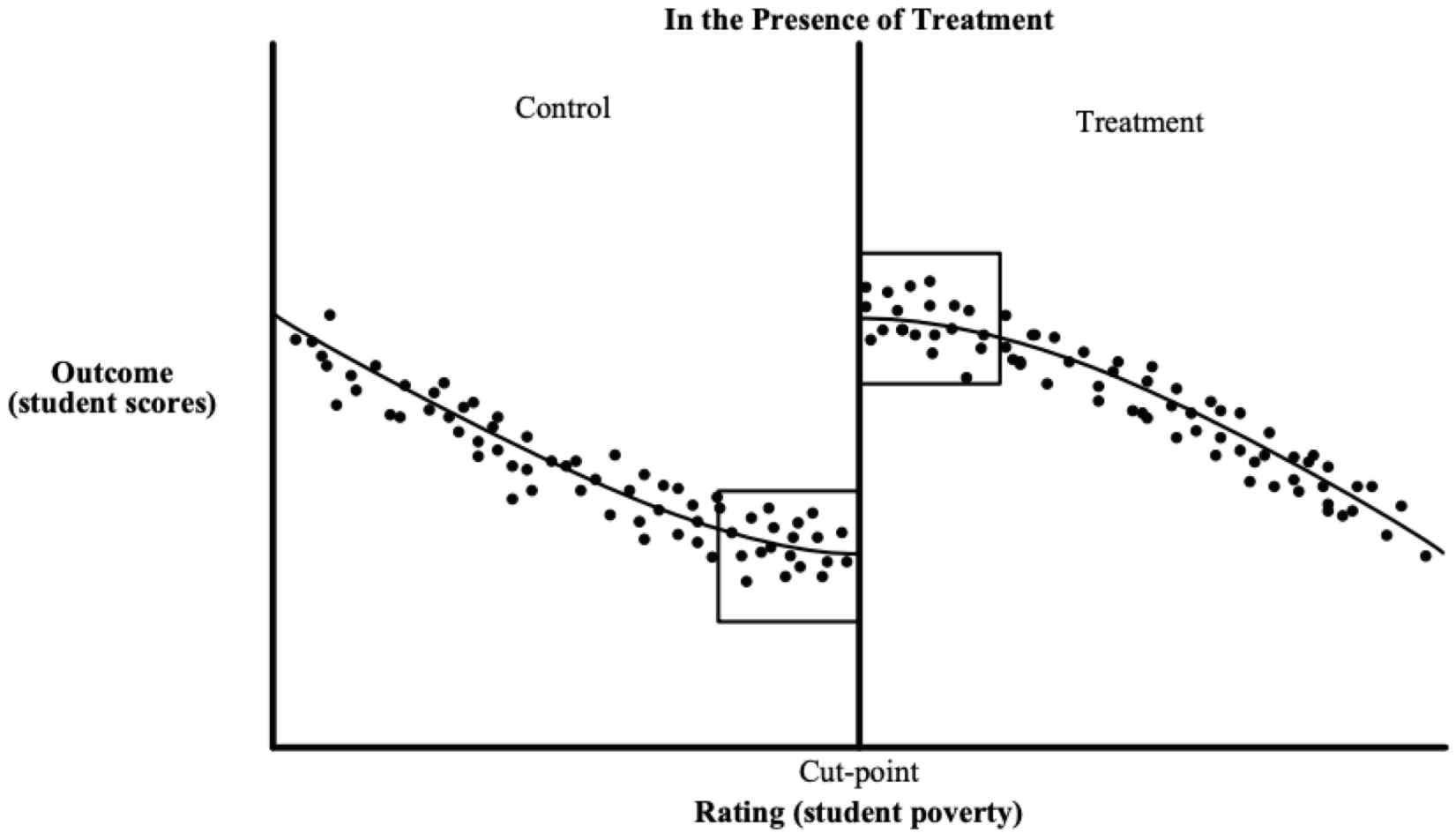
- Limitations?

Causal inference with observational data



https://www.mdrc.org/sites/default/files/RDD%20Guide_Full%20rev%202016_0.pdf

Causal inference with observational data



https://www.mdrc.org/sites/default/files/RDD%20Guide_Full%20rev%202016_0.pdf

Causal inference with experiments

- Standard “Randomized Controlled Trials”
- Trial Design
  - Outcome
  - Sample size
    - Power
  - Randomization

Ethical Concerns

Some things to keep in mind

Sample Size ↑

Power ↑ **Effect Size** ↓ **Significance Level** ↑

Effect size > what is practically useful and significant

Significance Level ~ cost of implementing the new system/model

Designing Trials to Validate ML Systems

- What can we do with historical data?
  - Model Selection
  - Performance Estimation
- How do we select the model to deploy?
  - Compare shortlisted models against baseline
- How do we validate whether our ML system will work when deployed?

Designing an experiment

- What are we trying to test?
- How do we create different arms of the experiment?
- How we do we run the experiment?
- What do we measure?
- How do we evaluate?

What are we testing in an experiment?

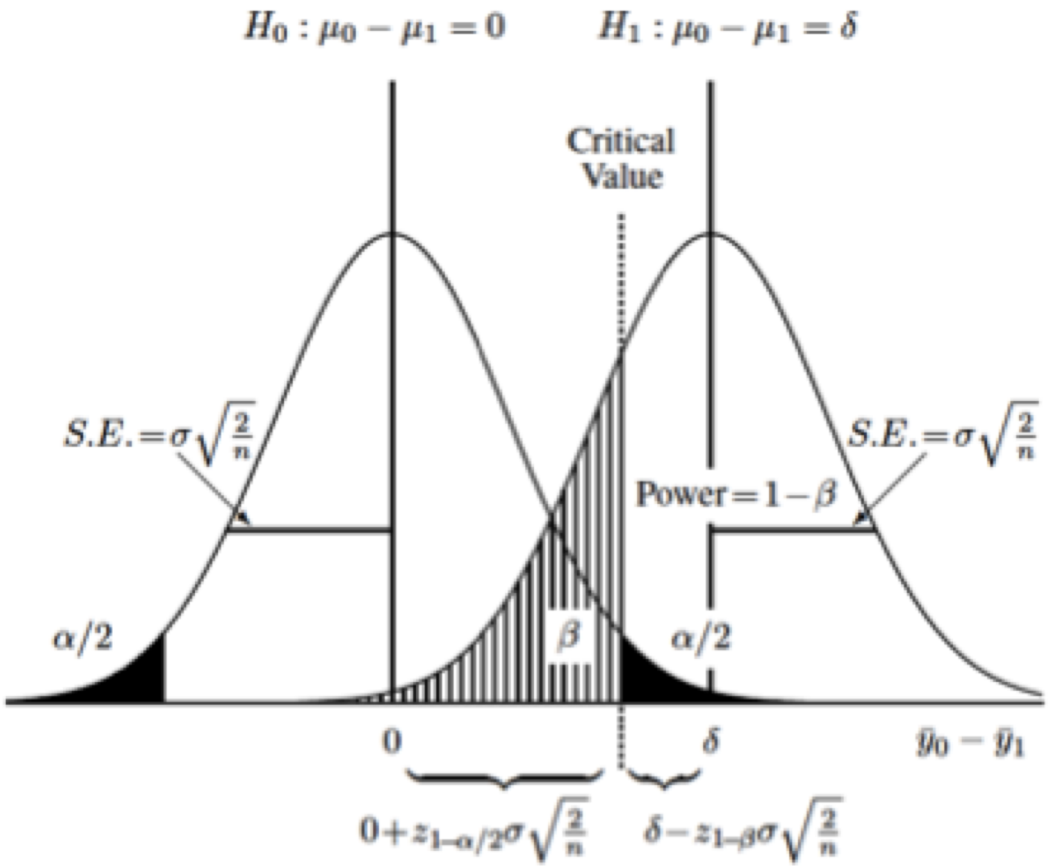
- Lift?
- Classification?
- Probability estimates?
- Ranking?

- How does the experiment design change?

How many [students, inspections, projects] should we include in our trial?

- Doing a trial over [y] years with [n] students every year will give us a [p%] chance to detect an [e%] increase in [outcome] with a statistical significance of [s%]

How many [students, inspections, projects] should we include in our trial?



https://clincalc.com/stats/samplesize.aspx

Some Readings

- References on causal inference
- The seven tools of causal inference, with reflections on machine learning by Pearl, J. Comm ACM. 2019
- Elements of Causal Inference by Peters et al. MIT Press. (Open Access Text)
- Counterfactuals and Causal Inference: Methods and Principles for Social Research by Morgan and Winship. Cambridge University Press.
- Data Analysis Using Regression and Multilevel/Hierarchical Models by Gelman and Hill. Cambridge University Press.