UNIVERSITY of WASHINGTON

Q-learning for Optimal Treatment Rules

Yuwen (Leah) Jia
Mentor: Nina Galanter

Suppose we want to know...

- Does caffeine increase work efficiency?
  - what about for someone who hasn't slept for 3 days? someone who can't consume caffeine?
  - who should / shouldn't drink caffeine?

OPTIMAL TREATMENT RULES

Our interest: apply treatment rule that is best for individuals

Causal inference conditions:

1. potential outcome matches observed outcome
2. no confounding
3. positive probability of being assigned to each of the treatment levels

Q-Learning

X: covariate
A: treatment
β: parameters

1. Outcome Y which we wish to maximize:
   - $Y = \beta_0 + \beta_1 X + (\beta_2 + \beta_3 X)A$
   - terms relate solely to the patient's information
   - interaction between patient and treatment A
2. obtain parameter estimates (β_2, β_3)
3. estimate the optimal treatment decision rule
   - "Treat (A = 1) if $\beta_2 + \beta_3 X > 0$; no treatment (A = 0) otherwise."

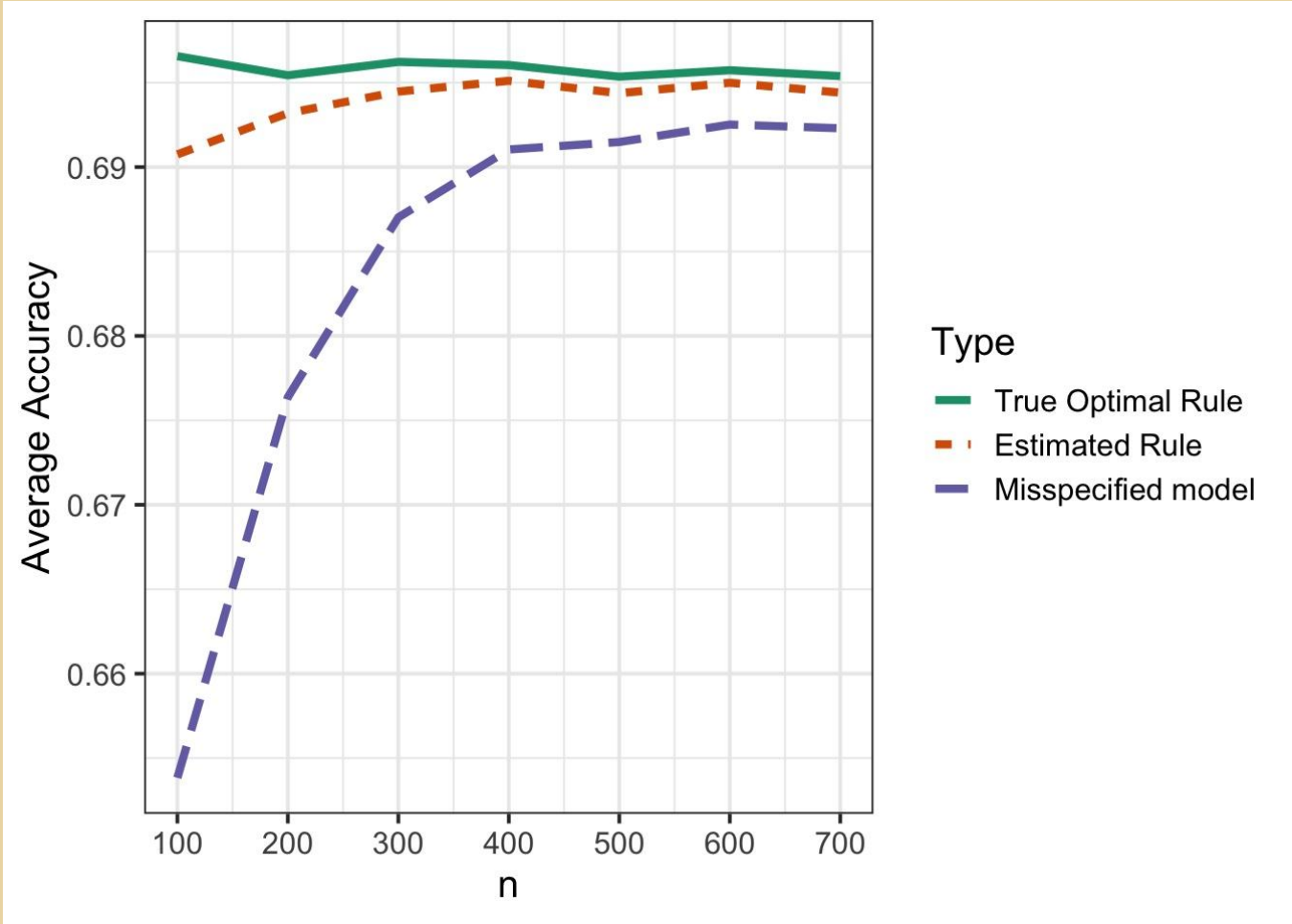
Simulation

- evaluate the performance of our statistical methods
- how do we use them:
  - generating data using random sampling
  - estimate outcome models using simulated data
  - Compare estimated rules to the true optimal rule

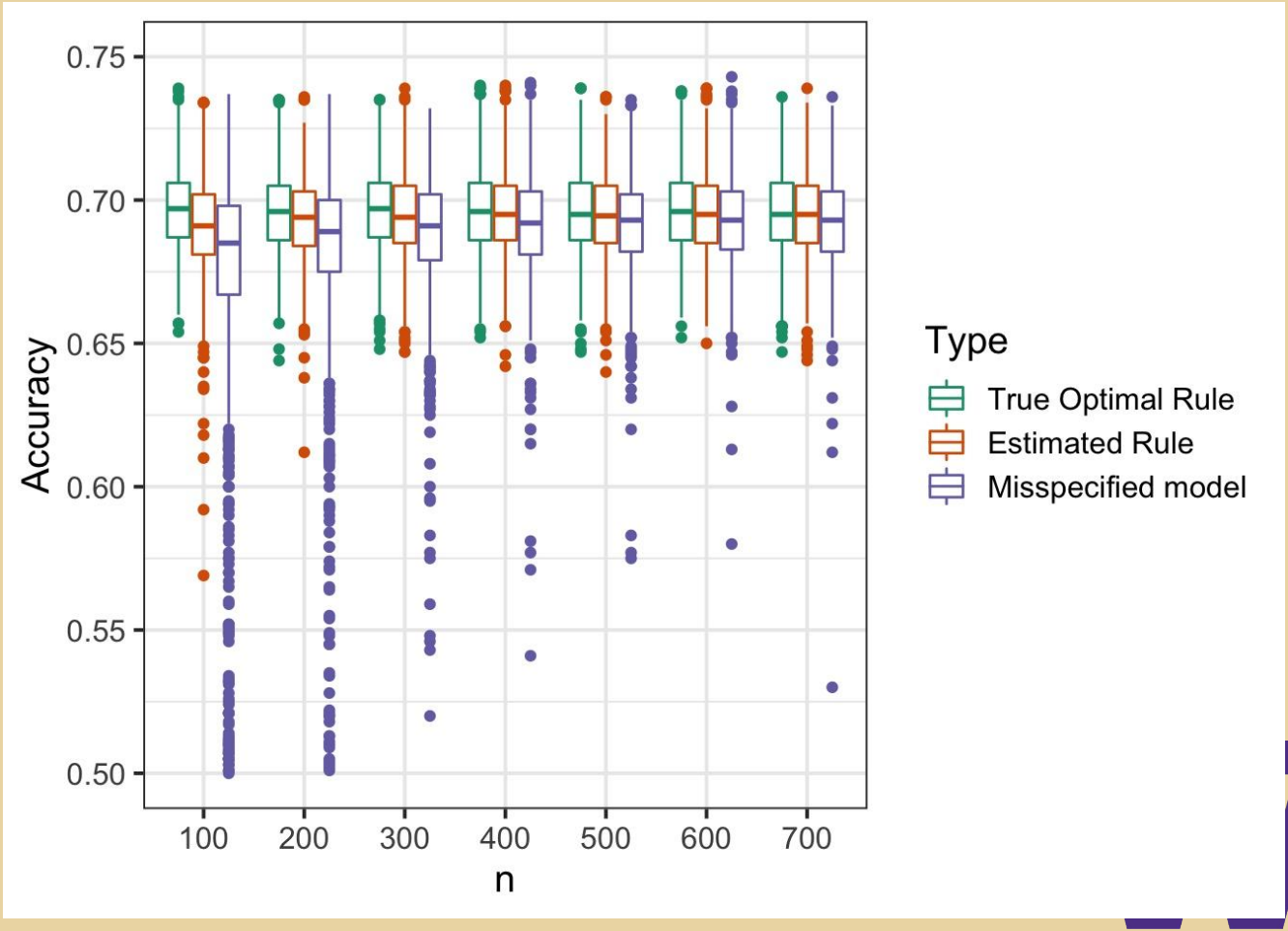
Simulation Setup (one of many ways)

- assign n random X & random treatments (A=0, A=1)
- $E[Y|X,A] = 0 + 1*X + (0 + 1*X)A + 1*X^2$
- consider 3 models to compare:
  - the estimated rule
    - $E[Y|X,A] = \beta_0 + \beta_1 X + (\beta_2 + \beta_3 X)A + \beta_4 X^2$
  - a rule based upon a mis-specified model
    - $E[Y|X,A] = \beta_0 + \beta_1 X + (\beta_2 + \beta_3 X)A$
  - the true best rule (treat when $X > 0$)

Result


Average Accuracy
0.69
0.68
0.67
0.66
100
200
300
400
500
600
700
n
Type
True Optimal Rule
Estimated Rule
Misspecified model

Result 2


0.75
0.70
0.65
0.60
0.55
0.50
Accuracy
100
200
300
400
500
600
700
n
Type
True Optimal Rule
Estimated Rule
Misspecified model

Summary

- Q-learning method:
  - Create a treatment rule without knowing the truth
  - Maximize the average outcome across all participants
- disadvantage of q-learning:
  - vulnerable to model misspecification

sources: Hernán MA, Robins JM (2020). *Causal Inference: What If*.
Michael P. Wallace and Erica E. M. Moodie. *Adaptive Treatment Strategies in Practice*