Linking Probabilistic Numerics to Causal Inference: Theoretical and Methodological Connections

Harrison Zhu

Probabilistic Numerics

- Model the conditional average treatment effect (CATE):

$$\tau(x) = \mathbb{E}[Y^{(1)} - Y^{(0)}|x]$$

where $Y^{(Z)}$ is the response for a treatment group Z

- Use a nonparametric framework:

where $Y = f(X, Z) + \epsilon$

$$Y = Y^{(1)}Z + (1 - Z)Y^{(0)}$$

- This is fundamentally a missing data problem and **selection bias** often occurs when setting up experiments

Probabilistic Numerics

- We want to integration some function f over some measure Π:

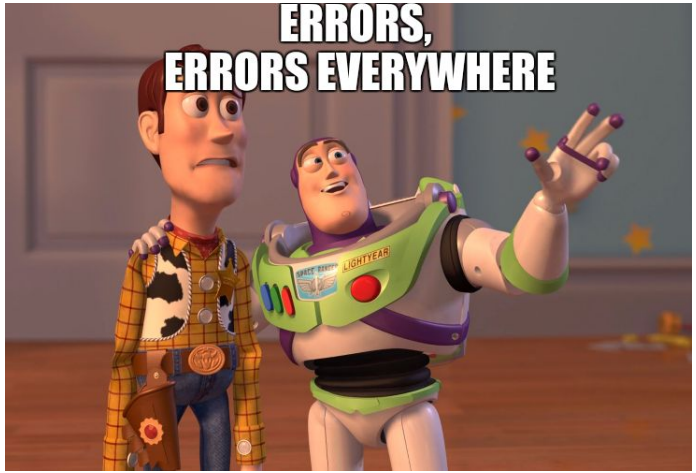
$$\Pi[f] = \int f d\Pi$$

- Given some dataset *(X, y)*, some stochastic process *g(x, w)* as a prior for f, find the posterior distribution

$$\Pi[g]|X, y$$

With point estimates $\hat{\Pi}_{BPNI}[f] = \Pi[\mathbb{E}[g|X, y]]$

- **Goals:** Study the error rates for different classes of functions, adaptive algorithms (sampling new points), uncertainty quantification etc...



Causal Inference Probabili

- Average Treatment Effect (ATE), the average effect of the treatment on a population with distribution Π:

$$\Pi[\tau] = \int \tau d\Pi$$

- Conducts different stages of clinical trials: within the control group $\rightarrow$ larger set of individuals $\rightarrow$ county $\rightarrow$ countrywide etc…

- Whether the new trials improve or consolidate the previous results?

Probabilistic Numerics

- Perform numerical integration on the function τ

- Adaptive/active learning to improve the estimate of both τ and $\Pi[\tau]$

- Theoretical error or contraction rates and the subsequent experimental error rates

Potential Applications and Work

- Integrate the theory of probabilistic numerics into the study of average treatment effects
- Borrow methods from probabilistic numerics for causal inference, and vice versa!
- Use cases:
  - Understanding vaccine effectiveness on different groups of individuals
  - Understanding effectiveness of interventions to reduce overall emissions

Thank you!

Any questions?