

Conclusion

- Three (tensor-valued) variables ought to be enough for everyone: stuff we know, stuff we do, stuff that happens.
- If we are careful to experiment with what we do, we only need very straightforward “causal inference” to learn everything we need to know.
- DAGs are better tools for humans to understand things than for helping us determine what to do from among a set of options.

What I haven't talked about

How do we estimate $E[Y | S, \text{do}(A)]$?

- Model checking procedures
- Heterogeneous treatment effect estimators
- Doubly robust estimators
- Neural networks for causal effects

How do we $\text{argmax}_A E[Y | S, \text{do}(A)]$?

- Optimization
- Differentiable programming
- Defining and aligning on goal metrics
- Estimating tradeoffs
- Eliciting preferences from users

How do we choose $P(A | S)$?

- Randomized experiments
- Off-Policy Evaluation
- Adaptive experimentation
- Bandit algorithms
- Bayesian Optimization

How do we estimate $\text{Var}[\hat{E}[\dots]]$?

- Gaussian process models
- Bayesian inference
- Conformal inference
- Deep ensembles
- Bayesian neural networks