Methods for Causal Inference

February 26, 2024

There are many methods out there for causal inference, and we'd like every group to try using at least two:

- propensity score matching;
- double machine learning.

We'd also like you to investigate (at least) one other method, which should be unique to your group.

Possibilities include:

- augmented inverse probability weighting (Robins and Rotnitzky, 1992);
  this is a commonly used approach, and there is an R package available (Zhong et al., 2021).
- Bayesian Additive Regression Trees (BART) (Hill, 2011);
  the R package for this method seems quite unstable: it might be better to try using Python.
- causal forests (Wager and Athey, 2018);
  an adaptation of random forests to causal problems.
- T-learners (or X-learners!) (Künzel et al., 2019);
  methods which learn separate models for the treated and untreated individuals ('two-learners').
- Targeted Maximum Likelihood Estimation Schuler and Rose (2017);
  a conceptually complicated method with some decent R packages (Gruber and Van Der Laan, 2012).
- some other (non-vanilla!) approach.

References

S. Gruber and M. Van Der Laan. tmle: an R package for targeted maximum likelihood estimation. *Journal of Statistical Software*, 51:1–35, 2012.

J. L. Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

S. R. Künzel, J. S. Sekhon, P. J. Bickel, and B. Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116 (10):4156–4165, 2019.

J. M. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS epidemiology: methodological issues*, pages 297–331. Springer, 1992.

M. S. Schuler and S. Rose. Targeted maximum likelihood estimation for causal inference in observational studies. *American Journal of Epidemiology*, 185(1):65–73, 2017.

S. Wager and S. Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Y. Zhong, E. H. Kennedy, L. M. Bodnar, and A. I. Naimi. AIPW: an R package for augmented inverse probability–weighted estimation of average causal effects. *American Journal of Epidemiology*, 190(12):2690–2699, 2021.