

Selective Inference in Propensity Score Analysis

Yoshiyuki Ninomiya, Yuta Umezu, Ichiro Takeuchi

**Department of Mathematical Analysis and Statistical Inference, The Institute of
Statistical Mathematics**

Selective inference (post-selection inference) is a methodology that has attracted much attention in recent years in the fields of statistics and machine learning. Naive inference based on data that are also used for model selection tends to show an overestimation, and so the selective inference conditions the event that the model was selected. In this presentation, we develop selective inference in propensity score analysis with a semiparametric approach, which has become a standard tool in causal inference. Specifically, for the most basic causal inference model in which the causal effect can be written as a linear sum of confounding variables, we conduct Lasso-type variable selection by adding an L1 penalty term to the loss function that gives a semiparametric estimator. Confidence intervals are then given for the coefficients of the selected confounding variables, conditional on the event of variable selection, with asymptotic guarantees. An important property of this method is that it does not require modeling of nonparametric regression functions for the outcome variables, as is usually the case with semiparametric propensity score analysis.