

Robust Semiparametric Causal Inference

Gaspard Bernard

Institute of Statistical Science, Academia Sinica, Taiwan

Abstract

The classical causal inference problem in which one aims to estimate an average treatment effect, while allowing the outcome of the treatment to be not necessarily independent of its assignment, is considered. This problem is closely linked to the semiparametric inference problem of estimating the expectation of Y , a Bernoulli random variable, from i.i.d. observations (R, RY, Z') . Here, R is a masking random variable following a Bernoulli distribution representing treatment assignation, and it is assumed that R and Y are independent conditionally on some vector of covariates Z . Various root- n consistent estimators have been proposed in the literature. However, we show that in certain scenarios where the model is (even slightly) misspecified, they tend to perform poorly. In particular, when the assumption that R and Y are independent is not exactly satisfied, these estimators tend to exhibit very poor performance.

A new robust estimator is therefore proposed. We study its nonasymptotic behavior under contamination and general misspecification, showing that both the estimation error and the approximation error can be controlled. We next turn to the question of the rate-optimality of the proposed estimator when the model is correctly specified. In some scenarios where the underlying model is correctly specified but highly irregular, we show that the minimax rate for estimating the functional of interest can be attained. When the underlying model is correctly specified and regular enough, we show that the optimal parametric root- n rate can be attained. Finally, still under correct model specification, we study in what sense a lack of regularity in the conditional expectation of Y can be compensated for by stronger regularity assumptions on R , and vice versa.