Confidence interval with possibly invalid instruments even after controlling for many confounders

Hyunseung Kang[*]

hskang@stanford.edu

The instrumental variable (IV) method is a popular method to estimate causal effects of a treatment on an outcome by using variables known as instruments to extract unconfounded variation in treatment, with recent application of the IV method in Mendelian randomization. These instruments must satisfy the core assumptions, including (A1) association with the treatment, (A2) no direct effect on the outcome, and (A3) ignorability. These assumptions may be made more plausible by conditioning on many, possibly high-dimensional covariates, an opportunity increasingly made available by the compilation of large data sets. However, even after conditioning on a large number of covariates, it's possible that the putative instruments are invalid due to their direct effects on the outcome (violating (A2)) or violations of ignorability (violating (A3)) and the resulting inference on the causal effect may be misleading.

We propose a general inference procedure that provides honest inference in the presence of invalid IVs even after controlling for many, possibly high dimensional covariates. We demonstrate our procedure on real and simulated data and find that it outperforms traditional methods, especially when the instruments are invalid.

[*]Department of Economics, Stanford University, John A. and Cynthia Fry Gunn Building, 366 Galvez Street, Stanford, CA 94305-6015, USA