Randomization Inference with An Instrumental Variable: Two Examples and Some Theory

Paul R. Rosenbaum, Department of Statistics, Wharton School
University of Pennsylvania, Philadelphia, PA 19104-6340 US

Talk based on several papers, including one with Guido Imbens and another with Robert Greevy, Jeffrey Silber, and Avital Cnaan.

Abstract

An instrument partially manipulates a treatment, but the instrument affects the response only indirectly through its manipulation of the treatment. Two examples of randomization inference with an instrumental variable will be discussed: a randomized trial with a strong instrument, and a widely discussed study from labor economics with a clever but extremely weak instrument. The randomized, placebo controlled trial concerns an attempt to use a drug, enalapril, to preserve cardiac function of children with cancer who had been treated with anthracyclines. As is often true in clinical trials, compliance with the protocol was imperfect: some children consumed less than the prescribed dose of drug. In this case, the randomization is the instrument, manipulating but not fully controlling the dose of drug received. The example from labor economics is due to Angrist and Krueger, who wished to study the economic returns to additional years of education. In the US, laws concerning compulsory education require children to remain in school until a particular birthday, but the school year begins in September for all children, so these laws require different students to attend school for slightly different periods of time based on their date of birth. In their study, Angrist and Krueger used quarter-of-birth as an instrument for the treatment, years of schooling. The instrument is, of course, extremely weak, because it adds at most a fraction of a year to schooling for a fairly small fraction of all students.

In the randomized trial, randomization inference with an instrumental variable permits randomization to form the ‘reasoned basis for inference,’ in Fisher’s phrase, entirely avoiding biases from self-selection in the decision to comply with the study protocol. Moreover, this analysis agrees with the so-called ‘intent-to-treat’ analysis, in the sense that both analyses report exactly the same significance level for testing the null hypothesis of no treatment effect.

A large literature shows that conventional methods give incorrect conclusions with weak instruments; for instance, 95% confidence intervals do not cover 95% of the time. In contrast, randomization inference does not have this problem, even in situations that are not identified: its 95% confidence intervals do cover 95% of the time, with intervals whose (possibly infinite) length appropriately reflect the strength of the instrument. Moreover, theory shows this is the only way to achieve such coverage in the absence of distributional assumptions. This is illustrated using the study from labor economics and simulations.