An Automated Approach to Causal Inference in Discrete Settings

ABSTRACT:

When causal quantities cannot be point identified, researchers often pursue partial identification to quantify the range of possible values. However, the peculiarities of applied research conditions can make this analytically intractable. We present a general and automated approach to causal inference in discrete settings. We show causal questions with discrete data reduce to polynomial programming problems, and we present an algorithm to automatically bound causal effects using efficient dual relaxation and spatial branch-and-bound techniques. The user declares an estimand, states assumptions, and provides data (however incomplete or mismeasured). The algorithm then searches over admissible data-generating processes and outputs the most precise possible range consistent with available information -- i.e., sharp bounds -- including a point-identified solution if one exists. Because this search can be computationally intensive, our procedure reports and continually refines non-sharp ranges that are guaranteed to contain the truth at all times, even when the algorithm is not run to completion. Moreover, it offers an additional guarantee we refer to as $\epsilon$-sharpness, characterizing the worst-case looseness of the incomplete bounds. Analytically validated simulations show the algorithm accommodates classic obstacles, including confounding, selection, measurement error, noncompliance, and nonresponse.