Combining experiments and observations in causal inference

Juha Karvanen

Jyväskylä University, Finland

Abstract

The knowledge on a real-world problem often consists of fragments of evidence that may come from experiments, observations and expert opinions, all including a varying degree of uncertainty and possibly missing data and selection bias. In causal inference, these fragments are combined to draw conclusions on the effects of causes. Do-calculus is a general-purpose tool for solving identification problems, i.e., deciding if the causal effect of interest can be expressed using available observational and experimental distributions. The rules of do-calculus do not themselves indicate the order in which they should be applied and therefore identification algorithms such as the ID algorithm by Shpitser and Pearl have been derived. Despite the rapid progress of the field, the current algorithms are not able to solve the general data fusion problem where the task is to identify the causal effect of interest combining arbitrary data sources.

In this presentation, the main algorithms for the identification of causal effects are reviewed in a non-technical level and the problems outside of their scope are discussed. Besides the ID algorithm, there are algorithms for transportability and selection bias recovery. These algorithms are implemented in the R package causaleffect created at University of Jyväskylä. The challenges for the future are related to extending these algorithms to more complicated setups and applying them in real-world problems. Often many alternative expressions exist for an identifiable causal effect but the properties of estimators arising from different expressions are mainly unknown.