Summary 23: Causal Inference

Susan Cherry

This paper reviews casual inference based on the Structural Causal Model (SCM). First, the authors compare causal analysis and statistical analysis. Unlike standard statistical analysis, causal inference is concerned with the changing of conditions, such as changes as a result of treatments or external interventions. It is important to remember that behind every causal conclusion, there are causal assumptions that are not testable in observational studies. Causal relationships can't be defined just from the distribution. Causal concepts include randomization, influence, confounding, instrumental variables, etc. Associational assumptions are testable in principle but causal assumptions cannot be tested without an experimental control.

Next, the paper discusses structural causal models and path diagrams, in which arrows are drawn from causes to their effects. The absence of an arrow implies that the value assigned to one variable is not determined by the value of another variable. Background factors that influence but are not influenced are called exogenous variables. Variables that are influenced are called endogenous. Next, the authors talk about confounding and causal effect estimation. First, the introduce covariate selection and the back-door criterion. Statisticians want to select a subset of factors for measurement and adjustment so that if we compare treated and untreated individuals with the same values of selected features, we can find a treatment effect in that subpopulation. The following criterion allow us us to define a sufficient set: 1) No element of S is a descendent of X. 2) The elements of S block all backdoor paths from X to Y. Then the authors move to counterfactual analysis in structural models.

Next, Potential Outcomes and Counterfactuals are discussed. This means we are interested in the value of the outcome variable if a subject had received a treatment x instead of what it actually received. Potential outcome approach is carried out by treating the observed distribution as the marginal distribution of a probability distribution of the counterfactual variable. The structural model views the intervention as an operation that changes the distribution but keeps the variable the same while the potential outcome approach views the outcome variable under the intervention to be a different variable. The main drawback of the black-box approach is that researchers must begin to articulate the science behind the problem their studying.

Finally, Direct and Indirect effects are discussed. The causal effect studied so far is the total effect of a variable X on the outcome variable. Sometimes, however, the direct effect does not represent the target of the investigation well enough and attention is directed at the direct effect of X on Y. Direct effect is the effect that is not mediated by other variables in the model (effect of one variable by holding others fixed). Indirect affects are the expected change in Y holding X constant and changing Z to what it would have been if X had ben set to X'. For example, a policy maker might be interested in the effect of a policy to a set of employees or in controlling the delivery of messages in a network of agents.

In summary, causal inference requires a scientific language for explaining causal knowledge and a mathematical model for processing the knowledge. The authors introduce randomization, intervention, direct/indirect effects, and counterfactuals.