Department of Decision Sciences

Statistics Seminar

Causal inference under case-control and other outcome dependent sampling

Vanessa Didelez
Leibniz Institute for Prevention Research and Epidemiology – BIPS GmbH

Thursday, 27th October 2016
12:30pm Room 3-E4-SR03 Via Rontgen 1 Milano

Abstract

In this presentation I will review when and how it is possible to draw causal conclusions from case–control designs; some of the results are valid more generally for other situations where by design or accident the sampling depends on the outcome. The main focus is on the question of identifiability: does the available data, at least in principle (for 'very large' samples sizes), allow us to consistently estimate the desired causal quantity? If the answer is 'no' then this is typically due to structural bias, i.e. to fundamental problems of design and available information, which is obviously strongly influenced by the sampling design. In case–control studies, we face the following potential sources of structural bias regarding causal inference: (1) Case–control studies are necessarily observational, so confounding is likely to be present. (2) Case–control studies are retrospective with sampling being conditional on disease status which means there is also a threat of selection bias. (3) A consequence of the retrospective sampling is that methods which depend on, or are sensitive to, the marginal distribution of the outcome cannot be used without some modification, since the required information is not generally available. This is potentially relevant to certain methods of adjusting for confounding as well as to the identifiability of typical causal effect measures, such as the average causal effect.

While confounding is a problem of any observational study and has been widely addressed in the causal inference literature, points (2) and (3) are more specific to case–control studies and will be the focus here. I will specifically consider: identification of the null-hypothesis of no causal effect which is closely related to the non-parametric identification of causal odds ratios; further I will address how certain structural knowledge can enable us to reconstruct the full joint distribution; finally I will briefly compare these approaches with those that rely on additional knowledge of the population prevalence, such as standardisation, propensity scores, and instrumental variables.

Department of Decision Sciences

Via Röntgen 1 - 20136Milano

Tel. 02 5836.5632
Fax 02 5836.5630