Methods for the Analysis of Exposure Effect on Secondary Outcomes in Case-Control Studies

Tamar Sofer, University of Washington

Case-control studies are designed towards studying associations between exposures and a single, primary disease outcome. Typically, information about secondary outcomes is also collected, but association studies targeting secondary outcomes should account for the case-control sampling scheme. The Inverse Probability Weighted (IPW) estimator is often used in such studies to prevent potential selection bias when estimating population effects. In this talk, I will present methods for analysis of secondary outcomes, which extend the IPW estimator. First, I suggest a "control function" assisted IPW estimator. In this approach we add a so-called control function to the usual IPW. This function incorporates information about the population disease model and the selection bias. The resulting estimator is more efficient than the usual IPW if the disease model and the selection bias function are correctly specified, and robust to misspecification of the selection bias function. Second, I present an IPW pseudo-likelihood approach for the analysis of the effect of a set of genetic variants on multiple correlated secondary outcomes. Using the pseudo-likelihood we develop a variance component test, for the set effect. Upon rejection of the null hypothesis of no effect, we propose to identify and estimate non-zero effects of genetic variants using oracle-penalized weighted pseudo-likelihood. The proposed estimators are evaluated in simulation studies and demonstrated on a case-control genome-wide association study of type 2 diabetes.