Semiparametric Bayesian Analysis of Studies of Gene-Environment Interaction

Bhramar Mukherjee

University of Michigan

E-Mail: bhramar@umich.edu

Abstract: We will discuss the problem of retrospective modeling of case-control data for studying gene-environment interactions in a semiparametric Bayesian framework. The special feature of gene-environment interaction studies is that in many situations it is scientifically plausible to assume that the genetic and environmental factors are independent in the underlying population. Under this additional constraint of gene-environment independence, one can derive more efficient estimation techniques than the traditional prospective logistic regression analysis (Piegorsch et al, 1994; Chatterjee and Carroll, 2005). However, the efficient estimates from the retrospective likelihood may be severely biased under the violation of the independence assumption. Stratification effects present in the population could potentially induce non-independence among genetic factors and environmental exposures. We will first provide a semiparametric Bayesian approach to model stratification effects under the assumption of gene-environment independence. We will then propose an alternative to relax the constraint of gene-environment independence in a natural Bayesian framework to strike a compromise between efficiency and robustness. We will analyze data from a case-control study on ovarian cancer patients conducted in Israel to illustrate our methods. Collaborators on material related to the presented work are: Nilanjan Chatterjee (National Cancer Institute), Malay Ghosh (University of Florida), Li Zhang (Cleveland Clinic Foundation) and Samiran Sinha (Texas A & M).