The University of Chicago

Department of Statistics

Seminar Series

LEI SUN
Departments of Public Health Sciences and Statistics
University of Toronto

"Non-discovery rate and stratified false discovery control with application to genome-wide association studies"

MONDAY, April 10, 2006 at 4:00 PM
133 Eckhart Hall, 5734 S. University Avenue
Refreshments following the seminar in Eckhart 110.

ABSTRACT

Current high-throughput genotyping technology provides efficient means to collect vast amounts of data with low cost. However, it also poses many interesting and challenging statistical and computational problems, in particular the multiplicity problem or the problem of multiple hypothesis testing. With the diminishing power associated with the traditional control of the family-wise error rate (FWER), the use of the false discovery rate (FDR) has become common in many applications, and genome-wide association (GWA) studies of complex human disorders are no exception. Using a recent GWA study of Parkinson disease by Maraganore et al. (2005) as an example, in which 198,345 SNPs were tested for association and 1,906 identified with p-values < 0.01, we first asked two questions: 1) how many are false among the 1,906 significant SNPs? 2) how many did we miss among all the true signals?

The analysis of 1) is essentially FDR estimation. The quantity asked in 2) is extremely important and meaningful, however it has been largely over-looked in the statistical literature. We recently proposed the non-discovery rate (NDR) to quantify the type II error/false negative rate for multiple hypothesis testing. Analyses showed that: 1) FDR $> 95\%$: most positives were false! 2) NDR $> 95\%$: most true signals were missed! These results lead to the third question: 3) can we do better?

We propose a stratified false discovery control approach for situations where there is an inherent stratification of the large number of tests to be performed. The proposed method is a simple way to incorporate available auxiliary information where the auxiliary variable is the stratum indicator. Both analytic results and application demonstrate the potential advantages of control or estimation of FDR by stratum.

This is joint work with Shelley Bull, Radu Craiu and Andrew Paterson of University of Toronto.

Please send email to Mathias Drton (drton@galton.uchicago.edu) for further information. Information about building access for persons with disabilities may be obtained in advance by calling the department office at (773) 702-8333.