Please Post

Ohio State - Cleveland Clinic Foundation – Case Western Reserve University

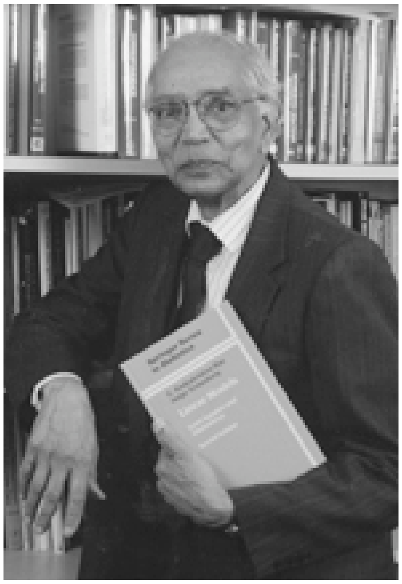
Biostatistics Joint Symposium
Thursday, May 18, 2006

Blackwell Hotel Ballroom,
2110 Tuttle Park Place, Columbus, Ohio

(see the symposium link at www.stat.ohio-state.edu/~seminar for directions and further abstracts)

Schedule

11:30 – 12:00	Robert C. Elston, Case Western Reserve University "Some statistical problems in genome-wide association analyses."
12:15 - 1:45	Buffet Lunch
1:45 - 2:15	Mei-Ling Ting Lee, The Ohio State University, School of Public Health “Generalized Rank Tests for Replicated Microarray Data”
2:15 - 2:45	Michael Kattan, Cleveland Clinic Foundation "Statistical Prediction Models and Medical Decision Making"
2:45 - 3:15	Break
3:15 - 4:15	C.R. Rao, Penn State University **“Bagging and Boosting – Applications to Classification and Regression Methodologies”**



Keynote Speaker: Calyampudi R. Rao
Eberly Professor Emeritus
Penn State University

Abstract

In this talk, the basic principles of bagging and boosting will be laid out in considerable detail. Some theoretical and empirical evidence will be presented that these methods do work well. Examples from the field of clinical data and microarrays will be used.

Tree structured methods like CART are attractive in that they produce flexible models of simple structure. They are, however, unstable in that small perturbations of training data can give different trees. The idea of aggregating trees over many data sets by bootstrap resampling (bagging and boosting) can lead to improved predictions. Some improvements of these methods will also be presented.

The 2006 Biostatistics Joint symposium is funded by the Statistics Department, the School of Public Health, and the Biostatistics Center of The Ohio State University

May 18, 2006 OSU-CCF-CWRU Biostatistics Joint Symposium

Further Abstracts

Some statistical problems in genome-wide association analyses

Robert C. Elston
Case Western Reserve University

New bead and chip technology for single nucleotide polymorphisms (SNPs), and the completion of the first phase of the HapMap project, have motivated the start (perhaps premature) of genome-wide association studies capitalizing on the presence of linkage disequilibrium. This has led to the development of multiple new methods of analysis, as well as novel designs, for using association analysis to find disease genes. Of particular concern has been the possibility of association studies being confounded by population stratification and the multiplicity of tests now possible. I shall give an overview of the statistical, genetic and computational issues involved.

Generalized Rank Tests for Replicated Microarray Data

Mei-Ling Ting Lee
Ohio State University, School of Public Health

In order to filter out unwanted noise and random error, there has been an increasing practice of replication in microarray studies. As a consequence, an important question arises on how to make inferences based on these replicated intensities measurements.
Generalized rank-sum tests or generalized signed-rank tests were proposed for analyzing replicated microarray data.

Statistical Prediction Models and Medical Decision Making

Michael Kattan
Cleveland Clinic Foundation

Quantitative approaches to medical decision making typically invoke the decision analytic framework. Central to this framework are probabilities and patient preferences. More accurate predicted probabilities should lead to improved medical decision making. However, predicted probabilities can come from a variety of sources, with varying degrees of accuracy.

The accuracies of various methods for determining predicted probabilities will be discussed and compared in case studies. These methods include data-driven approaches as well as clinician judgment. As a general rule, it seems that traditional regression models, with special attention paid to predictive accuracy, tend to provide the most accurate predictions presently available. Approaches for implementing them in standalone software are illustrated, as are the use of these models in clinical trial design and analysis.
