

BIOSTATISTICS SEMINAR

WINTER 2010

Three topics concerning large rectangular arrays of data

Richard Olshen, PhD

Department of Health Research & Policy (Biostatistics) and
Departments of Electrical Engineering and of Statistics
Stanford University

Wednesday, February 3rd, 2010

3:30pm - 4:30pm

Cookies and Coffee at 3:00 PM in room 51-254 CHS

13-105A CHS

ABSTRACT: I will give a survey of three topics, all regarding analyses of large rectangular sets of data. Concerns are motivated by studies of gene expression and assessing (low) levels of circulating proteins, separately by single nucleotide polymorphisms (SNPs). The first topic is that of successive normalization and applies directly to the first two topics. One mean polishes, then standard deviation polishes, alternatively and successively, rows and columns of an array. Why might we be interested in such normalization? When and to what does this set of transformations converge, and how fast? Are there dimensional constraints? The second topic is a genome-wide association study (GWAS) of mean arterial blood pressure in a group of white women. The third topic, time permitting, is a study of the impact of SNPs and other features on insulin resistance in a group of Chinese women.

The three works cited are joint with many coauthors, to whom are due much credit and no blame. They will be mentioned during my presentation.