Cheng Cheng

Exploratory Failure Time Analyses and Inference of Copy Number Variations

A recent approach to genomic research is Genome Wide Association Study (GWAS) that capitalizes on the high-density single nucleotide polymorphism (SNP) array technology. With this technology, currently one can obtain genotypes of up to about 1 million SNPs in a genome in hundreds even tens of thousands of subjects in a single study. The goal is then to identify the SNPs whose variations can significantly explain the variations in one or more phenotypes of interest. In human cancer genomics, an important phenotype is the risk (time to and probability) of treatment failure defined as cancer relapse. There will be a presentation and discussions of two novel nonparametric and robust methods for GWAS of failure time phenotypes, following a brief review of common failure time analysis methods. The first method is a SNP-by-SNP screening procedure and the other is a dimension reduction method based on a recent
sparse regression technique. Simulation studies and real data examples will be discussed.
A hallmark of cancer genomes is that there often exist gains or losses of relatively large DNA segments on chromosomes, termed as copy number variation (CNV, from the normal diploid state). The second part
of the lecture consists of a presentation of some recent developments in statistical inference of CNVs in somatic (tumor) and normal cells using SNP array signals.