Steve Qin, PhD
Emory University

Friday, November 20, 2015
12:00-1:00 p.m.
UW Biotech Center Auditorium, 425 Henry Mall

Improving Hierarchal models using historical data

Abstract:
Modern high throughput biotechnologies such as microarray and next generation sequencing produce massive amount of information for each sample assayed. However, in a typical high throughput experiment, only very limited amount of data are observed for each individual feature, thus the classical large $p$, small $n$ problem. Bayesian hierarchical model, capable of borrowing strength across features within the same dataset, has been recognized as an effective tool in analyzing such data. However, the shrinkage effect, the most prominent feature of hierarchical features, can lead to undesirable over-correction for some features. In this work, we discuss possible causes of the over-correction problem and propose several alternative solutions. Our strategy is rooted in the facts that in Big Data era, large amount of historical data are available which should be taken advantage of. Our strategies present a new framework to enhance the Bayesian hierarchical models. Through simulation and real data analysis, we demonstrated superior performance of the proposed strategies.