



UNC  
GILLINGS SCHOOL OF  
GLOBAL PUBLIC HEALTH

## **BIOSTATISTICS SEMINAR**



**Rajarshi Mukherjee, PhD**

Stein Fellow

Department of Statistics

Stanford University

### **Sparse Signal Detection with Binary Outcomes**

Massive genetic data presents us with several statistical challenges. One major challenge is to develop effective strategies for signal detection from large scale genetic and genomic data when signals are possibly weak and sparse. In this talk, motivated by next generation sequencing association studies with binary outcomes, we explore a delicate balance between minor allele frequencies, signal sparsity, and signal strength in deciding sharp detection thresholds. In particular, we observe a new phenomenon in the behavior of the sparse signal detection problems which does not occur in the case of continuous (e.g. Gaussian) outcomes. As a crucial ingredient, we will discuss a version of the Higher Criticism Test which is provably sharp up to optimal constants in the regime of sparse signals. The theoretical results will be illustrated using numerical simulations and data examples from genetic association studies. Time permitting, we will also try to address a related question of detecting differential attractiveness of vertices in sparse networks. \*This talk is based on joint work with Xihong Lin and Natesh Pillai.

**Thursday, February 9, 2017**

**3:30 pm - 4:30 pm**

**Blue Cross Blue Shield Auditorium**

**0001 Michael Hooker Research Center**