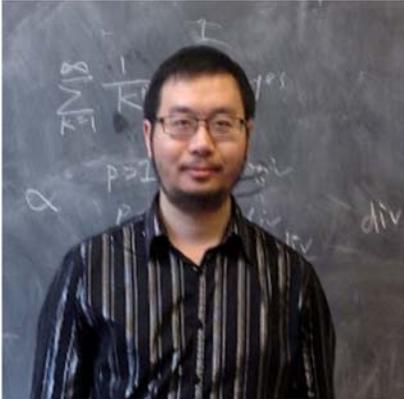


DEPARTMENT OF STATISTICS

**Xin Tong**

Department of Data Science and Operations
University of Southern California

Neyman-Pearson classification

September 25, 2019

11:45am – 12:45pm

Light refreshments will be served

110 Frelinghuysen Road

Hill Center, Room 552

Abstract: In many binary classification applications, such as disease diagnosis and spam detection, practitioners commonly face the need to limit type I error (that is, the conditional probability of misclassifying a class 0 observation as class 1) so that it remains below a desired threshold. To address this need, the Neyman-Pearson (NP) classification paradigm is a natural choice; it minimizes type II error (that is, the conditional probability of misclassifying a class 1 observation as class 0) while enforcing an upper bound, α , on the type I error. Although the NP paradigm has a century-long history in hypothesis testing, it has not been well recognized and implemented in classification schemes. Common practices that directly limit the empirical type I error to no more than α do not satisfy the type I error control objective because the resulting classifiers are still likely to have type I errors much larger than α . This talk introduces the speaker and coauthors' work on NP classification algorithms and their applications and raises current challenges under the NP paradigm.

Bio: Xin Tong is an assistant professor in the Department of Data Sciences and Operations at the University of Southern California. He attended the University of Toronto for undergraduate studies in mathematics and obtained a Ph.D. degree in Operations Research from Princeton University. Before joining the University of Southern California, He was an instructor in Statistics at the Department of Mathematics, Massachusetts Institute of Technology. His current research interest focuses on asymmetric statistical learning problems. His research is partially funded by the United States NSF and NIH. .

