Lecture 11.1 (03:15-03:40)

Large-scale Adaptive Multiple Testing for Sequential Data

Shyamal Krishna De (ISI)

In this talk, we shall discuss multiple testing procedures with simultaneous control of false discovery and nondiscovery rates when m-variate data vectors are observed sequentially or in groups. Existing multiple testing methods for sequential data uses fixed stopping boundaries that do not depend on the sample size, and hence, are quite conservative when the number of hypotheses m is large. We propose sequential tests based on adaptive stopping boundaries that shrink the continue sampling region as the sample size increases. Under a two-group mixture model assumption, we propose an oracle rule based on the local false discovery rate statistic and prove exact control of FDR and FNR at some prefixed levels. Further, we develop a data-driven rule which guarantees simultaneous control of FDR and FNR asymptotically as m tends to infinity. Both the oracle and the data-driven stopping times are shown to be finite with probability one for all finite m, and converge to a finite constant as m grows to infinity. Extensive analysis of simulated datasets illustrate the superiority of the proposed tests over some existing methods.