Mathematics, Statistics, and Computer Science **@ UIC**

Statistics and Data Science Seminar

Finite sample change point inference and identification for high-dimensional location-shift

Mengjia Yu (UIUC)

Abstract: I will discuss two approaches of the cumulative sum (CUSUM) statistics and the U-statistics in change point problems for high-dimensional location-shift. Both works are non-parametric, fully data-dependent and enjoying strong theoretical guarantees under arbitrary dependence structures. 1. Based on the ℓ^∞-norm of the CUSUM statistics, we study inference and identification for high-dimensional mean vectors. For the problem of testing existence of a change point in an independent sample generated from the mean-shift model, we introduce a Gaussian multiplier bootstrap to calibrate critical values of the CUSUM test statistics. For the problem of estimating the change point location once it is detected, two estimators are proposed by maximizing the ℓ^∞-norm of the generalized CUSUM statistics at two different weighting scales. In both problems, dimension impacts the rate of convergence only through the logarithm factors, and therefore consistency of the CUSUM location estimators is possible when p is much larger than n. 2. In cases where mean does not exist, we consider signal cancellations in the general U-statistics framework with anti-symmetric kernels of order 2, and proposed another test that is more robust to detect location-shift by selecting bounded kernels. The ℓ^∞-norm of the U-statistic and its Gaussian multiplier bootstrap approximation are investigated, and no tuning parameter is needed in this scheme. Subject to mild conditions kernels, we derive similar rates of uniform convergence to our CUSUM-based test. Connection of two approaches and numeric studies are also provided.

Wednesday, March 20 at 4:00 PM in 636 SEO