The Chinese University of Hong Kong
Department of Statistics

Seminar

Heavy-tailed distribution for combining dependent p-values with asymptotic robustness

By

Professor George Tseng
Professor in Biostatistics
University of Pittsburgh, Graduate School of Public Health

Abstract

Abstract: The issue of combining individual p-values to aggregate multiple small effects is prevalent in many scientific investigations and is a long-standing statistical topic. Many classical methods are designed for combining independent and frequent signals in a traditional meta-analysis sense using the sum of transformed p-values with the transformation of light-tailed distributions, in which Fisher's method and Stouffer's method are the most well-known. Since the early 2000, advances in big data promoted methods to aggregate independent, sparse and weak signals, such as the renowned higher criticism and Berk-Jones tests. Recently, Liu and Xie (2020) and Wilson (2019) independently proposed Cauchy and harmonic mean combination tests to robustly combine p-values under "arbitrary" dependency structure, where a notable application is to combine p-values from a set of often correlated SNPs in genome-wide association studies. The proposed tests are the transformation of heavy-tailed distributions for improved power with the sparse signal. It calls for a natural question to investigate heavy-tailed distribution transformation, to understand the connection among existing methods, and to explore the conditions for a method to possess robustness to dependency. In this paper, we investigate the regularly varying distribution, which is a rich family of heavy-tailed distribution and includes Pareto distribution as a special case. We show that only an equivalent class of Cauchy and harmonic mean tests have sufficient robustness to dependency in a practical sense. We also show an issue caused by large negative penalty in the Cauchy method and propose a simple, yet practical modification. Finally, we present simulations and apply to a neuroticism GWAS application to verify the discovered theoretical insights and provide practical guidance.

Date:	September 14, 2021 (Tuesday)
Time:	9:00 a.m. - 10:00 a.m.
Meeting ID:	606 898 8598
Passcode:	cuhkstat
Zoom link:	https://cuhk.zoom.us/j/6068988598?pwd=Q1VTL2MyWTNDWlhuVFAvQWx6dHkrUT09

ALL INTERESTED ARE WELCOME !!