

BIOENGINEERING SEMINAR

FALL 2021

Reliable feature selection and prediction under adverse contamination

Abstract

Biomarker discovery studies have seen a proliferation with increasingly affordable high-throughput proteomics and genome sequencing. Given the abundance of such studies, the number of generalizable and clinically relevant discoveries is lacking. In particular, many discovery studies are based on samples of limited size and from potentially heterogeneous populations, while at the same time hundreds or even thousands of genes are sequenced. In addition to the heterogeneity of the population, this increasingly large pool of candidate biomarkers also comes with a greater risk of outliers and other forms of contamination. In this talk I will elucidate some of the detrimental effects of such "data-artefacts", specifically on biomarker discovery and related feature selection tasks, and how they can lead to spurious discoveries. I will then discuss a novel method for reliably identifying relevant biomarkers in the presence of such artefacts. The new method harnesses as much information in the data as possible, while not requiring prior specification of the form or source of artefacts. It is shown to outperform other methods commonly employed for biomarker discovery studies and I will demonstrate the usefulness of the method in a proteomic biomarker discovery study.

Biography

Dr. David Kepplinger is an Assistant Professor in the Department of Statistics at George Mason University. David obtained his PhD in Statistics in 2020 from the University of British Columbia. His research agenda comprises methodological and computational aspects of robust estimation in high-dimensional problems as well as their application to Biomedical Sciences. David is working on statistical methods with reliable performance under the presence of adverse contamination or outliers anywhere in the data, with a particularly focus on the robustness of feature selection and ways to counter the effects of contamination on predictive models.



David Kepplinger, PhD

Assistant Professor, Department of
Statistics
George Mason University
Fairfax, VA

**Thursday, October 21
12:00-1:00 pm**

Join Zoom Meeting

<https://gmu.zoom.us/j/91001788853?pwd=enZXRXFneG1zdk5NN1VM5WtnQlJMUT09>

Meeting ID: 910 0178 8853
Passcode: 229574