

## THE HONG KONG UNIVERSITY OF SCIENCE & TECHNOLOGY

### **Department of Mathematics**

# SEMINAR ON STATISTICS AND DATA SCIENCE

# **Boosting e-BH via conditional calibration**

By

## **Prof. Zhimei REN**

University of Pennsylvania

#### <u>Abstract</u>

The e-BH procedure is an e-value-based multiple testing procedure that provably controls the false discovery rate (FDR) under any dependence structure between the e-values. Despite this appealing theoretical FDR control guarantee, the e-BH procedure often suffers from low power in practice. In this paper, we propose a general framework that boosts the power of e-BH without sacrificing its FDR control under arbitrary dependence. This is achieved by the technique of conditional calibration, where we take as input the e-values and calibrate them to be a set of "boosted e-values" that are guaranteed to be no less—and are often more—powerful than the original ones. Our general framework is explicitly instantiated in three classes of multiple testing problems: (a) testing under parametric models, (b) conditional independence testing under the model-X setting, and (c) model-free conformalized selection. Extensive numerical experiments show that our proposed method significantly improves the power of e-BH while continuing to control the FDR. We also demonstrate the effectiveness of our method through an application to an observational study dataset for identifying individuals whose counterfactuals satisfy certain properties. This is joint work with Junu Lee.

#### **Biography**

Zhimei Ren is an assistant professor in the Department of Statistics and Data Science at the Wharton School, University of Pennsylvania. From 2021-2023, she was a postdoctoral researcher in the Statistics Department at the University of Chicago, advised by Professor Rina Foygel Barber. She obtained her Ph.D. in Statistics from Stanford University, advised by Professor Emmanuel Candès. Her research interests lie broadly in multiple hypothesis testing, distribution-free inference, causal inference, and data-driven decision-making.

> Date : 16 July 2024 (Tuesday) Time : 4:00pm Venue : Room 4472 (Lifts 25/26)

> > All are Welcome!