

ACMS Statistics Seminar

Claire Bowen

University of Notre Dame

Tues, November 28

154 Hurley Hall

3:30– 4:30 PM



Statistical Tests to Elect Partitions Sequentially (STEPS) and Its Application in Differentially Private Release and Analysis of Voter Registration Data

Voter data is important in political science research and applications such as improving youth voter turnout and predicting the presidential election outcome. This kind of data often contains sensitive information about the individuals in the data sets. One way of mitigating the privacy concern is removing identifiers in the “anonymized” data by linking it to other public data sets such as healthcare data or the Personal Genome Project Data. **Differentially Private Data Synthesis (DIPS)** techniques produce synthetic data or pseudo individual records at a preset level of privacy protection. Although DIPS provides a strong and robust privacy guarantee, statistical inferences drawn from the synthetic data can be poor due to the large amount of noise added to the data. We propose and apply a new approach called **Statistical Tests to Elect Partitions Sequentially (STEPS)** on voter data that allocates the privacy budget based on the statistical significance of the data’s variables. From the simulation study, SAFE preserves the variability and variable associations of the non-parametric DIPS approaches.

The Department of Applied and Computational
Mathematics and Statistics

Please visit acms.nd.edu to view the full list of speakers.