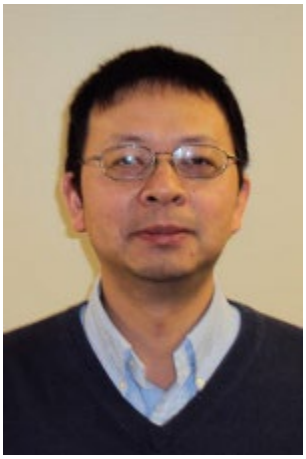
RUTGERS
School of Arts and Sciences

# Department of Statistics



**Min-ge Xie**
**Department of Statistics**
**Rutgers University**

## *Repro Samples Method for Addressing Irregular Inference Problems and for Unraveling Machine Learning Blackboxes*

**Wednesday, February 21th, 2024**
**11:50 AM**
**96 Frelinghuysen Road, CoRE Building, Room 431**

**Zoom Meeting: Meeting ID: 969 0606 4706**
**Password: 745339**
https://rutgers.zoom.us/j/96906064706?pwd=ZklvbExpRVBJQ3c5dUhhYTFuR2ZrZz09

**Light refreshments will be served in Hill 452, 11:15am**

**Abstract:** Rapid data science developments and the desire to have interpretable AI require us to have novel frameworks to tackle frequently seen, but highly non-trivial ``irregular inference problems,'' e.g., those involving discrete or non-numerical parameters and those involving non-numerical data, etc. This talk presents an innovative, effective and wide-reaching framework, called *repro samples method*, to conduct statistical inference for the irregular problems and more. We develop theories to support our development and provide effective computing algorithms for problems in which explicit solutions are not available. The method is likelihood-free and is particularly effective for irregular inference problems. For commonly encountered irregular inference problems that involve discrete or nonnumerical parameters, we propose a three-step procedure to make performance-guaranteed inferences for all parameters and develop a unique matching scheme that turns the disadvantage of lacking theoretical tools to handle discrete/nonnumerical parameters into an advantage of improving computational efficiency. The effectiveness of the proposed method is illustrated through case studies by solving two highly nontrivial classical inference problems in statistics: a) how to quantify the uncertainty in the estimation of the unknown number of components and make inference for the associated parameters in a Gaussian mixture; b) how to quantify the uncertainty in model estimation and construct confidence sets for the unknown true model, the regression coefficients, or both true model and coefficients jointly in high dimensional regression models. The method also has extensions to complex machine learning models, e.g., (ensemble) tree models, neural networks, graphical models, etc. It provides a new toolset to develop interpretable AI and to help address the blackbox issues in complex machine learning models.

**Bio:** **Min-ge Xie, PhD** is a Distinguished Professor at Rutgers, The State University of New Jersey. Dr. Xie received his PhD in Statistics from University of Illinois at Urbana-Champaign and his BS in Mathematics from University of Science and Technology of China. He is the current Editor of *The American Statistician* and a co-founding Editor-in-Chief of *The New England Journal of Statistics in Data Science*. He is a fellow of ASA, IMS, and an elected member of ISI. His research interests include theoretical foundations of statistical inference and data science, fusion learning, finite and large sample theories, parametric and nonparametric methods. He is the Director of the Rutgers Office of Statistical Consulting and has a rich interdisciplinary research experiences in collaborating with biomedical researchers, computer scientists, engineers, and scientists in other fields.