





## High Dimensional Nonstationary Time Series IRTG 1792 Short Course

## Wei Peng

## Asymptotic Distributions and Rates of Convergence for Random Forests

Random forests remain among the most popular off-the-shelf supervised learning algorithms. Despite their well-documented empirical success, however, until recently, few theoretical results were available to describe their performance and behavior. We push beyond recent work on consistency and asymptotic normality by establishing rates of convergence for random forests and other supervised learning ensembles. We develop the notion of generalized U-statistics and show that within this framework, random forest predictions can potentially remain asymptotically normal for larger subsample sizes than previously established. We also provide Berry-Esseen bounds in order to quantify the rate at which this convergence occurs, making explicit the roles of the subsample size and the number of trees in determining the distribution of random forest predictions.



Wei Peng is a Ph.D. student at the Department of Statistics at University of Pittsburgh. He received his Bachelor's degree in Mathematics at Nanjing University in 2015. In his research, Wei is interested in statistical inference of machine learning models, particularly random forests.

## February 9, 2021 | 15:00-16:30 | Online via Zoom

