MISSOURI S&T

Mathematics & Statistics Colloquium

Monday, December 11, 2023, 4:15pm-5:15pm

G5 Rolla Building and online via zoom[1]

Please join for refreshments and coffee (start at 4pm) before the talk!

Dr. Nilanjan Chakraborty

William Chauvenet Postdoctoral Lecturer
Department of Statistics and Data Science
Washington University

Statistical Inference for Subgraph Densities Under Induced Random Sampling from Network Data

Abstract. Network data arises in numerous applied disciplines viz., political science, sociology, biology, computer science, and economics. Real-life networks are large and often cannot be fully observed. Hence, statistical inference for large networks based on a suitable sampling scheme is considered to be an important problem in network analysis. We develop a framework for obtaining statistical guarantees for subgraph densities of a general population network under without replacement sampling (SRSWOR). The examples of such subgraph densities include edge density, triangle density, two-star density and other popularly studied graph summary statistics. Under this sampling scheme, we derive a Berry-Esseen bound to establish the asymptotic normality of the Horwitz-Thompson (HT) estimator for the population subgraph densities. The HT estimator is shown to be unbiased for population subgraph densities. To facilitate inferential procedures, we provide a jackknife estimator of the unknown population variance and establish its consistency. We provide explicit expressions for the Berry-Esseen bound under certain popular super population network models like sparse graphons, stochastic block models (SBM), preferential attachment models (PFA), random dot product graph models (RDPG) and others. We also establish the joint asymptotic normality of two subgraph densities which is crucial in establishing the asymptotic normality of the global clustering coefficient of the sampled graph. Our results find a useful application to the problem of testing the equality of two population graphs using the subgraph densities as the test statistic. Finally, we present a simulation study and a real data analysis to corroborate our theoretical findings. This is a joint work with Ayoushman Bhattacharya (WashU) and Prof. Soumen Lahiri (WashU).

Biographical Sketch. I am currently working as a William Chauvenet Postdoctoral Lecturer in the Department of Statistics and Data Science at the Washington University in Saint Louis. My Postdoctoral mentors are Prof. Soumen Lahiri and Prof. Robert Lunde. I received my Ph.D. in Statistics from the Department of Statistics and Probability at Michigan State University. My thesis advisors were Prof. Hira Koul and Prof. Lyudmila Sakhanenko. Prior to that, I obtained my Master degree in Statistics from the University of Calcutta and my Bachelor degree in Statistics from St. Xavier's College Kolkata, India. I am broadly interested in developing rigorous statistical methods for analyzing complex data, viz. High-Dimensional data and Networks when issues such as arbitrary dependence and noise accumulation in these datasets impede inferences using standard methods. The following are my current research interests : Resampling methods for complex data, Time Series Analysis, Network Regression, Gaussian Approximation, Missing Data and Neuroimaging.

[1]Zoom meeting ID: 98803918945
password: mathstat