BIOSTATISTICS SEMINAR

Oscar Hernan Madrid Padilla, PhD
Assistant Professor of Statistics
UCLA

Learning Gaussian DAGs from Network Data

Abstract:

Structural learning of directed acyclic graphs (DAGs) or Bayesian networks has been studied extensively under the assumption that data are independent. We propose a new Gaussian DAG model for dependent data which assumes the observations are correlated according to an undirected network. Under this model, we develop a method to estimate the DAG structure given a topological ordering of the nodes. The proposed method jointly estimates the Bayesian network and the correlations among observations by optimizing scoring function based on penalized likelihood. We show that under some mild conditions, the proposed method produces consistent estimators after one iteration. Extensive numerical experiments also demonstrate that by jointly estimating the DAG structure and the sample correlation, our method achieves much higher accuracy in structure learning. When the node ordering is unknown, through experiments on synthetic and real data, we show that our algorithm can be used to estimate the correlations between samples, with which we can de-correlate the dependent data to significantly improve the performance of classical DAG learning methods.

October 7, 2021   **133 Rosenua Hall**   **3:30-4:30 PM**

Zoom Link:
https://unc.zoom.us/j/93545206596? pwd=N1IKeVZjSFhuM2lhSDJCWjN3c21BUT09
Meeting ID: 935 4520 6596   Passcode: 823321

UNC | GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH