

R. Clifton Bailey Statistics Seminar Series

Sampling-Based Estimation of in-Degree Distribution in Directed Complex Networks

Vladas Ppiras

**Department of Statistics and Operations
University of North Carolina at Chapel Hill**

Student Union 1 - Room 3B
4400 University Drive, Fairfax, VA 22030
November 30, 2018
11:00 A.M. - 12:00 Noon

Abstract: The focus of this work is on estimation of the in-degree distribution in directed networks from sampling network nodes or edges. A number of sampling schemes are considered, including random sampling with and without replacement, and several approaches based on random walks with possible jumps. When sampling nodes, it is assumed that only the out-edges of that node are visible, that is, the in-degree of that node is not observed. The suggested estimation of the in-degree distribution is based on two approaches. The inversion approach exploits the relation between the original and sample in-degree distributions, and can estimate the bulk of the in-degree distribution, but not the tail of the distribution. The tail of the in-degree distribution is estimated through an asymptotic approach, which itself has two versions: one assuming a power-law tail and the other for a tail of general form. The two estimation approaches are examined on synthetic and real networks, with good performance results, especially striking for the asymptotic approach.