



THE UNIVERSITY OF
CHICAGO

Department of Statistics

STATISTICS COLLOQUIUM

DAVID DUNSON

Department of Statistical Science
Duke University

Probabilistic modeling of high-dimensional tables and networks

MONDAY, January 11, 2016, at 4:00 PM

Eckhart 133, 5734 S. University Avenue

Refreshments following the seminar in Eckhart 110.

ABSTRACT

High-dimensional discrete data are collected in many application areas, but have seen limited consideration in the literature. We focus in particular on two related problems: (i) high-dimensional categorical data that can be organized as a many way contingency table; and (ii) graph/network-valued data consisting of a binary adjacency matrix measured for each subject. The first problem arises in numerous applications ranging from survey research in social sciences and epidemiology to genomics and marketing. The second arises in analysis of brain connectomes and team dynamics. We propose a general approach for these problems based on probabilistic tensor factorizations, which are low rank and sparse. Taking a Bayesian nonparametric approach to inference, we show large support, posterior consistency and convergence rates in the high-dimensional setting. Efficient Markov chain Monte Carlo algorithms are developed for practical implementation, and the methods are illustrated through applications in epidemiology and neuroscience.