# MICHIGAN STATE UNIVERSITY <br> Department of Statistics and Probability 

## COLLOQUIUM

Xiaofeng Shao
University of Illinois at Urbana-Champaign

# Martingale Difference Divergence and Its Applications to Contemporary Statistics 

Tuesday, December 6, 2016<br>10:20 a.m. - 11:10 am<br>Refreshments 10:00 am<br>C405 Wells Hall


#### Abstract

Martingale difference divergence (MDD) is a metric that quantifies the conditional mean dependence of a random vector $Y$ given another random vector $X$. We shall present applications of Martingale difference divergence and its variant to two contemporary statistical problems: high dimensional dependence testing and dimension reduction for multivariate time series. In the first part, we propose a novel test to assess the conditional mean dependence of a response variable on a large number of covariates. Our MDD-based procedure is able to detect certain type of departure from the null hypothesis of conditional mean independence without making any specific model assumptions. We establish the asymptotic normality of the proposed test statistic under suitable assumptions that can be verified for covariates with banded dependence or Gaussian distribution. Power analysis and a wild bootstrap procedure will also be presented along with some simulation results. In the second part, we introduce a new methodology to reduce the number of parameters in multivariate time series modeling. In particular, we seek a contemporaneous linear transformation such that the transformed time series has two parts with one part being conditionally mean independent of the past information. Our dimension reduction procedure is based on eigendecomposition of the so-called cumulative martingale difference divergence matrix, which encodes the number and form of linear combinations that are conditionally mean independent of the past. We provide a simple way of estimating the number of factors and factor loading space, and obtain some theoretical results about the estimators. The finite sample performance will be illustrated from a real data analysis.


