

MICHIGAN STATE UNIVERSITY  
Department of Statistics and Probability

## COLLOQUIUM

**Kjell Doksum**

Univ. of Wisconsin, Madison

### **Ensemble Subspace Methods for High Dimensional Data**

Friday, October 7, 2016  
10:20 a.m. - 11:10 am  
Refreshments 10:00 am  
C405 Wells Hall

#### **Abstract**

We consider high dimensional regression frameworks where the number  $p$  of predictors exceed the number  $n$  of subjects.

Recent work in high dimensional regression analysis has embraced an approach that consists of selecting random subsets with fewer than  $p$  predictors, doing statistical analysis on each subset, and then merging the results from the subsets. This ensemble approach makes it possible to construct methods for high dimensional data using methods designed for low dimensional data. Moreover, penalty methods such as Lasso that are unstable when  $p > n$  unless very stringent conditions are imposed, often produce more efficient predictors when used in the ensemble approach. We examine the extent of the prediction risk improvement achieved by the ensemble approach when it is applied to Lasso, Lars, SCAD, the Elastic Net, and predictors based on variable selection methods.

In particular, we use simulation results and real data to examine the relative risks of these procedures as we: (1) increase  $p$ , (2) increase correlation between variables, (3) increase the number of weak signals.

This is joint work with Saeid Amiri and Ejaz Ahmed

*To request an interpreter or other accommodations for people with disabilities, please call the Department of Statistics and Probability at 517-355-9589.*