

# Department of Applied and Computational Mathematics and Statistics Colloquium



**Yunjin Choi**

Department of Statistics  
Stanford University

## ***SELECTING THE NUMBER OF PRINCIPAL COMPONENTS: ESTIMATION OF THE TRUE RANK OF A NOISY MATRIX***

Principal component analysis (PCA) is a well-known tool in multivariate statistics. One significant challenge in using PCA is the choice of the number of principal components. In order to address this challenge, we propose an exact distribution-based method for hypothesis testing and construction of confidence intervals for signals in a noisy matrix with finite samples. Assuming Gaussian noise, we derive exact results based on the conditional distribution of the singular values of a Gaussian matrix by utilizing a post-selection inference framework. In simulation studies we find that our proposed methods compare well to existing approaches.

**Monday, January 18, 2016**

**4:15 PM – 5:15 PM**

**129 Hayes-Healy Center**

Colloquium Tea 3:45 PM to 4:15 PM 154 Hurley Hall