Stanford University  
Department of Statistics

DEPARTMENTAL SEMINAR

4:15pm, Tuesday, July 23, 2013  
Sequoia Hall Room 200

Cookies served at 3:45pm, 1st Floor Lounge.

Speaker: Julie Josse, Agrocampus Ouest Rennes, France

Title: Missing values in PCA: from single to multiple imputation

Abstract:

Missing values are ubiquitous in statistical practice. They are problematic since most statistical methods can’t be applied directly on an incomplete dataset. In this talk, we focus on handling missing values in principal component analysis (PCA). First, we describe the regularized PCA imputation algorithm that estimates the scores and loadings and overcomes the issue of overfitting. Then, we point out how single imputation methods underestimate the variability and discuss multiple imputation solutions. The suggested methods can be used to complete data with continuous, categorical or mixed variables.