Syllabus for STA 232A, Fall 2013

Course Website: `www.stat.ucdavis.edu/~beran/s232a.html`

Instructor: Rudolf Beran (rjberan@ucdavis.edu). No office telephone.
Instructor Office Hours: Tue and Thurs 11 am–12 noon (and by appointment) in MSB 4224.

TAs: Apratim Ganguly (aganguly@ucdavis.edu) and Cong Xu (cgxu@ucdavis.edu).
TA Office Hours: Mon and Fri 3 pm–4 pm (AG); Mon 1–2 pm and Wed 2:30–3:30 pm (CX) in MSB 1117.

Lectures: Tue 2:10–3:00 pm and Thurs 2:10-4:00 pm in CHEM 166, starting Sept 26.
Discussion: Tue 3:10–4:00 pm in CHEM 166, starting Oct 1.

Prerequisites: This is a core course in the Statistics Ph.D. program. It presupposes an upper division background in regression and analysis of variance; linear algebra, including the singular value decomposition; ability to do numerically stable linear algebra using R or other professional software; advanced calculus; upper division probability theory, including joint distributions, the multivariate normal distribution, mean vectors and covariance matrices, central limit theorems and laws of large numbers.

Labs: The Lab assignments will be due in the Discussion Section on Tuesdays, starting October 8. These labs will bring course theory to bear on the analysis of data sets. Doing most labs will require writing numerically stable computer code to carry out the necessary linear algebra.

Midterms: Midterm 1 (on October 29) and Midterm 2 (on December 5) will each be administered in class at 2:10–4:00 pm. There will be no final exam.

Course Sketch: The prerequisites in linear algebra, probability theory, and R or equivalent software are essential. The course will highlight the following topics:

- Classical linear regression and analysis of variance, updating Part I of *Analysis of Variance* by Henry Scheffé (1959), for instance.
- Using the Moore-Penrose pseudoinverse and the singular value decomposition to express least squares theory and algorithms. *Matrix Analysis for Statistics* (second edition) by James R. Schott (2005) is a useful reference.
- The limitations of unbiased estimation. Superior biased estimation in linear models through bias-variance trade-off. Adaptive submodel fits, adaptive penalized least squares fits, and their embedding within the class of adaptive hypercube estimators for linear models. For the latter, see *Hypercube estimators: Penalized least squares, submodel selection, and numerical stability* at `http://dx.doi.org/10.1016/j.csda.2013.05.020`
- Numerical instability of penalized least squares estimators versus numerical stability of the equivalent hypercube estimators. *Matrix Computations* (third edition) by Gene Golub and Charles Van Loan (1996) provides extensive background on numerical stability.

Text: None. Lectures will outline the ideas and results. The references cited above are an in-depth resource for more.

Grading: The course grade will be based 35% on Midterm 1, 35% on Midterm 2, and 30% on the Lab assignments.

Software: Students may use any professional-level software that can perform the numerical linear algebra required in the Lab assignments. Solution code for the Labs will be provided solely in R. The site `http://cran.r-project.org` offers binaries and documentation for R. The link `http://bendixcarstensen.com/APC/linalg-notes-BxC.pdf` is an introductory tutorial on linear algebra in R.