Syllabus:
Topics in Experimental Design,
Stat 8313, Spring 2020.

November 20, 2019

In this course we review the fundamentals of experimental design and discuss contemporary topics, leading eventually to current research. A list of topics is give in this syllabus.

Course Instructors:

R. D. Cook, 397 Ford (5-7732).
Email: dennis@stat.umn.edu.
Office Hours: By appointment.

C. J. Nachtsheim, 3-227 Carlson School of Management (4-1077).
Email: nacht001@umn.edu.
Office Hours: By appointment

Lectures: Mon, Wed, Fri 11:15AM - 12:05PM; FORD HALL 110. Alternate times for some lectures/presentations may be scheduled, particularly near the end of the semester.

Prerequisites: Stat 5303 and background in linear models.

Text: None. Readings will be assigned from the literature.

Course Web Page: http://www.stat.umn.edu/∼dennis/Stat8313S20/ (Not yet available.)

Homework: Homework is a required part of the course. There will be homework assignments throughout the semester, portions of which will be graded. Selected students will likely be asked to present their solutions on the day the homework is returned.

Grading: A grade of "B" requires satisfactory completion of the homework problems and reading assignments, along with regular attendance and participation in classroom discussion. A grade of "A" requires completion of a class project involving detailed study of some aspect of the course material, followed by a paper and class presentation. Projects require a written proposal that must be approved in advance, and they should be well underway by April 1. Project suggestions will be given from time to time (see below). You should expect to spend about $1/4$ of your time on the project.

Exams: None planned at present. Depending on the number of projects, presentations will begin around April 20 and some might be scheduled after the last day of instruction, May 4. Presentations can be delivered before April 20 and the student's discretion. Written project reports are due at the time of presentation and copies should be distributed to the class.

Incomplete: Grades of 'I' will be given only in extraordinary circumstances, and then only by written agreement between the instructor and the student.

Computing: JMP, Matlab, and/or R will be employed for computing. These are available to all enrolled UMN students.

Topics: Roughly the first half of the course will be on foundations, with second half being more applied and computer oriented. Topics are subject to change, depending on the progress of the course.

1. Review of regression
2. Fundamentals of experimental design
   - Randomization
   - Replication
   - Local control (blocking generally)
   - Review of classical (catalog) design: Completely randomized, randomized complete block and latin square designs, balanced and partially balanced incomplete block designs, split plot and split block designs, ...
3. Optimal design basics
   - Optimal designs, criteria
   - Relative efficiencies
   - Variances of estimated parameters
   - Equivalence theorems for optimal designs
   - Fixed covariate designs

4. Design Algorithms
   - Approximate design, algorithms
   - Exact designs
   - Row and column exchanges, coordinate exchange algorithm
   - Optimal blocking
5. Diagnostics and design comparisons
   - Power
   - Correlation cell plots
   - FDS plots
6. Relaxing assumptions
   - Model robust designs
   - Bayesian designs
7. Screening and supersaturated designs
   - Regular vs. nonregular designs
   - Supersaturated designs
   - Minimum alias designs
   - Definitive screening designs
8. Nonlinear design
   - Nonlinear design, designs for logistic regression
   - Optimal block and split-plot designs
   - Discrete choice experiments
9. Designs for computer experiments
   - Latin hypercube
   - Uniform
   - Maximin and minimax distance
   - Sphere packing

Possible Projects: Many of the projects described below are stimulated by referenced articles. In such cases, the project must be more than just a recitation of the results in the articles, should not necessarily be confined to the paper cited and must include an element of original research. The same articles may be used for multiple projects, provided that they are meaningfully distinct. Additional projects may be given during lecture.

Project 1 *Construct a design protocol for turkey experiments to be described in lecture, and provide computer code for generating the designs.*

Project 2 *Study the role for experimental design in big data.*

Project 3 *Seeger, M. W. (2008). Bayesian Inference and Optimal Design for the Sparse Linear Model.* Journal of Machine Learning Research **9**, *759–813.*

Project 4 *FMRI Design:*

Kao, M-H,. Majumdar, D., Lazar, N. and Stufken, J. (2009). Multi-objective optimal experimental designs for event-related fMRI studies. NeuroImage **44** *849 – 856.*

Ming-Hung Kao (2009). Multi-Objective Optimal Experimental Designs for ER-fMRI Using MATLAB. Journal of Statistical Software **30**, *Issue 11.*

Kao, M-H,. Majumdar, D., Mandal, A. and Stufken, J. (2013). Maximin and maximum-efficient event-related FMRI designs under a nonlinear model. The Annals of Applied Statistics **7**, *1940 – 1959.*

Project 5 *There has been relatively little work done on optimal design for multivariate linear regression. Here are three papers on the topic. There is more literature, of course. Some will be discussed in class.*

Krafft, O. and Scheffer, M. (1992). D-Optimal Designs for a Multivariate Regression Model . Journal of Multivariate Analysis **42**, *130–140.*

Imhof, L. (2000). Optimum designs for a multi response regression model. Journal of Multivariate Analysis **72**, *120–131.*

Marget, W. M. and Morris, M. D. (2019). Centeral composite experimental designs for multiple responses with different models. Technometrics, *https://doi.org/10.1080/00401706.2018.1549102.*

Project 6 *Dette, H., Pepelyshev, A. and Zhigljavsky, A. (2013). Optimal design for linear models with correlated observations.* Annals of Statistics, **41**, *143–176.*

Project 7 *Dette, H., Melas, V. B., , and Wong, W. K. (2005). Optimal Design for Goodness-of-Fit of the Michaelis-Menten Enzyme Kinetic Function.* Journal of the American Statistical Associatio **100**, *1370–1381.*

Project 8 *Atkinson, A. and Cook, R. D. (1995). D-Optimum Designs for Heteroscedastic Linear Models.* Journal of the American Statistical Association **90***, 204 – 212*

Project 9 *Something on randomization theory.*

Project 10 *Study ways to deal with a situation in which there is no conclusive information on the end points of the design space* $\times_{j=1}^{p}[L_j, U_j]$ *for* p *predictors, providing a computer code to implement your suggestions.*

Project 11 *Study ways to proceed when two or more criteria are relevant and we cannot optimize all simultaneously. See for example*

Cook, R. D. and Wong, W. K. (1994). On the equivalence of constrained and compound optimal design. Journal of the American Statistical Association*,* **89***, 687–692.*

Clyde, M. and Chaloner, K. (1996). The Equivalence of Constrained and Weighted Designs in Multiple Objective Design Problems. Journal of the American Statistical Association **91***, 1236–1244.*

Project 12 *Develop the best strategy for augmenting Definitive Screening Designs for estimation of the full quadratic (response surface) model. See for example*

Jones, B., and C. J. Nachtsheim (2011). "A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects". Journal of Quality Technology, 43, pp. 1–14.

Jones, B., and C. J. Nachtsheim (2013). "Definitive Screening Designs with Added Two-Level Categorical Factors." Journal of Quality Technology, 45, pp. 121–129.

Project 13 *Develop exact designs for moving (nonparametric) regression. See for example*

Fedorov, V. V., Montepiedra, G., and Nachtsheim, C. J. (1999). "Design of Experiments for Locally Weighted Regression," Journal of Statistical Planning and Inference, 81, 363–382.

Project 14 *Are there multiple non-isomorphic conference matrices (of a given size) and if so do some, when used to construct definitive screening designs, have better model estimation properties than others. Do rotations of definitive screening designs lead to better model robustness properties? See for example*

Xiao, L., Lin, D. K., and Bai, F. (2012). Constructing definitive screening designs using conference matrices. Journal of Quality Technology *44(1), 2–8.*

Project 15 *Investigate designs for experiments based on linear models with* $n < p$*. See*

Shao, J. and Deng, X. (2012). Estimation in high-dimensional linear models with deterministic design matrices. Annals of Statistics *4, 812–831.*