

Contact info: Professor Hal Stern
 Department of Statistics
 2216 Bren Hall ZOT 1250
 email: sternh@uci.edu
 phone: 949-824-1568
 website: www.ics.uci.edu/~sternh/courses/210/

Class hours Lecture: MWF 11:00am-11:50am ET 204
 Discussion: T 4:00pm - 4:50pm SSPA 1170

Office hours: M 12:00pm - 1:00pm (and by appointment)

Course goals: Introduction to statistical methods for analyzing data from experiments, observational studies and surveys. Methods covered include two-sample procedures, analysis of variance, linear regression. Course covers both the underlying theory behind these methods and how the methods are applied in practice.

Prerequisites: Knowledge of basic statistical methods (at the level of Statistics 7), knowledge of undergraduate probability and statistics (at the level of Statistics 120ABC), calculus, linear algebra.

Course outline/schedule (tentative)

F Sept 25 - M Oct 5 (5 lectures): Introduction to inference / Inference in the two-sample case
 Randomization-based and model-based approaches to two-sample inference including discussion of study design and model checking.

(Reading: *Chapter 15, Appendix A.6-7*)

W Oct 7 - M Oct 19 (6 lectures): Analysis of variance (ANOVA)

One-way analysis of variance for inference concerning more than two samples/populations including discussion of planned contrasts and multiple comparisons. Use of pairing/blocking to reduce variance.

(Reading: *Chapters 16-18, 21*)

W Oct 21 - M Oct 26 (3 lectures): Simple linear regression and correlation

(Reading: *Chapters 1-5*)

T Oct 27 - Makeup lecture in discussion

F Oct 30 - No lecture

T Oct 27 - M Nov 9 (4 lectures): Multiple linear regression: basics

(Reading: *Chapters 6-7*)

W Nov 4 - MIDTERM EXAM

F Nov 6 - No lecture

T Nov 10 - Makeup lecture in discussion

W Nov 11 - No lecture (holiday)

T Nov 10 - W Nov 18 (4 lectures) - Multiple regression: model diagnostics and remedies

Evaluating regression assumptions and potential remedies; study of outlying and influential cases.

(Reading: *Chapters 10-12*)

F Nov 20 - W Nov 25 (3 lectures) - Multiple regression: model building, model selection

Indicator variables, polynomial regression, relationship of ANOVA and regression, model selection.

(Reading: *Chapters 8-9*)

F Nov 27 - NO lecture (holiday)

M Nov 30 - F Dec 4 (3 lectures) - Factorial experiments and ANOVA, makeup/review topics

(Reading: *Chapters 19-20, 23-24*)

Discussion: There is a discussion section associated with this class on Tuesdays at 4pm. This discussion serves several purposes: (1) It is an ideal time to ask questions about class/homework; (2) It is an ideal time for SAS instruction/discussion; (3) It can be used as needed for me to present additional material or make up for a missed lecture. You are not required to attend discussion UNLESS I ask you to attend on a given week (see schedule above). Attendance is highly recommended in any case – if you do come, please bring your questions!

Exams: There will be an in-class midterm exam. This is scheduled for Wednesday November 4, 2009. The final exam will consist of two elements: (1) Take-home data analysis; (2) Standard two-hour final exam. The latter will occur on Friday December 11, 2009 from 8:00-10:00am.

Homework: Students will be assigned (approximately) weekly homework. Homework will be generally due on Wednesday. These assignments will include some application-oriented questions and some theory questions. The homework will also **require** the use of the SAS statistical package. You are allowed to **discuss** homework with other students and with me. Every student **must** write-up the homework **independently**. Failure to complete your homework independently will be punished – students who turn in the same work will split the grade! Late homework will be penalized but it is much better to turn homework in late than not at all; homework turned in after solutions have been posted will not earn any credit.

SAS: Yes, you read that correctly, SAS is a required element of the course. This is done because SAS is widely used in industry and in medical research. Though you may not find it terribly fun and it will not be used in other courses, some experience with SAS is a good thing – really, trust me! I will provide SAS instruction but alas I must confess that I am not a regular user. SAS is available on computers in ICS 346 (approx 5 machines only) and via internet connection on the Broadcomm Distributed Unified Cluster (BDUC) (details on using BDUC to follow). You may also purchase SAS from NACS for \$45 (good for a one-year license). Latest version is Version 9.2.

Grading: Course grades are determined by performance on all work as follows: homework (approx 25%), midterm exam (approx 30%), final exam (approx 45%).

Course Evaluations: Course evaluations will be carried out on-line at the end of the semester. They will be available for you to complete from (approximately) Friday Nov 27 at 7pm and will remain available until Sunday Dec 6 at midnight. I mention this now because your feedback is valuable to me and I want you to remember that we will be asking for it!

Textbooks and References:

Primary text: *Applied Linear Statistical Models*, MH Kutner, CJ Nachtsheim, J Neter, W Li (5th ed), McGraw Hill Irwin, 2005.

Material will also be drawn from: *The Statistical Sleuth*, FL Ramsey, DW Schafer (2nd ed), Duxbury Press, 2002.

References (intro stat):

The Basic Practice of Statistics, DS Moore (2nd ed), WH Freeman, 2000.

Introduction to the Practice of Statistics, DS Moore, GP McCabe (4th ed), WH Freeman, 2003.

Mind on Statistics, JM Utts and RF Heckard (2nd ed), Thomson/Brooks Cole, 2004.

References (prob and stat):

Mathematical Statistics and Data Analysis, JA Rice (2nd ed), Duxbury, 1995.

An Introduction to Mathematical Statistics and its Applications, RJ Larsen, ML Marx (4th ed), Prentice Hall, 2005.