

Stat 431, Fall 2008

Statistical Inference

[Homework assignments](#) [Readings](#) [Schedule of topics](#) [Announcements](#)

People

			Office hours
Instructor	Mikhail Traskin	mtraskin@wharton	Monday/Wednesday, 10:30 am – 12:00 pm, 466 Jon M. Huntsman Hall
TA	Oliver Entine	entine4@wharton	TBA

Additional help

Visit the StatLab: location and schedule at <http://www-stat.wharton.upenn.edu/~braunsf/statlab.html>.

Lectures

1203 Steinberg Hall - Dietrich Hall. Monday/Wednesday, 1:30 – 3:00 pm.

Course homepage

Refer to <http://stat.wharton.upenn.edu/~mtraskin/courses/stat431/index.html> (this page) for announcements, handouts, homework assignments and other materials.

Course description

This course is about making decisions under uncertainty using statistical methods. The topics include estimation, confidence intervals, hypothesis testing, single and multiple linear regression, one-way and two-way analysis of variance, variable selection, logistic regression and categorical data analysis.

Interpretation of the results and analysis of assumptions is an important part of the course. [Statistical computing package](#) will be extensively used to carry out the computations. However no special emphasis will be made on the details of computations.

Prerequisites

Familiarity with basic probability theory is assumed. Stat 430 or equivalent should provide sufficient background. Otherwise chapters 2, 3, 4 and 5 of the [Devore's book](#) will be a good substitute together with a [short review](#) of fundamental concepts used by the course, written by Prof. David Freedman, University of California, Berkeley.

Statistical computing package

JMP version 7, available in the Wharton Computer Labs, F75/F80 Jon M. Huntsman Hall (Wharton

account required: see <http://accounts.wharton.upenn.edu>). Individual copies are also available for purchase at <http://estore.e-academy.com>. A six-month license costs \$29.95 and twelve-month is \$49.95.

You may also use R, an open-source statistical software that is available from [The R Project for Statistical Computing](http://www.r-project.org) web page.

Text book

J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 7th ed.

Assignments

Grading

- Homework assignments: 20% (lowest score excluded)
- Midterm: 30%
- Final: 50%

Homework assignments

Weekly assignments will be due at the *beginning* of lecture each Monday. Problems involving computer calculations should be worked using JMP or R. **No extensions to the due date will be given.** However, the lowest homework assignment score will be omitted from the final grade calculation. Unsubmitted work counts as a zero score.

You may work with and help each other, *however you must submit your own solutions, with your own writeup, unless otherwise noted.*

Exams

In-class midterm: date TBA, 1:30 – 3:00 pm, 1203 Steinberg Hall - Dietrich Hall. No make-up.

Final: Wednesday, December 17th, 9:00 – 11:00 am, location TBA.

Both exams are open notes. Calculators may be used but no laptops are allowed.

Readings

- Paul F. Velleman, "Truth, Damn Truth, and Statistics", *Journal of Statistics Education*, Volume 16, Number 2 (2008), <http://www.amstat.org/publications/jse/v16n2/velleman.html>

Schedule of topics

Lectures will closely follow the text which will be occasionally supplemented with handouts on topics going beyond those covered in the book. A review of basic probability theory and common distributions, e.g. chapters 2, 3, 4 and 5 in Devore, might be useful. Topics discussed in sections 3.4 (binomial distribution), 4.3 (normal distribution) and 5.4 (central limit theorem) are of special interest.

Lec#	Date	Topic	Text
------	------	-------	------

01	Wed 03 Sep	Introduction/overview	
02	Mon 8 Sep	Normality; boxplots; QQ plots (probability plots)	1.4, 4.6
03	Wed 10 Sep	Confidence intervals: known variance; Large-sample CIs	
04	Mon 15 Sep	Confidence intervals for population proportion; CIs: unknown variance; Non-normal population distribution	
05	Wed 17 Sep	One-sample hypothesis testing	
06	Mon 22 Sep	One-sample hypothesis testing	
07		Two-sample inference: testing and intervals	
08		Two-sample inference: testing and intervals	
09		Single factor (one-way) ANOVA	
10		Single factor (one-way) ANOVA	
11		Two factor (two-way) ANOVA	
12		Two factor (two-way) ANOVA	
13		Simple linear least-squares regression	
14		Simple linear least-squares regression; Correlation coefficient	
15		Simple regression and variable transformation	
16		Simple nonlinear regression; polynomial regression	
17		Multiple linear least-squares regression	
18		Multiple linear least-squares regression	
19		Variable selection	
20		Logistic regression	
21		Categorical data analysis: goodness of fit	
22		Categorical data analysis: goodness of fit	
23		Bootstrap	
24		Distribution-free procedures	

Announcements