

**Department of Statistics
The Wharton School
University of Pennsylvania**

Statistics 431

Spring 2007

Professor: Paul Shaman
Email: shaman@wharton.upenn.edu
Office: 468 JMHH
Telephone: 215-898-8749

Office hours: M 3-6, TTh 4:30-6 and by appointment

Classes meet: Section 001, MW 10:30-11:50, in F65 JMHH
Section 002, MW 1:30-2:50, in F65 JMHH

Teaching Assistants and Stat Lab

Name:	Kartikeya Ghia	Juntian Xu
Email:	kghia@wharton	juntianx@wharton
Office/Phone:	432.2 JMHH/215-898-6810	432.4 JMHH/215-898-1252

The Stat Lab is available for help. It is located in F96 JMHH. Hours are listed at stat.wharton.upenn.edu/~juntianx/statlab.html.

Course website

Statistics 431 is using webCafe. You can gain access by going to <http://webcafe.wharton.upenn.edu> and following the link to the Statistics Department. All materials for this course will be distributed and managed via the website, and you will be able to monitor your grade entries throughout the semester.

Note for non-Wharton students: If you do not have a Wharton computing account, you will need to establish one to access the website. The account also provides access to the computing labs in Wharton and to the intranet. To get an account, go to

<http://accounts.wharton.upenn.edu>

After you have obtained your account, allow some time for activation.

Wharton students and students who have recently taken a Wharton course have existing accounts.

Course Overview

This course develops ideas for helping to make decisions using statistical methods. The topics are estimation, hypothesis testing, regression analysis, the analysis of variance and logistic regression. If time permits, we will consider categorical data analysis and nonparametric methods. Students beginning the course are expected to have some familiarity with data display (including boxplots, histograms and scatterplots), summary statistics (including mean, standard deviation and quantiles) and the binomial and normal distributions, but these topics will be reviewed and discussed as we encounter them.

The course will emphasize critical interpretation and analysis of assumptions. We will use JMP IN to carry out computations. The course does not dwell much on the details of computation—its main focus is understanding, interpretation and communication of statistical results.

Materials

Probability and Statistics for Engineering and the Sciences, 6th ed., by J. L. Devore, Brooks/Cole, Belmont, CA, 2004. Sections 1.4, 3.4, 4.3, 4.6, 5.4 and 6.1, and Chapters 7-13 (several sections from these chapters will be skipped) will be covered. If time permits, we will address some topics in Chapters 14 and 15.

Class notes. The text will be supplemented with notes, which will be posted on webCafe.

JMP IN 5.1 statistical software and handbook. I *highly recommend* you buy the software so that you have it on your own computer. We will use it extensively in class and you will need to know how to read its output and use it for assignments and for reading and interpreting quiz and examination questions. When you install the software on your computer you will also have installed five manuals, all in pdf format. JMP IN is available on machines in the Wharton Computer Labs located in F75 (60 seats) and F80 (29 seats), both in JMHH.

JMP IN 5.1 is available at the bookstore. As an alternative, you may purchase JMP 6 at www.e-academy.com. A six-month license costs \$29.95 and a twelve-month license sells for \$49.95. For our purposes the differences between releases 5.1 and 6 are very minor.

Homework

- There will be eight homework assignments.
- Each homework will be assigned at a lecture and will be due in class a week later unless otherwise noted.
- *Homework will not be accepted late.*
- The homework is designed to teach and you are encouraged to seek help from the instructor and the TAs if you have questions. You may also work with and help each other. *Unless otherwise instructed, though, you must submit your own solutions, with your own writeup.*

Examinations

Two midterm examinations are scheduled:

Both will be in the evening, 6-8 pm, dates to be announced.

The final examination will be Thursday, 26 April, 3-5 pm.

All examinations are open book and open notes. Laptops are not permitted at the examinations.

Grading and Grading Policy

- The course grade will be calculated as 20% homework, 25% for each midterm examination and 30% final examination.
- Your lowest homework score will be dropped. Nonsubmission counts as a zero score.

Calendar

Classes will be held Monday and Wednesday all weeks except the following:

- The week of January 15th—Monday is Martin Luther King, Jr. Day, and no class is scheduled.
- The week of March 5th—Spring Break is March 5-9 and thus the Monday and Wednesday classes will not meet.

The last class day is Wednesday, April 18th. Altogether there are 27 class days (Monday-Wednesday schedule).

Drop and Withdraw Dates

The drop deadline is Friday, 9 February. The withdraw deadline is Friday, 23 March.

Schedule of Topics (tentative)

Week	Topic	Text
8 Jan	Introduction/review	1.4, 3.4, 4.3, 4.6, 5.4, 6.1
15 Jan	Confidence intervals	Chapter 7
22 Jan	Confidence intervals, hypothesis testing	Chapters 7, 8
29 Jan	Hypothesis testing	Chapters 8, 9, 15.1
5 Feb	Hypothesis testing	Chapter 9, 15.2
12 Feb	Simple linear regression	Chapter 12, 13.3
19 Feb	Simple linear regression	Chapter 12, 13.3
26 Feb	Multiple regression	Chapter 13
12 Mar	Multiple regression	Chapter 13
19 Mar	Multiple regression	Chapter 13
26 Mar	Logistic regression	Chapter 13
2 Apr	One-way analysis of variance	Chapter 10
9 Apr	Multifactor analysis of variance	Chapter 11
16 Apr	Categorical data analysis	Chapter 14