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

**Statistics 431**                                        **Spring 2009**

_____

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

**Office hours:**  TTh 3-6 and by appointment

**Classes meet:**  Section 001,  TTh  9-10:20,  in F50 JMHH
                   Section 002,  TTh  10:30-11:50,  in F50 JMHH

**Teaching Assistant and Stat Lab**

**Name:**         Sivan Aldor-Noiman
**Email:**        sivana@wharton.upenn.edu
**Office/Phone:** 432.1 JMHH/215-898-1246

The Department operates a Stat Lab.  Location and hours are posted on webCafe (see next paragraph).

**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. 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:**  You will need to establish a Wharton Computing Account to access webCafe.  The account also allows use of the computing labs in Wharton.  You must have a current Penn ID card to obtain an account.  To get an account, go to

        http://apps.wharton.upenn.edu/accounts/class

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

**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 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*, 7th ed., by J. L. Devore, Brooks/Cole, 2008. Sections 1.4, 3.4, 4.3, 4.6, 5.4 and 6.1, and Chapters 7-13 (some sections from these chapters will be skipped) will be covered. If time permits, we will address some topics in Chapters 14 and 15.

*Business Analysis Using Regression: A Casebook*, by D. Foster, R. Stine, and R. Waterman, Springer-Verlag, 2001 (revised printing). We will cover the first 11 classes.

JMP 7 statistical software. 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 place the software on your computer you will also have installed five manuals and two cards for quick reference, all in pdf format. JMP 7 is installed in Wharton classrooms and is available on machines in the Wharton Computer Labs, located in F75 and F80 JMHH.

JMP 7 may be purchased from estore.e-academy.com. A six-month license costs $29.95 and a twelve-month license sells for $49.95. If you have JMP version 5 or 6, it will suffice.

**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 TA if you have questions. You may also work with and help each other. *Unless otherwise instructed, though, you must submit your own writeup.*

**Examinations**

Two midterm examinations are scheduled:

> Both will be in the evening, 6-8 pm: Thursday, 19 February; Thursday, 2 April.

The final examination will be Thursday, 7 May, 6-8 pm.

In-class quizzes will be held approximately every other week. Each quiz will last ten minutes and will begin at the start of class. There will be six quizzes during the semester. Each quiz will be announced in advance in class, and a reminder will be posted on webCafe.

All examinations and quizzes are open book and open notes. At the examinations and quizzes laptops are not permitted and cell phones should be turned off.

**Grading and Grading Policy**

- Your course grade will be calculated as 20% homework, 20% quizzes, 30% midterm examination score, and 30% final examination.
- Your lowest homework score will be dropped. Nonsubmission counts as a zero score.
- Your lowest quiz score will be dropped. A missing quiz counts as a zero score.
- The midterm examination score is the greater of your two individual midterm examination scores. A missing midterm examination counts as a zero score. There will be no makeup midterm examinations.

**Calendar**

Classes will be held Tuesday and Thursday all weeks except the following:

- The week of January 12th—The first class is Thursday, January 15th.
- The week of March 9th—Spring Break is March 9-13.
- The week of April 27th—The last class is Tuesday, April 28th.

Altogether there are 28 class days (Tuesday–Thursday schedule).

**Drop and Withdraw Dates**

The drop deadline is Friday, 20 February.  The withdraw deadline is Friday, 3 April.


**Schedule of Topics (tentative)**

| Week | Topic | Devore (D)/FSW |
|------|-------|----------------|
| 12 Jan | Introduction/review | D 1.4, 3.4, 4.3 |
| 19 Jan | Introduction/review<br>Confidence intervals | D 4.6, 5.4, 6.1<br>D Chapter 7 |
| 26 Jan | Confidence intervals, hypothesis testing | D Chapters 7, 8 |
| 2 Feb | Hypothesis testing | D Chapter 8, 15.1 |
| 9 Feb | Hypothesis testing | D Chapter 9, 15.2 |
| 16 Feb | Simple linear regression | FSW Classes 1, 2<br>D Chapter 12, 13.3 |
| 23 Feb | Simple linear regression | FSW Classes 2, 3<br>D Chapter 12, 13.3 |
| 2 Mar | Multiple regression | FSW Class 4<br>D Chapter 13 |
| 16 Mar | Multiple regression | FSW Classes 5, 6<br>D Chapter 13 |
| 23 Mar | Multiple regression | FSW Classes 6, 7<br>D Chapter 13 |
| 30 Mar | Multiple regression | FSW Classes 7, 8<br>D Chapter 13 |
| 6 Apr | Analysis of variance | FSW Classes 9, 10<br>D Chapters 10, 11 |
| 13 Apr | Analysis of variance, logistic regression | FSW Classes 10, 11<br>D Chapter 11, 13.5 |
| 20 Apr | Logistic regression | FSW Class 11<br>D 13.5 |
| 27 Apr | Categorical data analysis | D Chapter 14 |