Course: STAT W4700
Title: Probability and Statistics
Semester: Fall 2013

1 Course Description

Objective: Statistics is the language in which data is analyzed and interpreted, and thus any serious data scientist must have a firm understanding of the mathematical principles of probability and statistics. Assiduous students of this course will build this critical foundation.

Details: This course is a self-contained introduction to probability and statistics with a focus on data science. The topics covered include fundamentals of probability theory and statistical inference, including: probabilistic models, random variables, useful distributions, expectations, the law of large numbers, the central limit theorem, point and confidence interval estimation, maximum likelihood methods, hypothesis tests, and linear regression (as time permits).

Prerequisites: Enrollment in the IDSE certificate program (see `idse.columbia.edu`); a background in calculus up to multivariate integration; a background in basic optimization up to the use of Lagrange multipliers; a background in linear algebra up to matrix multiplication and eigenvalues. An ungraded quiz will be given during the first class to allow students to self-assess.

Note: This syllabus will change throughout the term.

2 Course Logistics

Time: MW, 6:10pm-7:25pm
Location: TBA
Instructor:

> John P. Cunningham
> Email: `jpc2181@columbia.edu`
> Office: Department of Statistics 1026, 1255 Amsterdam Ave

Teaching Assistant:

> Evan Blumgart
> Email: `eib2114@columbia.edu`
> Office: TBA
> Office Hours: TBA

Textbook (Required):

> *Probability and Statistics*, Fourth Edition
> M. H. DeGroot and M. J. Schervish
> Publisher: Pearson (2012)
> ISBN: (ISBN-10) 0321500466; (ISBN-13) 978-0321500465
> Note: This textbook is available at Book Culture.
> Note: Many homework problems will be drawn from this text, so the correct edition is required.

3 Grading and Academic Integrity

We take the honor code very seriously; students caught cheating or otherwise in violation will face disciplinary action. Please note the Barnard honor code text:

> "We... resolve to uphold the honor of the College by refraining from every form of dishonesty in our academic life. We consider it dishonest to ask for, give, or receive help in examinations or quizzes, to use any papers or books not authorized by the instructor in examinations, or to present oral work or written work which is not entirely our own, unless otherwise approved by the instructor.... We pledge to do all that is in our power to create a spirit of honesty and honor for its own sake."
> http://barnard.edu/node/2875
> https://www.college.columbia.edu/academics/academicintegrity

Grading: Grading will be on a curve and weighted according to:

Weekly Homework (drop lowest 2)	35%
Midterm Exam	30%
Final Exam	35%

Regrading Policy: Except in the case of obvious grader error, students are discouraged from requesting regrades of homeworks or exams. In the event of a request, please be advised that the entire homework or exam will be regraded, with the likely result of a lower score after the regrade.

Homework: Students are encouraged to work together, but homework write-ups must be done individually and must be entirely the author's own work. Homework is due at the **beginning** of each Monday class. **Late homework will not be accepted under any circumstances.** To receive full credit, students must thoroughly explain how they arrived at their solutions.

Midterm and Final Exams: Students may bring one sheet of 8.5×11in paper, on which the student can write notes (front and back). No other materials are allowed (including textbooks, computers, and other electronics). To receive full credit, students must thoroughly explain how they arrived at their solutions.

Electronics: Taking notes on an electronic device is optional, but students using these devices for unrelated or distracting purposes will lose this privilege.

4 Schedule and Reading

Week	Date	Content (subject to change)
1	9/3 - 9/8	• Topics: introduction, sample spaces, axioms, combinatorics • Reading: Ch 1.1-1.11 (skim: 1.1-1.2; skip ❖ in 1.4 only) • Assignment: HW01 out: **due 9/9**. • Problems: TBA
		continued on next page...

Week	Date	Content (subject to change)
2	9/9 - 9/15	• Topics: conditional probability, independence, Bayes rule, random variables • Reading: Ch 2.1-2.4; Ch 3.1 • Assignment: HW01 due 9/9; HW02 out: **due 9/16**. • Problems: TBA
3	9/16 - 9/22	• Topics: random variables, distributions, cdf's, marginals, conditionals • Reading: Ch 3.1-3.6 (intentional repetition of 3.1) • Assignment: HW02 due 9/16; HW03 out: **due 9/23**. • Problems: TBA
4	9/23 - 9/29	• Topics: multivariate distributions, functions of a random variable, Markov chains • Reading: Ch 3.7-3.10 (particularly 3.7) • Assignment: HW03 due 9/23; HW04 out: **due 9/30**. • Problems: TBA
5	9/30 - 10/6	• Topics: expectation, moments, covariance • Reading: Ch 4.1-4.7 • Assignment: HW04 due 9/30; HW05 out: **due 10/7**. • Problems: TBA
6	10/7 - 10/13	• Topics: popular and useful distributions • Reading: Ch 5.1-5.10 (particularly 5.10; skip ❖ in 5.4) • Assignment: HW05 due 10/7; HW06 out: **due 10/14**. • Problems: TBA (will include midterm review questions)
7	10/14 - 10/20	• Topics: normal distribution revisited and expanded • Reading: Ch 5.6; Ch 5.10; handouts • Assignment: HW06 due 10/14; HW07 out: **due 10/21**. • Problems: TBA (will be shortened due to midterm) • Note: **MIDTERM EXAM** Wednesday Oct 16. • Note: Monday 10/14 class TBD (Columbus Day)
8	10/21 - 10/27	• Topics: law of large numbers, central limit theorem • Reading: Ch 6.1-6.4 (skip all ❖ in 6.2 and 6.3) • Assignment: HW07 due 10/21; HW08 out: **due 10/28**. • Problems: TBA
9	10/28 - 11/3	• Topics: inference, conjugacy, maximum likelihood • Reading: Ch 7.1-7.7 (skip all ❖ in 7.6) • Assignment: HW08 due 10/28; HW09 out: **due 11/04**. • Problems: TBA

Week	Date	Content (subject to change)
10	11/4 - 11/10	• Topics: sampling distributions, Fisher information • Reading: Ch 8.1-8.8 (skip ❖ in 8.8) • Assignment: HW09 due 11/04; HW10 out: **due 11/11**. • Problems: TBA • Note: Monday 11/4 class TBD (election night)
11	11/11 - 11/17	• Topics: hypothesis testing 1 • Reading: Ch 9.1-9.2; 9.5 • Assignment: HW10 due 11/11; HW11 out: **due 11/18**. • Problems: TBA
12	11/18 - 11/24	• Topics: hypothesis testing 2 • Reading: ; Ch 9.6-9.8 • Assignment: HW11 due 11/18; HW12 out: **due 11/25**. • Problems: TBA
13	11/25 - 12/1	• Topics: least squares regression, inference, GLM, ANOVA • Reading: Ch 11.1 - 11.6 • Assignment: HW12 due 11/25; HW13 out: **due 12/2**. • Problems: TBA • Note: Wednesday 11/27 class TBD (Thanksgiving)
14	12/2 - 12/8	• Topics: sampling, simulation, bootstrap • Reading: Ch 12.1-12.6 (skip ❖ in 12.4) • Assignment: HW13 due 12/2; HW14 out: **due 12/9**. • Problems: TBA
15	12/9 - 12/20	• Topics: final review • Reading: review all readings • **FINAL EXAM** TBA (Dec 13-20) • Assignment: HW14 due 12/9.