

# PS350A: PROBABILITY AND MATHEMATICAL STATISTICS

## DESCRIPTION

Introduction to probability and statistical inference, with applications to social sciences and public policy.

## TIMES AND LOCATIONS

Lectures: MW 1:15-3:05 in ART4 (NOTE: room changed back to original)

Section I (computational): T 12:00-1:00 in Building 160, Room 315

Section II (analytical): Th 12:00-1:00 in Building 160, Room 315

## CONTACT INFORMATION

[Jonathan Wand](#), Assistant Professor

Office: Encina Hall West, Room 308

Office hours: W 3-4, Th 3-4, and by appointment

Email: [wand\(at\)stanford.edu](mailto:wand@stanford.edu)

Sarah Anzia, TA

Office: Encina Hall West, Room 313

Office hours: TBA

Email: [sanzia\(at\)stanford.edu](mailto:sanzia@stanford.edu)

Saul Jackman, TA

Office: Encina Hall West, Room 317

Office hours: TBA

Email: [sjackman\(at\)stanford.edu](mailto:sjackman@stanford.edu)

## COMPUTING

Computational homework assignments will develop familiarity with a software package for statistical computation and plotting. All computation will be presented and discussed in terms of **R**; software and related materials can be freely found at <http://www.r-project.org>.

## TEXTS AND READINGS

We will rely on two main texts, which will be referred to in the outline by **DS** or **V**. Additional supplementary resources are listed at the end of the syllabus.

**DS** DeGroot, M.H. and M.J. Schervish. 2002. *Probability and Statistics*. Addison Wesley, Boston, MA, 3rd edition

**V** Verzani, J. 2005. *Using R for Introductory Statistics*. Chapman & Hall/CRC

For those who have had a prior treatment of this material, the parallel readings are provided from:

**CB** Casella, George and Roger L. Berger. 2001. *Statistical Inference*. Belmont, CA: Duxbury Press, 2nd edition

# COURSE OUTLINE

Required readings are marked with ●; optional reading are marked with \*.

## 1. INTRODUCTION AND OVERVIEW (September 22, M)

*I will present examples of research from different fields to illustrate types of questions and topics which are related to this course. Technical details will be considered during the remainder of the course. Emphasis on inference based on experiments versus observational data.*

*Reading:*

- Freedman, Pisani, and Purves, Statistics. Chapters 1-2. (on coursework)

## 2. PROBABILITY: AXIOMATIC FOUNDATIONS AND SAMPLE SPACES (September 24, W)

*Key concepts of probability are introduced; set theory, axioms, sample spaces; counting*

*Reading:*

- **DS** 1.1-1.6, 1.7-1.8, 1.10
- **V** 1
- \* **CB** 1.1-1.2

## 3. PROBABILITY: ADDITIONAL KEY CONCEPTS (September 29, M)

*Concepts of conditional probability and independence are introduced; key results of Law of Total Probability and Bayes Theorem emphasized.*

*Reading:*

- **DS** 2.1 - 2.3
- \* **CB** 1.3

## 4. RANDOM VARIABLES: UNIVARIATE DISTRIBUTION FUNCTIONS (October 1, W)

*Properties of functions of sample space are considered. Discrete, continuous random variables; quantiles of RV.*

*Reading:*

- **DS** 3.1 - 3.3
- **V** 2, 6.1-6.2, 6.4-6.5
- \* **CB** 1.4-1.6

5. RANDOM VARIABLES: EXPECTATIONS (October 6, M)

*Expectations: functions of expectations, expectations of functions. Mean, median, and loss functions. Transformations. Variance.*

*Reading:*

- **DS** 3.8, 4.1–4.3, 4.5, 4.4
- \* **CB** 2.1-2.2, 2.3

6. SOME NAMED UNIVARIATE DISTRIBUTIONS (October 8, W)

*Lets collect the ideas so far and apply what we know to some named distributions—many of which we have seen before—as well as think about the relevance/application of these distributions.*

*Reading:*

- **DS** 5.1-5.6, 5.9-5.10, 11.1-11.2,
- \* **CB** 3.1-3.3, 3.4

7. RANDOM VARIABLES: MULTIVARIATE DISTRIBUTIONS FUNCTIONS (Oct 13, M)

*Characterizing a relationship between multiple random variables. Special case of Independence. Conditional Distributions. Covariance and Correlation.*

*Reading:*

- **DS** 3.4 - 3.7, 3.9, 4.6-4.7, 5.11-5.12
- **V** 3.1-3.3, 3.4.3, 4
- \* **CB** 4

8. RANDOM VARIABLES AND RANDOM SAMPLING (October 15, W)

*Conclusion of comments on multivariate distributions (from previous lecture), and then proceeding into the properties of random sampling (continued in next lecture).*

*Reading:*

- **DS** 4.8, 5.7, 4.4
- **V** 5, 6.3
- \* **CB** 4.7, 5

9. RANDOM SAMPLING (October 20, M)

*Sampling, LLN. Order statistics, CLT*

*Reading:*

- **DS** 4.8, 5.7-5.8, 4.4
- **V** 5, 6.3
- \* **CB** 5

10. ADDITIONAL TOPICS, REVISITING IDEAS, AND REFINEMENTS (October 22, W)

11. MIDTERM (October 27, M)

12. SUMMARY STATISTICS AND SAMPLING DISTRIBUTIONS (October 29, W)

*Method of moments estimation.  $t$  distribution. Confidence intervals*

*Reading:*

- **DS** 7.1, 7.3-7.5
- **V** 7
- \* **CB** 5

13. TESTING HYPOTHESES (November 3, M)

*How to formulate and test a theory. How well can you discriminate from alternative theories. How likely are you to make different types of errors. Tests of means of normally distributed variables. One sample and two sample tests. Omit (for now) subsections on LRT.*

*Reading:*

- **DS** 8.1, 8.5 - 8.7, 8.9
- **V** 8
- \* **CB** 8

14. POINT ESTIMATION, ML (November 5, W)

*Methods are considered for estimating unknown parameters for a given distribution. Emphasis on maximum likelihood (ML).*

*Reading:*

- **DS** 6.1, 6.2, 6.5
- **V** Appendix E
- \* **CB** 6.3, 7.2

15. PROPERTIES OF ESTIMATORS (November 10, M)

*General properties are considered, with applications to ML. Fisher Information, Asymptotic results for ML.  $\chi^2$  distribution*

*Reading:*

- **DS** 6.6, 7.2, 7.7-7.8
- \* **CB** 10

16. INTERVAL ESTIMATION AND HYPOTHESIS TESTING, ML (November 12, W)

*Classical tests, LRT, Wald. Inverting a test statistic. Size and coverage.*

*Reading:*

- **DS** 8.5-8.7
- \* **CB** 8-9

17. CONTINGENCY TABLES AND GOODNESS-OF-FIT (November 17, M)

*Reading:*

- **DS** 9.1-9.5

- **V** 9
  - Jasjeet S. Sekhon. "Quality Meets Quantity: Case Studies, Conditional Probability and Counterfactuals." *Perspectives on Politics*, June: 281-293. 2004.
18. NONPARAMETRIC STATISTICS (November 19, W)
- Komolgorov-Smirnov. Contamination and Breakdown. Sign and rank tests.*
- Reading:*
- **DS** 9.6-9.8
19. APPLICATIONS (December 1, M)
- Experiments vs observational work. Extrapolation, Selection, Mixing processes. Matching. Triangulation.*
- Reading:*
- Manski. Chapters 1-3.
  - Rosenbaum, PR and DB Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American statistician*, 39(1):33-38
  - Gerber, Alan S. and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3): 653-63.
  - Wand, Jonathan, Kenneth Shotts, Jasjeet S. Sekhon, Walter R. Mebane, Jr., Michael Herron, and Henry E. Brady. 2001. The Butterfly Did It: The Aberrant Vote for Buchanan in Palm Beach County, Florida. *American Political Science Review*, 95:793-810
20. APPLICATIONS, CONT'D (December 3, W)

## SUPPLEMENTARY READING MATERIAL

The following books are not required, but they may be of interest during the course. It is often very useful to read the same material covered by a variety of authors. Within each section, books are approximately ordered by increasing sophistication.

### COMPUTER AND R BOOKS

- <http://cran.r-project.org/other-docs.html>
- <http://cran.r-project.org/manuals.html>
- Krause, Andreas and Melvin Olson. 2002. *The Basics of S-PLUS*. 3rd ed. New York: Springer-Verlag. ISBN: 0387954562.
- Spector, Phil. 1995. *An Introduction to S and S-Plus*. Wadsworth Publishing Company. ISBN: 053419866X.
- Venables, W.N and Brian D. Ripley. 2002. *Modern Applied Statistics with S*. New York: Springer-Verlag. ISBN: 0387954570

### PROBABILITY AND STATISTICS

- Freedman, David, Robert Pisani, and Roger Purves. 1997. *Statistics*. 3rd edition. Norton, W.W. & Company.
- Wonnacott, Thomas H. and Roland J. Wonnacott. 1990. *Introductory Statistics*. 5th ed. Wiley, John & Sons. ISBN: 0471615188.
- Richard J. Larsen and Morris L. Marx. 2001. *An Introduction to Mathematical Statistics and Its Applications*, 3rd ed.
- Feller, William. 1961. *An Introduction to Probability Theory and Its Applications*, volume 1. New York: Wiley, 2 edition