

## **Linguistics 330: Spring 2008**

### **Inductive Statistics**

Instructor:	Matt Goldrick	TA:	Celina Troutman
Email:	goldrick@ling.northwestern.edu	Email:	c-troutman@northwestern.edu
Office Hours:	TTh 11-12:30 and by appointment	Office hours:	M 12:30-2:30
Office:	2016 Sheridan Rd. Room 22	Office:	Swift 007

Class time: TTh 2-3:20  
Class location: Library Mac/PC lab

#### What's this course all about?

Learners—everyday people as well as scientists—make generalizations on the basis of incomplete data. Children learn to speak an infinite range of utterances based on a finite sample of sentences. Scientists make generalizations regarding people based on detailed observations of small groups participating in experiments. What is the logic that allows learners to draw such inferences? We'll pursue this goal by answering two interrelated questions:

1. How can one determine the expected behavioral reflexes of some underlying physical (cognitive) process?
2. How can one assess whether observed data conform to the predicted behavioral reflexes of some process?

#### What am I going to do in this class?

We'll work towards these goals by:

1. Becoming familiar with the R statistical package. This is a powerful tool for visualizing and analyzing data. It's also an open-source, platform independent program, free to all users.
2. Reviewing the theory and application of statistical methods. Using R, readings and lectures, we will examine methods that allow us to generate the behavioral reflexes of some underlying processes (both within an individual and across an entire population) and assess how well these predictions accord with observed data.
3. Exploring real data sets. Laboratory analyses will draw on real data sets from language structure at various level (from sound to meaning), music, and clinical populations.
4. Using statistical inference to tackle research problems in language and music systems. Using R, reading and lectures we'll examine the theory and application of these methods to problems across linguistic and other communication systems. In a final project, you'll examine this connection with respect to either: (a) a method not covered in the rest of the course; or (b) data of your own choosing.

Prerequisites: One course at the 200 level in Linguistics, CSD, or Music Cognition, or the equivalent, or permission.

Webpage: Blackboard. You should be able to access it by going to [courses.northwestern.edu](http://courses.northwestern.edu). Let me know if you have trouble.

### Assessments: What You Need to Do

<u>Assessment Type</u>	<u>% of Final Grade</u>
Participation in and preparation for class meetings	30%
Labs (5)	60%
Final paper	
Presentation	2%
Paper	8%

1. Participation in and preparation for class meetings.
  - a. Participation will be assessed through attendance and ability to meaningfully participate in discussion. This doesn't mean you have to know everything before you come to class. Come prepared with clarification questions!
  - b. Preparation will be assessed by
    - i. Completion of readings and assignments from the textbook.
    - ii. Updating the course Wiki (a practical statistics manual that we will build up over the course of the term).
2. Labs. These will labs will be assigned throughout the quarter. You will have about 1 week (sometimes more) to complete each assignment. Due dates for the 5 labs are listed on the syllabus.
3. Final project. In a few weeks, you'll decide on how you want to extend the inferential logic we've learned in the course to other issues. There are two basic options:
  - a. As part of a group you will do background reading on a statistical method not covered in the course. You'll present your review of this work in one of the last two class sessions and write it up for a final project.
  - b. As part of a group you'll apply the techniques you've learned to data from outside the course.

### Course Policies

1. You have to do your work on time. Due dates are firm; attendance in class is not optional. If this is a problem, see #2. The highest possible letter grade for work handed in during the first 24 hours after class will be a C. I will give comments and suggestions on work turned in later than this, but you will receive no credit for the assignment.
2. If you have a problem, give me sufficient time to help you! If you've tried to do an assignment, but can't seem to complete it, come to me well before it's due. If you are unable to come to a class, try to let me know beforehand, or as soon as possible after class. Your reasons for class absences need to be verifiable.
3. You should work together, but the finished product must be your own. Working together is a big part of our in-class work; we hope this will extend outside of the classroom. However, for most assignments, your written work must be your own.
4. Show your work. Whenever you hand in an assignment, don't ever answer a question with a simple "yes" or "no." *Explain* how you got your answer.

## Tentative Course Plan

Readings downloadable from Blackboard

Primary text:

**V:** Vasishth, S. The foundations of statistics: A simulation-based approach

Supplemental:

**B:** Baayen, R. H. Analyzing linguistic data: A practical introduction to statistics using R

Date	Topic	Reading	Assignments
4/1	Introduction: Distributional signatures		
4/3	Basic R syntax	V1 (B 1, 2.1)	
4/8	Visualization		Lab 1 assigned: Distributions in speech and music
4/10	Generating behavior: Simple mechanisms		Lab 1 due
4/15	Generating behavior: More complex mechanisms	V2	Lab 2 assigned: Generating distributions
4/17	Generating distributions over a population		
4/22	Evaluating distributions over single and multiple samples		Lab 2 due.
4/24			
4/29			Lab 3 assigned: Assessing distributions
5/1	Parametric methods for assessing distributions	V3	
5/6		V4	Lab 3 due.
5/8			Final project discussion.
5/13	Non-parametric methods		Final project plan due. Lab 4 assigned: Assessing distributions parametrically
5/15			
5/20	Power		Lab 4 due.
5/22		V8	Lab 5 assigned: Non-parametric methods
5/27	Final project presentations and discussion		
5/29			Lab 5 due.
6/9			5pm: Final paper due