An Introduction to Statistics 571

Bret Hanlon and Bret Larget

Department of Statistics
University of Wisconsin—Madison

September 6, 2011
Welcome

- Two lectures, same notes, same assignments, same exams, *different Brets*.
- If space is available, you may attend either lecture and the discussion section of your choice. *(Indicate on your homeworks which session you plan to attend.)*
- Discussion section will begin the week of September 12.
- First discussion section will include an introduction to R, so bring a laptop if you have one; these discussion sections will be in computer classrooms for those without laptops.
The Scope of the Course

Statistics 571 is:

- an introductory course;
  - we will not cover every method that each of you needs for your own research;
- a graduate course;
  - we expect you to be mature students who take primary ownership of your education;
  - the textbook is our primary reference, but we alter the order of topics and include topics not in the textbook;
  - exam questions may not mimic homework questions; we want to examine your understanding in new contexts.
- a course for students across the biological sciences.
  - we take examples from many biological disciplines;
  - most examples will not be directly related to your own research area.
Web

- See http://www.stat.wisc.edu/courses/st571-larget/ for notes and assignments.
- See https://learnuw.wisc.edu/ for grades.
- See https://wischolar.wisc.edu/stat571/ for discussion/blog.
- Please use the blog and not regular e-mail for content questions so answers can be accessed by all.
- Restrict use of email to instructors for private correspondence.
The Discussion Board/Blog

- We are piloting a new UW-Madison tool, WiScholar, for the course.
- Login at https://wischolar.wisc.edu/stat571/ to gain access and the ability to add posts.
- The blog is:
  - a forum for peer-to-peer discussion;
  - a supplement to face-to-face communication with course instructors and TAs;
  - used well for posting:
    - issues with the course;
    - questions about homework problems hoping for hints to solutions;
    - questions about logistics and policy.
- Do not expect rapid responses from TAs and instructors.
- The blog is a supplement to regular face-to-face meetings during lecture, discussion section, and office hours and is a vehicle to encourage peer-to-peer teaching and learning.
Course Objectives

- Students will gain an understanding of statistical concepts and methods;
- Students will be able to apply what they learn in the course to data analysis and inference problems common in biological science and their own research;
- Students will become competent with a statistical computing package.
The textbook is *The Analysis of Biological Data* by Michael Whitlock and Dolph Schluter.

We will cover most material in the text and some material not in the text, *but the order we cover things will be quite different.*

This is the second year we have used this textbook and our development of the course continues to evolve.

The schedule of topics should be seen as tentative: we may need to make adjustments as we learn about the pace that works well.

Lecture is the primary source of information for the course, but the text ought to be easy for you to read.

We will expect a bit more mathematically and computationally than the textbook assumes.
Typical Lecture

A typical set of lectures on a topic will include:

1. A motivating set of data and biological question;
2. Graphs and numerical data summaries that illuminate the question;
3. A conceptual look at *The Big Picture*;
4. A detailed description of methods;
5. Application and interpretation for the example;
6. Details about R;
7. Flies in the ointment (and other concerns);
8. Extensions and further topics.
Computing

- You may use whichever statistics package you wish for the course.
- We use R ourselves and encourage its use by you.
- We will teach some R and offer support in its use.
- You are on your own with other statistics packages.
- R is free and available on all common platforms.
- R is the standard statistical computing package in graduate programs in statistics.
- There is a steep learning curve, but the effort is worthwhile.
- Teaching effective use of R is a course objective, and we will take time to do it in lecture, homeworks, and discussion.
Homework

- Assignments are typically given on Thursdays, due the following Friday by 4pm in your TA’s mailbox. (There may be exceptions around Thanksgiving and exam weeks.)
- Solutions are typically posted Monday afternoons.
- There will be 12 assignments during the semester.
- You may drop two assignment scores (might be missing due to illness or other commitments) or the lowest two scores.
- You may *collaborate* on homework, but each student must write out their own solution.
- What you turn in must be neat.
- Do not simply turn in computer output; answer questions in plain English and support with computer calculations or graphics when merited.
Grading

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm Exam 1</td>
<td>20%</td>
</tr>
<tr>
<td>Midterm Exam 2</td>
<td>20%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>40%</td>
</tr>
</tbody>
</table>

- Course grades are based on the above percentages.
- There is no fixed grading scale, but indications of a scale will be given with each exam.
Statistical Consulting

- Please be aware that most of you are eligible to use the **free statistical consulting service** offered to people in CALS, L&S Biology, and the IES program to support their research.
- (The Vet School chose not to support the service this year and are no longer eligible to use the service, except for select labs that have made their own arrangements.)
- Consulting is available both at the design stage of a study and for data analysis.
- The statistical consulting service involves making an appointment to talk about your data analysis and to get advice; it is not a service for someone to do your statistical analysis for you.
- Occasionally, a statistical consulting relationship can become a collaborative research effort.
- You may not need this service now, but do not forget about it as you continue your graduate education and research.
- Learn more at this web site: http://www.cals.wisc.edu/calslab/stat.html
Biometry Masters Program

- The Biometry Masters program is a master degree program offered by the Statistics Department aimed at PhD students in the biological sciences on campus.
- The program is distinct from the Masters Program in Statistics.
- Biometry students have co-advisors, one a Statistics Professor in the Biometry group, and biology professor.
- The Biometry masters thesis is often a chapter in the PhD dissertation.
- Prerequisites for the program include a three-semester sequence in calculus and the program includes several advanced statistical courses.
- Students who complete Statistics 571–72 and do well and who are very interested in deeper and more formal statistical training are ideal candidates for the program.
- See http://www.stat.wisc.edu/Department/phd-masters/biometry.html for more details.