S320/520 syllabus

Brad Luen

Summer 2015

1 Instructor

Dr. Brad Luen
Room 200, Stat House, 309 N Park Ave (at 7th)
bradluen at indiana.edu (put S320 or S520 in the subject line)

2 Lectures

The class will take place during the second six-week section (June 22nd to July 30th.) Lectures will be held from 12:40 pm to 2:30 pm Monday to Thursday in Geology Building 143.

3 Description

This course introduces the basic concepts of statistical inference through a careful study of several important procedures. Topics include 1- and 2-sample location problems, the one-way analysis of variance, and simple linear regression. Most assignments involve applying probability models and/or statistical methods to practical situations and/or actual data sets.

S320 is the basic version of this course, intended for undergraduates. It is the gateway to more advanced courses offered by the Department of Statistics. It differs from typical methods courses in the following respects:

- Greater emphasis on why a method works. Many courses explain how, but provide little explanation of why.

- Greater depth, less breadth. Many courses provide superficial coverage of a great many topics; S320 covers fewer topics, but in considerably more detail. Students desiring knowledge of procedures not covered in this course are strongly encouraged to take additional statistics courses. S320 is the gateway to majoring in statistics.

- More math. S320 is not a theoretical course (like STAT S420) and it uses a minimum of sophisticated mathematics. However, S320 does introduce a good deal of mathematical notation and it does assume that students are comfortable plugging numbers into formulas.

- Interactive computing. Rather than use a statistical computer package as a "black box," S320 relies on computer tools that simplify the computational burden but which require the student to understand how the analysis is to be performed.

S520 is an expanded version of S320 that covers additional material (see syllabus below). S520 serves two constituencies:

- Graduate students in quantitative disciplines who are looking for a solid introduction to statistics and who may want to take additional courses in statistics.

- Graduate students pursuing an M.S. in Applied Statistics who desire a more gentle introduction to the fundamental principles of statistical inference than is provided in the more theoretical STAT S620. These students will complete the standard S520 requirements, plus supplementary readings and exercises on more theoretical topics.

If you want a deeper introduction to statistics that will prepare you to take more advanced statistics courses, then take S520.

4 Prerequisites

No previous knowledge of probability is assumed; S320/S520 is recommended for students who wish to take a single, self-contained semester of statistics that emphasizes analyzing data. We will use several basic concepts from calculus; hence, S320/S520 has a prerequisite of MATH M212. In practice, if you remember or can quickly learn what a limit is and how to integrate functions like x^2, then you can succeed in this course. But please note: This tends to be a difficult course for students with limited mathematical preparation.

5 Registration

Students who register for credit are expected to attend class, do homework, and take tests. Students who register to audit are expected to attend class and do homework, but taking tests is optional. You do not need my permission to register for S320/S520; however, someone in the Department of Statistics may check enrollment to verify that each registered student has taken M212. If you have not taken M212, then you should check with me to verify that S320/S520 is a suitable course.

Please check with the lecturer if you want to sit in on the course without enrolling. If the class is full, enrolled students have priority for seats.

6 Office hours

It's easiest to just bug me after class. Email me to make an appointment.

7 Textbook

The required text is Michael Trosset's *An Introduction to Statistical Inference and Its Applications with R*, hardback, 2009. Exercises will be assigned from the book (note that the PDF floating around the Internet does not have the correct exercises). The book's webpage, which includes R functions, data sets, and a list of errata, is at `http://mypage.iu.edu/~mtrosset/StatInfeR.html`.

8   Computer software

We will make selective use of the statistical programming language R. R is free, Open Source software available for Windows, Mac OS X, and Linux. You should download and install R from

<div align="center">

`http://cran.r-project.org/`

</div>

after the first class, or earlier.

I also strongly recommend use of the RStudio environment:

<div align="center">

`http://www.rstudio.com/`

</div>

Information about R is included in the text, particularly in Appendix R. Please see Appendix R for information about how to read data in a text file into your R workspace. Additional R functions will be made available as required. You are encouraged to bring a laptop to class, though this is not essential.

9   Syllabus

The table of contents of the textbook will serve as a syllabus:

<div align="center">

`http://mypage.iu.edu/~mtrosset/StatInfeR/toc.txt`

</div>

The following sections will be optional:

- 3.6 Case Study: Padrolling Dice
- 10.3 The Symmetric 1-Sample Location Problem
- 11.2 The Case of a General Shift Family
- 12.3 Planned Comparisons
- 12.4 Post Hoc Comparisons
- 14.3 Monotonic Association
- 15.5 Assessing Linearity
- 16.2 The Bootstrap
- 16.3 Case Study: Adventure Racing

10   Grading

Homework assignments (25%): There will be approximately nine homeworks, announced in class and posted on Canvas. All students are required to do the exercises labelled 320. Students enrolled in S520 must also do the exercises labelled 520. Students enrolled in S320 may do the additional 520 exercises for extra credit. Collaboration on homework with other students is both permitted and encouraged, but you must write up your own solutions.
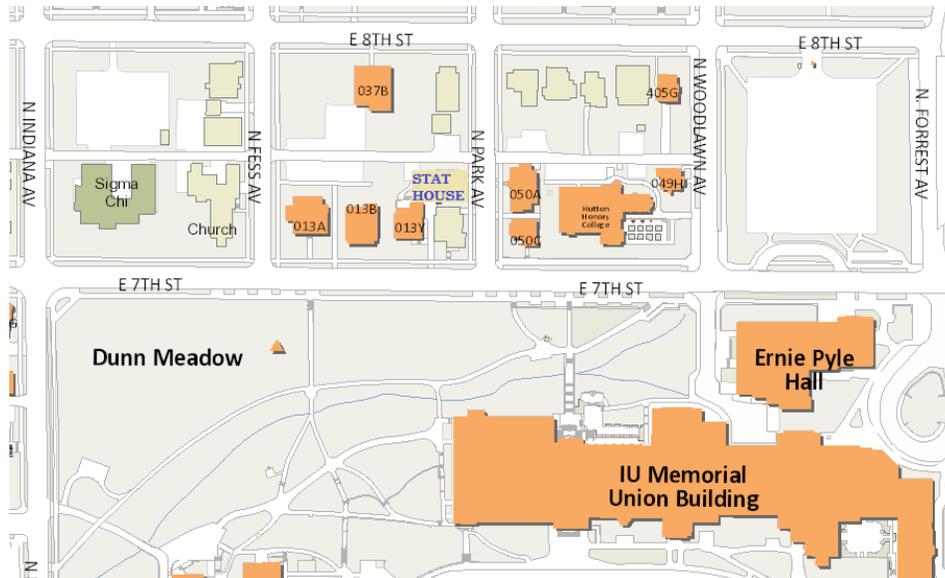
Figure 1: Location of the Stat House.

Homework will be due in class on Mondays and Wednesdays (except on test days.) Late homework will be penalized at 10% of the total possible score per day.

Midterm tests 1 (20%) and 2 (25%): Tests will have two parts, one take-home and one in-class, not of equal value. The take-home part will be open book and open notes; the in-class part will be closed book and closed notes, but a formula sheet will be allowed. Collaboration on tests with other students is not permitted under any circumstance. The in-class part of Midterm 1, covering Chapters 1–6, is provisionally scheduled for Monday 6th July. The in-class part of Midterm 2, covering Chapters 7–11, is provisionally scheduled for Monday 20th July.

Final exam (30%): The final exam will be comprehensive, but with strong emphasis on the material in Chapters 12-16. It will have two parts, one take-home and one in-class. The take-home part will be open book and open notes; the in-class part will be closed book and closed notes, but a formula sheet will be allowed. Collaboration on the final exam with other students is not permitted under any circumstance. The in-class exam will take place during the final class (Thursday 30th July.)
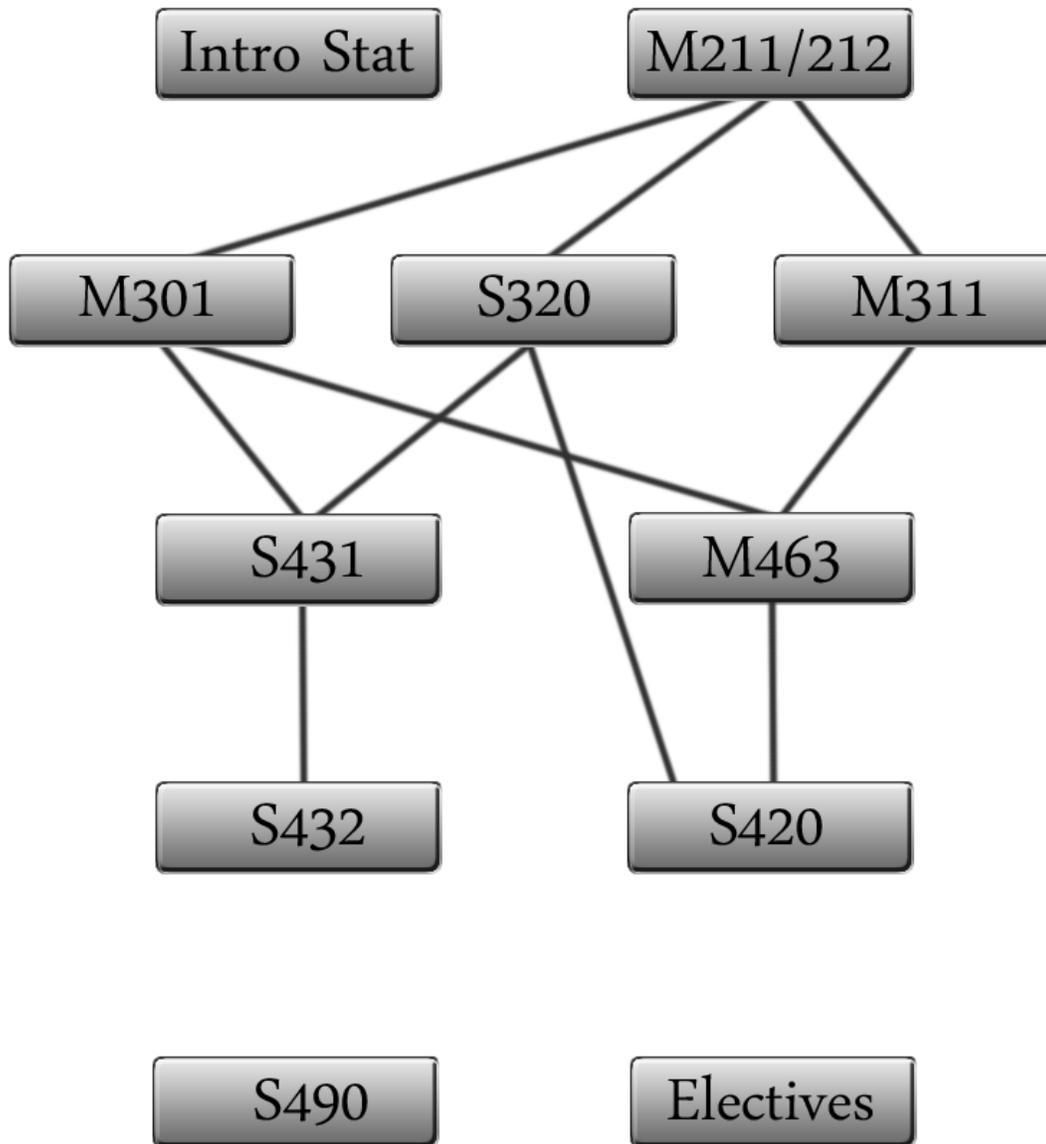
Figure 2: Flowchart of IU undergrad stat courses.