

46-923 STATISTICAL INFERENCE

MINI 2, Fall 2013

Instructor: Rebecca Nugent

Baker Hall 232C, (412) 268-7830, rnugent@stat.cmu.edu

Office Hours: TBA

TAs: Stefa Etchegaray Garcia, Lingxue Zhu

betchega@stat.cmu.edu, lzhu1@andrew.cmu.edu

Office Hours: Wednesday 4pm (Porter 117)

Saturday 11:30am-1pm (Tepper/Posner 145, NY1)

Class Meetings: Mondays, 5:30pm-8:30pm

10/28, 11/4, 11/18, 11/25, 12/2: Pittsburgh (Tepper/Posner 145)

11/11, 12/9: NYC (NY1)

Website: <http://www.cmu.edu/blackboard>

Prerequisites: Calculus up to and including integration of functions of several variables; basic knowledge of matrix algebra; 46-921 Probability

Textbook: *Statistics and Data Analysis for Financial Engineering* by Ruppert (Springer, 2010). This text is required for the course but will be used more extensively in later Statistics courses.

General Course Plan: This is the second part of a two-part course introducing probability and statistics. You now have the tools to describe uncertainty, and a collection of distribution useful for modeling real data sets and approximating the distribution of statistics. If instead, we were to take these actions in the opposite order, we are performing statistical inference. That is, we first observe some repeated realizations of the random variable and then make a decision about the underlying probability model based on this sample. We now focus on the use of data to estimate model parameters and to test postulated probability model assumptions.

Course Objectives/Goals:

1. You will understand the risks of overfitting probability distributions to data sets, recognizing the bias that can result from such mistakes.
2. You will understand the bias/variance tradeoff via the mean squared error.
3. You will be able to construct maximum likelihood estimators using both analytical and numerical approaches.
4. You will be able to construct approximate confidence regions for estimators.
5. You will understand how the bootstrap can be used to approximate standard errors and confidence intervals for a wide range of situations, with minimal assumptions.
6. You will understand the terminology and ideas behind statistical hypothesis testing, including a clear understanding of what a p-value is (and is not).
7. You will be able to implement Bayesian methods of inference in simple situations, and understand the philosophical underpinnings of this approach.

8. You will be able to implement methods of testing the validity of assumptions regarding the distribution of data.
9. You will grow in your understanding and usage of R to conduct statistical inference.

Course Work: Your grade will be determined by homework assignments and the final exam.

- Weekly homework assignments are due at the beginning of class (5:30pm) on Mondays. The type of homework submission (paper vs electronic, etc) depends on the location of the student. Deviation from the requested format requires instructor permission.
 - Pittsburgh students: give a paper copy to Abby
 - NYC students: give a paper copy to Diffy or Seida
 - online students: email a pdf of the HW to rnugent@stat.cmu.edu. **Note that, if homework is scanned, students must use an actual copier/scanner, not a phone.** Phone scans are very difficult to read/open/grade, etc.

Homeworks must be legible; it is not our responsibility to figure out “what you meant”. Answers that are unclear will be graded as such.

Similarly, you should include justification for any steps taken to solve the problem (aside from basic arithmetic and calculus). Not doing so may result in a loss of points. Writing down just the answer will not receive full credit (unless specifically allowed).

You should expect the difficulty level of the homework problems to be greater on average than the problems worked in class. Please attempt to see the TA or instructor during office hours or the TA sessions for help with homework problems. Questions posed by email must be sent at least 24 hours before the time an assignment is due in order to reasonably expect a response. Note though that a response is not guaranteed. Try to ask early.

- The final exam is cumulative and will be Saturday, December 14th, 5:30pm-8:30pm. More details later. In particular, we will discuss what you are allowed to bring and what will be provided (e.g., a formula sheet). You are responsible for any material covered in lecture and homework regardless of whether it appears in the text.

Grading policy: You may and should discuss homework problems with your fellow students, however the work you submit must be your own. Acknowledge any help received on your assignments. **Copied assignments, including the use of solutions from previous years, are a violation of the MSCF Honor Code.**

You have one week from the day an assignment is handed back in class to bring any grading issues, comments, complaints, etc to the attention of the instructor. Please note that if you are absent the day something is handed back, this deadline will not be extended unless arrangements have been made in advance with the instructor.

Final grades will be computed with these weights: Homework Average 50%, Final Exam 50%.

However, note that you must receive a passing grade on the final exam (60%) to pass the course.

Final letter grades will be determined as usual: [90,100] = A- / A / A+, [80,89] = B- / B / B+, [70, 79] = C- / C / C+, [60, 69] = D, [< 60] = R.

Computing: The statistical computing package we will use in this course is R. R is available on many campus computers, and you may download a free version from www.r-project.org.

R References: manuals available on R website;

<http://www.stat.cmu.edu/~rnugent/teaching/introR>

Introductory Statistics with R, Peter Dalgaard; Springer-Verlag

Modern Applied Statistics with S-Plus Venables, Ripley; Springer

Laptop Policy: Students are expected to be participating in class; any laptop use during class should pertain directly to the class. Instructor reserves the right to not allow laptop use during class. When the class has a guest speaker, laptops must be turned off and put away.

Cellphones/Pagers, etc: All cellphones, pagers, beepers, and anything else that makes noise should either be turned off or silenced during class.

Communication: Assignments and class information will be posted on Blackboard. Help with using blackboard is available at www.cmu.edu/blackboard/help/.

Email: Sending email to your professor or teaching assistants should be treated as professional communication. Emails should have an appropriate greeting and ending; students should refrain from using any kind of “shortcuts”, abbreviations, acronyms, slang, etc. in the email text. Emails not meeting these standards may not be answered.

Emails with questions about the material or grading should be directed to the professor, even if it's something as simple as an error in grade tabulation. I may post questions that seem useful and appropriate for the class on Blackboard (Discussion Board) or in an email to the class. Of course the original question will be anonymized.

Academic Integrity: All students are expected to comply with the CMU policy on academic integrity. This policy is online at www.studentaffairs.cmu.edu/acad_integ/acad_int.html

Cheating, copying, etc will not be tolerated; please ask if you are unsure of whether or not your actions are complying with assignment/exam instructions. Always ask if you are unsure; always default to acknowledging any help received.

Video/Audiotaping: No student may record or tape any classroom activity without the express written consent of the professor. If a student believes that he/she is disabled and needs to record or tape classroom activities, he/she should contact the Office of Equal Opportunity Services, Disability Resources to request an appropriate accommodation.

Disability Services: If you have a disability and need special accommodations in this class, please contact the professor. You may also want to contact the Disability Resources office at 8-2013.