

Computational Statistics and Data Analysis

STAT 341

Lecture Time/Room 02:30-03:50TTh / MC 4040
Instructor: Ali Ghodsi
Dept. of Statistics & Actuarial Science
Office: M3 4208
E-mail: aghodsib@uwaterloo.ca
Office hours: 4:00-5:00 T

Textbook

There is no required textbook for the class. Recommended books that cover the similar material are:

- **Elements of Computational Statistics** by James E. Gentle
- **Computational Statistics** by Geof H. Givens, and Jennifer A. Hoeting
- **Simulation** (Statistical Modeling and Decision Science) by Sheldon M. Ross

We will also be covering material similar to a variety of chapters from a few other books which I will point out in class.

Prerequisite

MATH 237 or 247, (STAT 230 with a grade of at least 60% or STAT 240), STAT 231 or 241

Antirequisite

CS 437/STAT 340

Computation

Computing parts of the assignments will be done in Matlab, but prior knowledge of Matlab is not essential.

Communication

All communication should take place using the *Piazza* discussion board.

Piazza is a good way to discuss and ask questions about the course materials, including assignments, in a public forum. It enables you to learn from the questions of others, and to avoid asking questions that have already been asked and answered. It also provides a forum for course personnel to make announcements and clarifications about assignments and other course-related topics. Students are expected to read *Piazza* on a regular basis.

Enrolling in Piazza

You will be sent an invitation to your UW email address. It will include a link to a web page where you may complete the enrollment process.

Piazza Guidelines

Here are some guidelines that you should keep in mind when posting items to Piazza:

1. Please remember that everything you post is public - everyone enrolled in this course will be reading it. As a result, in any posts you make, do not give away any details on how to do any of the assignments. This could be construed as cheating, and you will be responsible as the poster. If you have questions about an assignment that require you give specific details of your solution, you may still post to Piazza, but check *This is a private post - only visible to class instructors* (and TAs). If the instructors and/or TAs feels that posting it to everyone is appropriate, they will do so.
2. Keep posts related to the course, concise, and topical. As students are all expected to read Piazza on a regular basis, try not to waste the time of readers.
3. Please be diligent about attempting to find the answer before you post a question. Piazza includes excellent search facilities – use them! Scan all of the questions that have already been asked. Better yet, read them along with the answers. You'll learn lots! Please do all you can to avoid duplicates.
4. Make it easy for other students to find your question – just in case they have the same question and want to see the answer.
 - Use a meaningful subject heading. "Help" and even "Help for A3Q2" is not very meaningful. "Clarify parameter order for A3Q2" is much better.
 - Tag your post with all the applicable tags. Start a tag by typing the hash character (#). A drop-down list of tags that are currently in use will appear. Use one of them, if applicable. If not, create a new one. However, any tag you create should be applicable to many posts not just yours.
5. Please don't post things to the group that provide no useful information to readers. Posts like "I have the same question as this one just posted", or "I agree with this comment" serve no useful purpose, and waste people's time.
6. Keep complaints about the course out of Piazza or mark them with the *This is a private post - only visible to class instructors* checkbox. If you have a concern about anything to do with the course, the best way to deal with it, and to get results, is to take it to the course instructor. Piazza is not a complaint forum.

Assignments and grades will be handled through *Learn*. Please log on frequently to Piazza and Learn. You are responsible for being aware of all STAT 341 material, information and email messages found on *Learn* and *Piazza* throughout the semester.

Tentative Marking Scheme

30% Assignments (6 assignments)

20% Midterm exam (Thursday Oct 22, 2:30 pm -3:50 pm in MC4040 and CPH3602)

50% Final exam

Note

- A passing mark must be achieved on the final exam in order to receive a pass on the course.
- Final exam will be closed book. Calculators are allowed but to ensure that all students have equal access to calculators, you may only use one of the following three calculators: TI 30X II (S or B both allowed), Casio fx-300MS or Sharp EL-531X, all of which require a pink tie sticker or embossing (Science's blue googles sticker or Engineering's sticker are also allowed).

Please see the following link for details:

<https://uwaterloo.ca/math/current-undergraduates/regulations-and-procedures/calculator-regulation>

- A grade of INC is awarded ONLY to a student who has completed course work during the term well enough that they could reasonably be expected to earn a passing mark in the course, but who was unable to complete the final exam for reasons beyond his or her control. You will receive DNW (not INC) in all of the following cases.
 - Do not write the final exam and have no valid excuse.
 - Do not write the final exam, have a valid excuse but the average of your course work during the term is less than 60%.
 - Do not write both the final and the midterm exam, have a valid excuse for both the final and the midterm exam.
 - Do not write the final exam, have a valid excuse for two or more assignments and for the final exam.
- Marking scheme, dates, and topics are tentative and subject to change.

Note for Students with Disabilities

The Office for Persons with Disabilities (OPD), located in Needles hall, Room 1132, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with the OPD at the beginning of each academic term.

Tentative Topics

- Sampling (Generating random numbers)
- Monte Carlo Simulation
- Markov Chain Monte Carlo (MCMC)
- Introduction to Supervised Statistical Learning
- Introduction to unsupervised Statistical Learning
- Introduction to Feature Extraction Methods
- Discrimination Methods