Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
6.008 Introduction to Inference
Fall 2016

Information Sheet

Lecturers

Prof. Polina Golland	Prof. Gregory W. Wornell	Dr. George H. Chen
polina@mit.edu	gww@mit.edu	georgehc@mit.edu
Office: 32-D470	Office: 36-677	Office: 32-D472

Teaching Assistants
Ben Eysenbach (bce@mit.edu)
Danielle Pace (dfpace@mit.edu)
David Qiu (davidq@mit.edu)

Lab Assistants
Ali Soylemezoglu (alican@mit.edu)
Kevin Yang (yangk@mit.edu)

Administrative Assistant
Tricia O'Donnell (tricia@mit.edu), Office: 36-677, Tel: 253-2297

Lectures:	Monday and Wednesday, 10-11am, Room 32-155
Recitations:	Tuesday and Thursday, 10-11am (3-442) or 1-2pm (34-302) or 2-3pm (34-302)
Lab:	Thursday 3-5pm (38-500) or Friday 10-noon (38-500)
Office Hours:	Days, times and locations posted on web site.

Welcome to 6.008 !

We live in a world in which we constantly need to extract information from data. This is the central problem of inference. And computationally efficient methods for such inference are enabling technologies for an enormous range of applications.

Example domains abound, and include search and retrieval, data mining, computer vision and imaging, voice recognition, communication and compression, natural language processing, robotics and navigation, computational biology and bioinformatics, medical diagnosis, distributed sensing and monitoring, and finance.

Many of the most successful inference algorithms arise out of probabilistic modeling and analysis. If you want to learn the fundamentals of this discipline and see some of what you can do with it, this subject is the place to start. Indeed, it will provide a solid foundation for more advanced subjects that build on this framework of reasoning. As such, the subject is targeted at (and likely to strongly appeal to) students both across and beyond Course 6 (EECS).

Lectures, Recitations, Lab Sessions, and Office Hours

As with other core subjects, 6.008 has lectures and recitations, which are designed to complement each other. Recitations begin the first week of classes. There are three possible recitation times from which you can choose. Select to attend whichever suits your schedule best; if strong imbalances result, we will make adjustments.

6.008 also has scheduled lab hours that while optional are highly recommended and will help you with the laboratory component of the class. Lab hours also start the first week of classes. You are also free to choose either of the two weekly lab sessions to attend.

In addition, there will be several scheduled staff office hours throughout the week, the details of which are provided on the class website. You are welcome and encouraged to come to any and all of them you think might be helpful to you in digesting the material.

Prerequisites

The official prerequisite is calculus. Effectively, we are assuming you are comfortable with (and fluent in) basic mathematical analysis to the level of, e.g., 18.02. In addition, we assume you are familiar with the basics of the python programming language, but will provide tutorials the first week of the semester to help you come up to speed on the aspects of python that will be needed in the laboratory component of the subject. Finally, we emphasize that no prior exposure to probability is required, as we develop the necessary foundations in probability as part of this class.

Degree Requirements

For EECS students, the requirements satisfied by 6.008 will change in the future as the department implements the new undergraduate curriculum. The requirements specified in this section are true for the Fall 2016 offering of 6.008.

EECS students in 6-2 program may use 6.008 as one of their EE or CS foundation subjects. All EECS students may use 6.008 as one of their math elective or free elective subjects.

All EECS students may petition to use 6.008 as an introductory subject instead of 6.01/6.02/6.03/etc. Note that the petition must be submitted and approved (by the lecturers) before the Fall 2016 add date.

All EECS students may petition to take 6.008 instead of 6.042 as one of their math elective subjects and use it as a prerequisite for more advanced subjects that require 6.042. Note that the petition must be submitted and approved (by the lecturers) before the Fall 2016 add date. In approving this petition, we need to verify that you either have sufficient background in constructing formal proofs (an essential component of 6.042 that more advanced subjects rely on) or that you are committed to acquiring such background on your own.

M.Eng. students may use 6.008 as one of their restricted elective subjects. As always, no double counting is allowed.

Non-EECS students should consult their departments about what degree requirements 6.008 can satisfy for them.

If you have additional questions about degree requirements, please contact the lecturers. We will be happy to meet with you and discuss.

Reading

There is no existing text that matches the content of this relatively new subject and the style in which we teach it. However, we have been actively developing detailed course notes for the class, which we will distribute in parts throughout the term. In this development stage, the notes will be necessarily rough in places and contain bugs, which we will count on you to help us catch.

In addition, you may find the following text a useful auxiliary resource for the subject:

> D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*, 2nd ed., Athena Scientific, 2008.

If having your own copy would be handy, it may be purchased through, e.g., Amazon. We will also place this book on reserve at the MIT libraries (Barker). Other possibly useful readings will be posted on the course web site as appropriate.

We should emphasize that the course notes do *not* replace the lectures and recitations. The notes will be necessarily incomplete, and most of the discussion, interpretation, and insights will take place in class. Conversely, lectures and recitations are not a substitute for the notes. Indeed, we will defer many details of our development to the reading materials. Thus, you should view class and the written materials as strongly complementary.

Problem Sets and Computational Labs

There will be 11 problem sets (1-11) and 5 computational labs (I-V). The problem sets emphasize analysis aspects of the material, while the computational labs involve a programming component, based on python (version 3), to help you explore computational aspects of the material.

Problem sets and computational labs will be concurrent (and jointly designed), so that at any time you will be working on one problem set and one computational lab. It will be important to manage your time carefully, and allow enough time for them.

Problem sets will be due by 10am on Wednesday, in lecture. Computational labs will be due by 5pm on Fridays and are to be submitted electronically. Solutions will be available on the course web site immediately after the due dates and times, and thus we cannot accept late submissions for grading. To help with your planning, due dates are listed on the schedule at the end of this handout.

While you should do all the assigned problems and computational exercises, only a subset will actually be graded. Also, you will find a set of "practice" problems in every problem set. These are not required, but you may find it helpful to work through them if you are looking for more practice working with the concepts introduced in class.

Don't be misled by the relatively few points assigned to homework in the final grade calculation. While the grade you get on your homework is only a minor component of your final grade, working through (and, yes, often struggling with!) the homework is a crucial part of the learning process and will invariably have a major impact on your understanding of the material.

In undertaking the problem sets and computational labs, moderate collaboration in the form of joint problem solving with one or two classmates is permitted provided your writeup (and all code) is your own and that you identify your collaborators in your writeup.

Exams

There will be two evening quizzes and a final exam in the subject. Dates for the quizzes are Wednesday, October 5, 7:30-9:30pm (conflict date: October 6, 9-11am), and Wednesday, November 9, 7:30-9:30pm (conflict date: November 10, 9-11am). Both quizzes will be held in 32-124. Please let us know at least three weeks in advance if you have a conflict with the quiz time and need to make use of the conflict date. Note also that because of the evening quizzes, the corresponding Wednesday, October 5 and Wednesday, November 9 lectures are canceled.

The final exam will be during Finals Week, the exact date of which we will announce when scheduled by the Registrar.

The quizzes and final exam will all be *closed book*. You will be allowed to bring *one* 8.5×11-inch sheet of notes (both sides) to Quiz 1, *two* 8.5×11-inch sheets of notes (both sides) to Quiz 2, and *three* 8.5×11-inch sheets of notes (both sides) to the Final Exam.

Course Grade

The final grade in the course is based upon our best assessment of your understanding of the material during the semester. Roughly, the weights used in grade assignment will be:

Quiz 1	20%
Quiz 2	20%
Final Exam	30%
Problem Sets	10%
Computational Labs	15%
Classroom participation	5%

with the additional property that if you do better on the Final Exam than either quiz, and you have done all the problem sets, then the corresponding quiz will not count, i.e., the quizzes can only help you if you are doing all the homework. Also, please note that completing all the computational labs is a subject requirement.

Finally, as always, other factors such as contributions to the discussion in class and other interactions can make a difference in your final grade.

On-line Material Development

The staff has been actively producing on-line resources for 6.008, to be used in both the residential and edX versions of the subject. While we will not be relying on this material in the residential version of the class this semester, we are looking for volunteers from the class to help us debug and refine these new resources.

Those who volunteer will be expected to work through the new on-line material we have developed so far (which correspond to roughly the first half of the subject), and the associated on-line problems and homework assignments, and provide us with feedback in a special weekly office hour with the staff. By volunteering, you can earn an additional 10% toward your grade, based on your on-line homework grade and engagement (bringing the overall total to 110%). This opportunity will appeal most to students who relish the challenge of a new experience and playing a role in helping shape the future of 6.008. If this is something you'd be interested in being a part of, let us know!

Course Web Site, Email, and Piazza

We will make announcements via email, and we will post information and handouts on the course web site.

You should first make sure that you have an active Athena account (by visiting `http://ist.mit.edu/support/accounts/` if necessary) as well as a personal certificate (by visiting `https://ca.mit.edu/ca/` if necessary). If you have problems or if you are not a regular MIT student, please contact one of the TAs for assistance.

The course web site is:

`http://web.mit.edu/6.008`

You will need to have a valid certificate *and* be on the official course list to access the web site. If you have pre-registered for 6.008, this should have already been set up. You should shortly receive an invitation to join the site. Otherwise, contact one of the TAs. If you can access content on the web site, you should also be receiving all of the course announcements.

If you have any questions during the term, you can reach us by sending email to

`6.008-staff@mit.edu`

In addition, we will also use piazza as an on-line discussion forum. Through piazza you can post questions or comments about any aspects of the material, which your peers and the staff can respond to. It is great way to clarify potential points of confusion in course notes, problem sets, recitation handouts, computational labs, etc. You can access the piazza forum through the link provided on the class web site. Once you enroll yourself, you can fully participate in the discussions.

Tentative Syllabus and Schedule

	Date	Topic	HW out	HW due
W	9/7	L1: Introduction and overview	1 (L1), I	
M	9/12	L2: Probability of events		
W	9/14	L3: Discrete random variables	2 (L2,3)	1
F	9/16			I
M	9/19	L4: Joint distributions, marginals, conditionals, Bayes		
W	9/21	L5: Marginal and conditional independence structure	3 (L4,5), II	2
F	9/23	*Student holiday – no class*		
M	9/26	L6: Measures of randomness		
W	9/28	L7: Decision-making, most probable configurations, MAP rule	4 (L6,7)	3
M	10/3	L8: Graphical models, message-passing, hidden Markov models		
		Optional Quiz Review, 7:30-9:30pm, 6-120		
W	10/5	**Quiz 1** (through L7), 7:30-9:30pm, 32-141 *no class*		
F	10/7	*Add Date*		II
M	10/10	*Columbus Day – no class*		
W	10/12	L9: HMM marginalization: forward-backward algorithm	5 (L8,9)	4
M	10/17	L10: HMM most probable configuration: Viterbi algorithm	III	
W	10/19	L11: Parameter estimation, Maximum Likelihood method	6 (L10,11)	5
M	10/24	L12: Model learning; Naïve Bayes Models and HMMs		
W	10/26	L13: Learning structure in graphical models	7 (L12,13)	6
M	10/31	L14: Markov/Chebyshev bounds, law of large numbers		
W	11/2	L15: Typical sets	8 (L14,15)	7
F	11/4		IV	III
M	11/7	L16: Joint typicality		
		Optional Quiz Review, 7:30-9:30pm, 6-120		
W	11/9	**Quiz 2** (through L15), 7:30-9:30pm, 32-141 *no class*		
F	11/11	*Veterans Day — no class*		
M	11/14	L17: Atypical sequences, large deviations and max-entropy		
W	11/16	L18: Markov chains and random walks	9 (L16,17)	8
F	11/18			IV
M	11/21	L19: Sampling and approximate inference		
W	11/23	L20: Markov chain Monte Carlo and Gibbs sampling	10 (L18-20), V	9
		Drop Date		
M	11/28	*no class*		
W	11/30	L21: Continuous random variables		
M	12/5	L22: Joint PDFs; continuous inference		
W	12/7	L23: Jointly Gaussian random variables, innovations	11 (L21-25)	10
F	12/9			V
M	12/12	L24: Gaussian inference and modeling; linear inference		
W	12/14	L25: Central limit theorem		
T	12/20	**Final Exam**, 9:00am-12:00noon, DuPont		