Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
6.437 Inference and Information
Spring 2011

Information Sheet

Lecturers

Polina Golland
Office: 32-D470
E-mail: polina@mit.edu

Gregory W. Wornell
Office: 36-677
E-mail: gww@mit.edu

Teaching Assistants

Archana Venkataraman (pega85@mit.edu)
Da Wang (dawang@mit.edu)

Administrative Assistant

Tricia O'Donnell
Office: 36-677
Tel: 253-2297
E-mail: tricia@mit.edu

Lectures:	Tuesday and Thursday, 9:30am–11am, Room 32-155
Recitations:	Friday, 10-11am or 11am-noon, Room 26-322
Office Hours:	Days, times and locations posted on web site.

Welcome to 6.437 !

This course offers a graduate-level introduction to the principles of statistical inference, with an emphasis on information theoretic perspectives. As such, it is a core graduate subject for students in both Area I and the relevant subfields of Area II. The material in this course constitutes a common foundation for work in, for example, machine learning, signal processing, artificial intelligence, control, and communication.

It is worth stressing that 6.437 is an *introductory* graduate subject: it is not an advanced graduate subject for students who have already mastered both estimation and decision theory and information theory, yet want to understand such material at an even more sophisticated level. Nevertheless, by its structure, 6.437 will ultimately reveal connections among these fields, and provide background for more advanced treatments.

Ultimately, the course is about teaching you contemporary approaches to—and perspectives on—problems of statistical inference. The development of the material that forms the basis for this subject has historically been very much driven by applications. However, our focus in the course will not be on these applications—which form the basis for entire courses of their own—but rather on the common problem solving frameworks that they share. Nevertheless, we will cite various relevant applications as we develop the material and sometimes extract simplified examples from these contexts.

Note that the course has both lectures and recitations, which are designed to complement each other. Recitations begin the first week of classes. There are two possible recitation times to choose from, as indicated above. Select to attend whichever suits your schedule best. In addition, there are staff office hours scheduled throughout the week. You are welcome and encouraged to come to any and all of them you think might be helpful to you in clarifying your understanding of the material.

Prerequisites

The official prerequisite is either 6.041/6.431 or 6.436. The effective prerequisite is *fluency* with basic quantitative probabilistic reasoning and analysis, together with the kind of mathematical maturity that often comes from taking at least one higher level undergraduate subject that has a significant mathematical component. As such, a student having had 6.436 would be sufficiently well prepared, while a student having only had 6.041/6.431 and no subsequent subjects of a strong mathematical flavor would likely need additional preparation. As an example, having had 6.041/6.431 together with an introductory subject in analysis (e.g., 18.100), would be sufficient, but not necessary, preparation.

When in doubt, students whose undergraduate degrees are not from MIT should consult the staff to determine if they have had subjects that are effectively equivalent to the official prerequisites.

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 developing a set of course notes, which we will be distributing in parts as we go along. These notes are under active development, and as such are necessarily rough in places and contain bugs, which we will count on you to help us catch.

This semester, we will experiment with the course annotation site

`http://nb.csail.mit.edu`

that allows you to annotate course material online. We encourage you to to point out typos and points of confusion in the course notes. The website has a link to a tutorial that will help you to get started. The access to the site is invitation-based; all registered students will receive an invitation email shortly after the term starts.

You will also find sections of the following books to be useful and more in-depth auxiliary references for parts of the term. We will make essentially no use of these for the first several weeks of the term, so you will have plenty of time to browse through them beforehand to gauge their usefulness to you. We have placed all these books on reserve at the MIT libraries (Barker).

D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms,* Cambridge University Press, UK, 2003.

T. M. Cover and J. A. Thomas. *Elements of Information Theory,* Wiley, 2nd ed., 2006.

J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*, Wiley, 2000.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, 2nd ed., 2004.

If you are interested in further reading, either to strengthen your background, reinforce some of the concepts from lecture, or to probe some topics in more detail, you might want to take a look at the additional references on the course web site. In particular, you'll find several papers containing a variety of useful insights, which are worth the effort to work through.

Problem Sets

There will be 9 problem sets. Problem sets will be due in lecture. Problem sets must be handed in by the end of the class in which they are due. Problem set solutions will be available at the end of the due date's lecture.

While you should do all the assigned problems, only a randomly chosen subset will actually be graded. You will find some problems in the problem sets marked as "practice". These are not required, but you might 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 grades 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 at length!) the homework is a crucial part of the learning process and will invariably have a major impact on your understanding of the material. Some of the problem sets will involve a MATLAB component, to help you explore different aspects of the material.

In undertaking the problem sets, moderate collaboration in the form of joint problem solving with one or two classmates is permitted provided your writeup is your own.

Exams

There will be two (evening) quizzes in the subject. Dates for the quizzes are Wednesday, March 16, 7-10pm, and Wednesday, May 11, 7-10pm. The quizzes will be designed to require 1.5 hours of effort, but we'll use the three hour format to minimize the effects of time pressure. The quizzes will both be *closed book*. You will be allowed to bring *two* 8.5×11-inch sheet of notes (both sides) to Quiz#1, and *four* 8.5×11-inch sheets of notes to Quiz#2. The quizzes will be held in the lecture room (32-155).

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	40%
Quiz #2	45%
Homework	15%

with an additional property that if you do better on the Final Quiz than the Midterm Quiz, and you have done all the problem sets, then the Midterm Quiz will not count, i.e., the Midterm quiz can only help you if you are doing all the problem sets.

As always, other factors such as contributions to the lecture discussion and other interactions can make a significant difference in the final grade.

Course Web Site and Email

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

You should first make sure that you have an active Athena account (by visiting `http://web.mit.edu/accounts/` if necessary) as well as a personal certificate (by visiting `http://web.mit.edu/ist/topics/certificates/` 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/6.437`

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 6.437, this should already be set up; just double-check that you can access the web site (try to download a handout, for example). Otherwise, contact one of the TAs and they will add you to the list.

The student email list is

`6.437-students@mit.edu`

and will be kept in sync with the web site access list. 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

`6.437-staff@mit.edu`

This semester, we will also experiment with the web site

`http://www.piazzza.com`

that provides an easy interface (with LaTeX support) for students to discuss course materials. It is a much better alternative than the Stellar discussion forum. We encourage you to take advantage of the site to discuss course-related questions with your classmates. Note that we will neither monitor nor moderate this forum. The access to the site is invitation-based; all registered students will an receive invitation email shortly after the term starts.

Syllabus and Schedule

	Date	Topic	PS due	PS out
T	2/1	L1: Introduction and overview		1
R	2/3	L2: Bayesian hypothesis testing		
T	2/8	L3: NonBayesian decision theory	1	2
R	2/10	L4: Minimax decision theory		
T	2/15	L5: Bayesian parameter estimation	2	3
R	2/17	L6: NonBayesian parameter estimation		
T	2/22	*Monday schedule – no class*		
R	2/24	L7: Exponential families		
T	3/1	L8: Sufficient statistics	3	4
R	3/3	L9: The EM algorithm		
F	3/4	*Add Date*		
T	3/8	L10: Inference as decision	4	5
R	3/10	L11: Information geometry		
T	3/15	L12: Modeling as inference	5	6
W	3/16	**Evening Quiz 1** in 32-155 (through L10 and PS5)		
R	3/17	*no class*		
3/21-3/25		*Spring Break*		
T	3/29	L13: Extensions to continuous parameters		
R	3/31	L14: Priors		
T	4/5	L15: Alternating projections	6	7
R	4/7	L16: Approximations: deterministic		
T	4/12	L17: Approximations: stochastic	7	8
R	4/14	L18: Asymptotics: typical sequences, large deviations		
T	4/19	*Patriots Day – no class*		
R	4/21	L19: Method of Types and Sanov's Theorem		
		Drop Date		
T	4/26	L20: Asymptotics of hypothesis testing, estimation	8	9
R	4/28	L21: Asymptotics of model capacity		
T	5/3	L22: Introduction to parametric modeling		
R	5/5	L23: Model selection		
F	5/6		9	
T	5/10	L24: Review		
W	5/11	**Evening Quiz 2** in 32-155 (through L23 and PS9)		
R	5/12	*no class*		