Harvard University
School of Engineering and Applied Sciences

ENG-SCI 201/APMTH 231: Decision Theory

Spring 2019
Course Information

Course information

Instructor:
Demba Ba. (demba@seas.harvard.edu)
33 Oxford st, Room MD 143
Cambridge, MA 02138
(617) 495-1228 [Voice]
Office Hours: W 4:00-6:00 PM

Teaching Fellows:
Bahareh Tolooshams. (btolooshams@g.harvard.edu)
Office Hours: T 2:00-4:00 PM, 60 Oxford st # 330
Section: F 1:30-3:00 PM, MD 323

Tanvi Ranjan. (tanvi_ranjan@g.harvard.edu)
Office Hours: M 8:30-9:30 PM, Cruft 309
Section: M 7:00-8:30 PM, Cruft 309

Sai Qian Zhang. (zhangs@g.harvard.edu)
Office Hours: F 5:30-6:30 PM, MD 206

Lecture:
T-Th 10:30 AM–11:45 AM (Biolabs 1080)
Course Administrator:
Molly Marshall. (mmarshall@seas.harvard.edu)

Course Website:
`https://canvas.harvard.edu/courses/47876/`

Course overview

Statistics can be described as the science of making decisions under uncertainty. Given a model of outcomes of games in a league, odds from bookies and a starting capital, how does one maximize return on investment while mitigating the underlying randomness? This course will teach the mathematical foundations of decision making under uncertainty. ES 201 is a course in statistical inference and estimation from a signal processing perspective. The emphasis of the course will be on the entire pipeline from writing a model, estimating its parameters and performing inference. The focus will be on generative models, and on proving certain results (e.g. when/why does ℓ_1 regularization work? study of the convergence properties of popular optimization methods used to solve learning problems). The course will teach students how to develop new learning models, as well as to come up with learning algorithms and understand their properties. Topics include linear regression, logistic regression, maximum likelihood and kernel methods, compressive sampling and sparsity, Gaussian scale mixtures and hierarchical models, iterative re-weighted least-squares, EM and MM algorithms, state-space models, Bayesian regression, generative models of deep networks. An important component of the course will be a final project that will apply the course topics to data from Yelp, the NBA and many other exciting sources.

Prerequisites

The official prerequisites for this course are APPLIED MATH 21a or MATH 21a, and STATS 110 or equivalents. This is a highly interdisciplinary graduate-level course that will involve a combination of theory *and* computational modeling that are both motivated by data analysis problems. The key requirements are intellectual curiosity and a desire to learn to think about data in new ways. A certain level of mathematical maturity is assumed. In particular, prior exposure to abstract linear algebra and real analysis will deepen your understanding of the materials.

Textbook

Machine Learning: A Bayesian and Optimization Perspective, by Sergios Theodoridis, Academmic Press, 2015.

Policy on collaboration

To get the most out of this course, you are encouraged to struggle with the course assignments on your own and reach out to the course staff during Office Hours. **You are allowed to *discuss* the *content* of assignments with fellow students but *not* their *solutions*.** Your write-up of assignments must entirely be your own. Moreover, at the top of every assignment, you are kindly asked to acknowledge the students you have discussed said assignment with. We also ask you to acknowledge the use of books, articles, websites, lectures, discussions, etc., that you have consulted to complete your assignments.

Grading information

The final grade for this course will be based on your performance on problem sets, a midterm examinations and a final project. There will not be a final exam in this class.

1. **Problem sets**: There will be **6** bi-weekly **problem sets** that will count towards **45% of your final grade**. Problem sets will be due at the beginning of class on the date stated in the course calendar. There will be a non-negotiable grace period of 10 minutes. Problem sets will consist of "pen-and-paper assignments" and/or computational assignments. We will accept two methods of submission for the "pen-and-paper" component. You may either hand it in at the beginning of class or write it up electronically and submit it electronically along with the computational component.

2. **Midterms**: There will be **1 midterm examination** that will each count towards **30% of your final grade**.

3. **Final project**: The **final project** will count towards **15% of your final grade**. The final project can either be performed alone or in a group of up to three people. All members of a group will get the same grade on the final project. The bigger the group, the higher the expected output of the project. The goal of the final project is three-fold: we ask that you (a) formulate an interesting question that involves real data, (b) gather the data required to answer this question and (c) used concepts from the course to tell an interesting story in the form of an ipython notebook and short (10 to 15-minute) presentation to be delivered in front of the class. The grade will be based on your ability to apply the concepts and tools taught in class, and your ability to integrate (a), (b) and (c). We strongly encourage that you consult with the course staff as early as possible. We will collect final projects as Jupyter notebooks.

4. **Project proposal**: We ask that you put together a short one-page project proposal to be discussed with the course staff following Spring Recess (see course calendar). The project proposal will count towards **10% of your final grade**.

Disclaimer

While the above weights are used for computing the final grade, I reserve my right to scale the grades based on the performance of the entire class. Naturally, this possible scaling will not have an adverse effect on the grades and can only increase the raw grades.

Table 1: **Course Calendar**

	Date	Topic	Assignments
		Regression, Classification and Optimization	
T	01/29	Course overview	
Th	01/31	Maximum likelihood estimation	**Pset 1 out**
T	02/05	Linear regression	
Th	02/07	Intro to Convex Optimization I	
T	02/12	Intro to Convex Optimization II	
Th	02/14	Exp. family and weighted least-squares	**Pset 1 due, Pset 2 out**
T	02/19		
Th	02/21	SVMs: classification meets optimization	
T	02/26	Classification in RKHS	
Th	02/28	Decision trees classification	**Pset 2 due**
		Regularization, hierarchical models and Bayesian thinking	
T	03/05	Ridge regression	**Pset 3 out**
Th	03/07	ℓ_1-regularized regression, GSMs and IRLS	
T	03/12	EM, MM and optimization	
Th	03/14	**Midterm 1**	**Pset 3 due, Pset 4 out**
T	03/19	**Spring Recess**	**Reading assingment**
Th	03/21	**Spring Recess**	**Reading assingment**
T	03/26	ℓ_1-regularization theory	
Th	03/28	Bayesian logistic regression and IRLS	**Project proposal due**
		State-space models	
T	04/02	Linear-Gaussian state-space models and the Kalman filter	**Pset 4 due, Pset 5 out**
Th	04/04	ML estimation of Linear-Gaussian SSMs	
T	04/09	Binomial SSMs of neural spiking	
Th	04/11	<u>**Guest lecture**</u>**: Prof. Pierre Jacob**	
		Neural networks	
T	04/16	Perceptron and logistic regression as NN	**Pset 5 due, Pset 6 out**
Th	04/18	Intro to deep networks I	
T	04/23	Intro to deep networks II	
Th	04/25	Neural networks and sparse representations	
		Project presentations	
T	04/30	Group 1	**Pset 6 due**
Th	05/02	Group 2	
T	05/07	Group 3 (**if necessary**)	
Mon	05/13	**Final project due**	