PROB 140

Probability for Data Science

Fall 2020

WEEK 11 STUDY GUIDE

The Big Picture

Estimation from the frequentist and Bayesian perspectives

- In the frequentist world, a parameter is a fixed but possibly unknown number. The method of *maximum likelihood* identifies the parameter that makes the data most likely.
- In the Bayesian world, the data scientist's degree of uncertainty about unknown quantities is described by probability distributions. Unknown parameters are therefore random variables, and inference consists of updating our distribution of the parameter based on the observed data. For these calculations we need some basic methods for conditioning on continuous variables.
- The main example is inference for the unknown *p* of a coin. Independence is affected by the randomization of the parameter.
- The beta family is a rich class with which to describe our *prior* opinions about p; it then turns out that the same family describes our *posterior* opinion which is the prior updated based on the observed heads and tails.
- If the prior is uniform, the MAP estimate (the mode of the posterior) is the same as the maximum likelihood estimate with which the week started.

Week At a Glance

Mon 11/2	Tue 11/3	Wed 11/4	Thu 11/5	Fri 11/6
	Instructor's Session		Instructor's Session	
		GSIs' Sessions		GSIs' Sessions
Checkpoint Week 11 (Due Thu 11/5)			**Checkpoint Week 11 Due**	
HW 9 (Due Mon 11/9)				HW 9 Party 6-7PM
Lab 6A (Due Mon 11/9)		**Lab 5B Due**	Lab 6A Party 6-7PM	
	Skim Chapter 20	Read Chapter 20	Skim Chapter 21	Read Chapter 21

Note: This is an unusual week in the world as well as in 140. Chapters 20 and 21 together have only about as much content as a single one of the recent chapters. So you can take it a little easy.

Reading, Practice, and Live Sessions

Sections	Topic	Live Sessions: Prof. A.	Live Sessions: GSIs	Recommended Practice
Ch 20	**Approaches to inference** - 20.1 is about the method of maximum likelihood - 20.2 introduces conditioning on a continuous variable, and shows once again that randomizing a parameter affects dependence and independence (you saw this earlier with Poissonization) - 20.3 formalizes the concepts of prior and posterior distributions of parameters, and compares the MAP estimate and the MLE	Tuesday 11/3 - Maximum likelihood - Random parameters: conditioning and independence **Checkpoint is based on Chapter 20** Thursday 11/5 - Inference for the random p of a coin - Conjugate priors; prediction - Relations between the beta and the binomial, including the beta-binomial distribution	Wednesday 11/4 - Ch 20 Ex 1 - Ch 19 Ex 2 - Sec 19.4 example on Chernoff bound for the tail of a normal Friday 11/6 - Ch 20 Ex 5 - Ch 21 Ex 2 - Ch 20 Ex 6	Ch 20 - Ex 2, 4, 6
Ch 21	**Inference for the *p* of a random coin** - 21.1 picks up from 20.3, with a general beta prior instead of uniform - 21.2 is about the unconditional distribution of the number of heads, which is called beta-binomial - 21.3 is omitted this term			Ch 21 - All exercises not done in section