DATA 140 Probability for Data Science

Spring 2024

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.
- Our 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 4/1	Tue 4/2	Wed 4/3	Thu 4/4	Fri 4/5
	Lecture	Section	Lecture	Mega Section
Lab 7A Due Lab 7B (Due Mon 4/8)			Lab 7B Party 9 AM to 11 AM	
HW 10 Due HW 11 (Due Mon 4/8)				HW 11 Party 2 PM to 5 PM
Skim Section 20.1	Work through Chapter 20	Work through Chapter 20	Work through Section 21.1	Work through Sections 21.1 and 21.2

Reading, Practice, and Class Meetings

Book	Topic	Lectures: Prof. A.	Sections: TAs	Optional Additional 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 4/2 - Maximum likelihood - Random parameters: conditioning and independence	Wednesday 4/3 - Ch 20 Ex 1 - Ch 20 Ex 4 - Ch 20 Ex 6	Ch 20 - Ex 2, 6
Ch 21	**Inference for the random *p* of a 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	Thursday 4/4 - Inference for the random *p* of a coin - Conjugate priors; prediction - Relations between the beta and the binomial, including the beta-binomial distribution	Friday 4/5 - Ch 20 Ex 5 - Ch 21 Ex 3 - Ch 21 Ex 4	Ch 21 - All exercises not done in section