Spring 2024

WEEK 1 STUDY GUIDE

The Big Picture

We begin the course with a formal mathematical framework for defining and combining probabilities, and a mathematical definition of “random quantity.”

- The basic rules of probability are the same as those for proportions. To find a probability, you have to figure out an appropriate combination of rules to use.
- Numerical calculations quickly get large. Even in this age of powerful computers, mathematical approximation is often important for computation and insight.
- One of the rules enables you to update probabilities in the light of new information. This is a fundamentally important skill in data science. Assumptions matter, for identifying the right methods to use as well as for interpreting results.
- In Data 8 you saw a *statistic* defined as a number that you compute based on a sample. The more general concept is that of a *random variable*, which is a function on the outcome space.
- *Distributions* describe how probability is spread over a set of values. Every random variable has a distribution; pairs and larger groups have *joint distributions*.

Week At a Glance

Tue 1/16	Wed 1/17	Thu 1/18	Fri 1/19
Lecture	Sections	Lecture	Mega sections
HW 1 (due 5 PM Mon 1/22)			HW 1 Party 2 PM to 4 PM
Lab 1 (due 5 PM Mon 1/22)		Lab 1 Party 9 AM to 11 AM	
Work through Chapter 1	Work through Chapter 2	Skim Chapter 3	Work through Chapter 3 and skim Chapter 4

Reading, Practice, and Class Meetings

Book	Topic	Lectures: Prof. A.	Sections: TAs	Optional Additional Practice
1.1, 1.2	Probability as a function - 1.1 defines the domain - 1.2 shows how to find probabilities assuming equally likely outcomes	Tue 1/16 1.3-1-5 with an emphasis on the math more than the computation 2.1, 2.3, 2.5: The relation between axioms and rules; conditioning .	Wed 1/17 - "Balls in boxes": how this helps with visualization in problems that look very different from each other - Exponential approximation - What is common to Ch 1 Ex 7, 4, 6, 8, and Ch 2 Ex 12	Chapter 1 1, 2, 10
1.3, 1.4	An example of an exact calculation, using the product rule of counting - 1.3 has the general calculation - 1.4 has the numerical computation in a special case			
1.5	The first of many exponential approximations in the course			Chapter 2 1, 5, 6 If you have time, try 14. It's popular with quant interviewers.
2.1, 2.3	The axioms and basic rules - 2.1 is about addition, and hence also subtraction - 2.3 is about multiplication, and hence also division (crucial for conditional probabilities)			
2.5	Bayes' Rule: updating probabilities by conditioning probabilities by conditioning			
2.2, 2.4	Examples. Don't just read them – work them out			

Book	Topic	Lectures: Prof. A.	Sections: TAs	Optional Additional Practice
Ch 3	Random variables - 3.1 has the definition - 3.2 defines the distribution of the random variable, and shows how to find probabilities of events based on the random variable - 3.3 shows how random variables can have two kinds of equality	Thu 1/18 - The key ideas in Chapters 3 and 4, focusing more on the math than the code	Fri 1/19 - Conditioning and Bayes: points to notice - Random variables and equality - Chapter 2 Ex 13 - Chapter 3 Ex 3, 5ab - Chapter 4 Ex 5	Chapter 3 4, 7
Ch 4	Pairs of random variables - 4.1 is the two-variable version of 3.2: joint distributions, and finding probabilities - 4.2 has examples you should study - 4.3 shows how to extract the behavior of one random variable from the combined behavior of two - 4.4 shows how to update chances for one random variable given the value of another - 4.5 looks at how joint distributions help us understand dependence and independence; note the acronym “iid”			Chapter 4 Do as much as you can of all five exercises, but it’s fine to spread that over Week 2 as well.

Chapters 3-4 aren’t difficult technically, but they contain many basic concepts and essential terminology.