

# PROB 140

Spring 2022

## WEEK 2 STUDY GUIDE



Probability for Data Science

### The Big Picture

We continue to develop the basic toolkit: how to work with collections of random variables and collections of events. This gives us the tools to study some fundamentally important families of distributions.

- If events have a complicated dependence structure, you might not be able to calculate exact or even approximate chances. Sometimes the best you can do is find *bounds* for a chance.
- *Symmetry* in random permutations and simple random samples greatly simplifies calculations.
- There is a formula for the exact chance of the union of overlapping events, with a famous application.
- Distributions on a large finite number of values can be approximated by distributions on infinitely many values; a fundamentally important example of this is introduced.
- Random samples often result in random counts. The distribution of the count depends on the method of sampling.
- If the sample is a fixed number of i.i.d. success/failure trials, the distribution of the number of successes is *binomial*. The shape of the distribution can be understood by using consecutive odds ratios.
- In some situations, the binomial distribution is well approximated by a *Poisson* distribution, introduced earlier.

### Week At a Glance

Mon 1/24	Tue 1/25	Wed 1/26	Thu 1/27	Fri 1/28
	Lecture		Lecture	
		Sections		Sections
HW 1 Party 9am to 12 noon <b>HW 1 Due</b> HW 2 (Due Mon 1/31)				
Lab 1A (Due Mon 1/31)				Lab 1A Party 3pm to 5pm
Finish working through Ch 4; Skim Ch 5	Work through Ch 5	Finish working through Ch 5; skim Ch 6	Work through Ch 6	Fill any holes you left in working through Ch 4-6

## Reading, Practice, and Live Sessions

Book	Topic	Lectures: Prof. A.	Sections: GSIs	Optional Additional Practice
Ch 5	<ul style="list-style-type: none"> <li>- 5.1: Simple bounds for chances of unions and intersections</li> <li>- 5.2: The exact chance of a union, overlapping or not (requires the chances of all the overlaps)</li> <li>- 5.3: One of the most famous applications of inclusion-exclusion is to <i>fixed points</i> of a <i>random permutation</i>, also known as <i>matches</i>; this can be approximated by a distribution on infinitely many values</li> <li>- 5.4: Summary of results on symmetry in random permutations and simple random sampling</li> </ul>	<p>Tue 1/25</p> <p>Highlights of Ch 5</p>	<p>Wed 1/26</p> <ul style="list-style-type: none"> <li>- Lab 1A Part 1: a new look at the TVD</li> <li>- Chapter 5 Exercises 1, 9, 12</li> </ul>	<p>Chapter 5</p> <p>5, 6, 10, 13</p>
Ch 6	<ul style="list-style-type: none"> <li>- 6.1: In a fixed number of i.i.d. 0/1 trials, the number of successes has a <i>binomial</i> distribution</li> <li>- 6.2: Examples you should read</li> <li>- 6.3 extends the binomial to the <i>multinomial</i> case where each trial has several possible outcomes</li> <li>- 6.4 compares the number of successes when the sampling is with replacement (binomial) and when the sampling without replacement (<i>hypergeometric</i>)</li> <li>- 6.5 uses odds ratios to study the shape of binomial histograms, and finds the mode</li> <li>- 6.6 uses odds ratios to show that under some conditions the binomial has a <i>Poisson</i> limit</li> </ul>	<p>Thu 1/27</p> <p>Highlights of Ch 6</p>	<p>Fri 1/28</p> <ul style="list-style-type: none"> <li>- Ch 6 Ex 2, 4, 10, 11, 5</li> </ul>	<p>Chapter 6</p> <p>1, 4, 5, 11, 12</p>