

# PROB 140

Spring 2021

## WEEK 3 STUDY GUIDE



Probability for Data Science

### The Big Picture

This week is largely about two of the principal families of distributions: the binomial and the Poisson.

- Random samples often result in random counts, such as the number of voters who favor a candidate.

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. The Poisson is our first distribution on infinitely many values.
- Randomizing parameters can have dramatic effects on dependence and independence. A Poisson number of i.i.d. success/failure trials has beautiful and powerful properties.

### Week At a Glance

Mon 2/1	Tue 2/2	Wed 2/3	Thu 2/4	Fri 2/5
	Instructor's Session		Instructor's Session	
		GSI's Sessions		GSI's Sessions
Checkpoint Week 3 (Due Wed 2/3)		<b>Checkpoint Week 3 Due</b>		
HW 2 Party 7PM <b>HW 2 Due</b> HW 3 (Due Mon 2/8)				
<b>Lab 1B due</b> Lab 2A (Due Mon 2/8)				Lab 2A Party 5PM
Skim 6.1 through 6.4	Read/watch 6.5, 6.6	Read Chapter 6	Skim 7.1, 7,2	Read Ch 7

## Reading, Practice, and Live Sessions

Chapter	Topic	Live Sessions: Prof. A.	Live Sessions: GSIs	Recommended Practice
Ch 6	<b>Binomial and its relatives</b> <ul style="list-style-type: none"> <li>- 6.1: a fixed number of i.i.d. success/failure 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, not two</li> <li>- 6.4 compares the number of successes when sampling with replacement (binomial) and the number of successes when sampling without replacement (hypergeometric)</li> <li>- 6.5 examines the shape of the binomial histogram, and identifies the mode, by studying odds ratios</li> <li>- 6.6 uses odds ratios to show that under some conditions the binomial has a <i>Poisson</i> limit</li> </ul>	<p><b>Tue 2/2</b></p> <ul style="list-style-type: none"> <li>- Straightforward but important observations about success counts</li> <li>- Deeper dive into the math to explain what we see</li> <li>- An approximation, leading to a new class of distributions</li> </ul> <p><b>Checkpoint is based on Chapter 6</b></p>	<p><b>Wed 2/3</b></p> <ul style="list-style-type: none"> <li>- Ch 6 Ex 2</li> <li>- Lab 2A Part I: a new look at total variation distance</li> <li>- Ch 6 Ex 5</li> </ul>	<p><b>Chapter 6</b> 1, 4, 11, 12</p>
Ch 7	<b>Poissonization</b> <ul style="list-style-type: none"> <li>- 7.1 has properties of the Poisson distribution</li> <li>- 7.2 asks the same questions as 6.1, but with a Poisson number of trials</li> <li>- 7.3 extends this to trials with more than two categories, analogous to 6.3</li> </ul>	<p><b>Thu 2/2</b></p> <p>Poissonization:</p> <ul style="list-style-type: none"> <li>- Beautiful calculations with surprising results</li> <li>- Pay attention to the math because you'll need the methods again</li> </ul>	<p><b>Fri 2/5</b></p> <p>Poisson approx and Poissonization:</p> <ul style="list-style-type: none"> <li>- Ch 6 Ex 10</li> <li>- Ch 7 Ex 2</li> <li>- Ch 7 Ex 8</li> </ul> <p>Relations with other exercises</p>	<p><b>Chapter 7</b> 3, 4, 5</p>

Lab 2 is about measuring the quality of the Poisson approximation in 6.6.