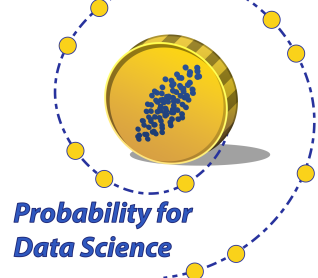


# DATA 140



Spring 2023

## WEEK 8 STUDY GUIDE

### The Big Picture

We move to random variables with a continuum of values, via one of the most important theorems in probability.

- We know how to find expectations and variances of sums of random variables. To find the distribution of a sum, we can use partitioning as before. But a more abstract math concept of a *probability generating function* lets us quickly calculate distributions of sums in special cases.
- Many of the simulations in Data 8 are evidence of the *Central Limit Theorem* in action: the distribution of the sum of a large i.i.d. sample is roughly normal. We use this to construct confidence intervals for the population mean.
- The normal is a continuous curve that acts as a probability distribution. We formally define the *density* of a random variable with a continuum of values, and extend the concepts of cdf and expectation to this situation.
- Along with the normal, we study two major distribution families: the uniform and the exponential.

### Week At a Glance

Mon 3/6	Tue 3/7	Wed 3/8	Thu 3/9	Fri 3/10
	Lecture	Sections	Lecture	Sections
Lab 5 (Part A due Mon 3/13)			Lab 5A Party (10 AM to noon)	
<b>HW 7 Due</b> HW 8 (Due Mon 3/13)				HW 8 party (3 PM to 5 PM)
Skim Sections 14.1-14.2	Work through Sections 14.1-14.2, skim the rest of Ch 14	Work through Ch 14, skim Section 15.1	Work through Sections 15.1-15.2	Work through Sections 15.1-15.4

## Reading, Practice, and Live Sessions

Sections	Topic	Lectures: Prof. A.	Sections: GSIs	Optional Additional Practice
Ch 14	<b>Sums and the CLT</b> <ul style="list-style-type: none"> <li>- 14.1-14.2 cover an abstract math method for understanding probability distributions; 12.2 finds exact distributions of i.i.d. sample sums.</li> <li>- 14.3 states the Central Limit Theorem and formally defines the normal curve</li> <li>- 14.4 shows how to work with the normal curve in Python; <b>this is for you to read by yourself</b></li> <li>- 14.5-14.6 cover the distribution of the i.i.d. sample mean, and hence the use of the sample mean in confidence intervals</li> </ul>	<b>Tuesday 3/7</b> <ul style="list-style-type: none"> <li>- Our first generating function: a math technique for understanding distributions</li> <li>- The CLT and some consequences</li> </ul>	<b>Wednesday 3/8</b> Ch 14: - Ex 1, 4, 6	<b>Ch 14</b> - 2, 3, 5, 7
Ch 15	<b>Random Variables with Densities</b> <ul style="list-style-type: none"> <li>- 15.1-15.2 define a “continuous” probability histogram, and generalize the concept of density from Data 8 histograms</li> <li>- 15.3 covers expectation (including variance) and has examples including the uniform distribution family</li> <li>- 15.4 covers the exponential distribution family</li> <li>- 15.5 shows how to do calculus in SymPy, included in your lab</li> </ul>	<b>Thursday 3/9</b> Random variables on a continuum of values: extending all previous concepts to this case, and recognizing a benefit of the continuous world: single points don’t affect probability calculations	<b>Friday 3/10</b> Ch 15: - Ex 1, 3, 5	<b>Ch 15</b> - 2, 9, 10, 11