PROB 140

Probability for Data Science

Fall 2022

WEEK 7 STUDY GUIDE

The Big Picture

We define and establish properties of the main measure of the variability in a distribution.

- The *standard deviation*, familiar to you from Data 8 as a measure of the spread in a data distribution, is defined as a measure of spread in the distribution of a random variable.
- *Variance*, which is the mean squared error and the square of the standard deviation, has better computational properties.

Non-standard week, non-standard guide

Week At a Glance

Mon 10/3	Tue 10/4	Wed 10/5	Thu 10/6	Fri 10/7
	Midterm		Lecture	Sections
HW 6 Due			HW 7 (Due Mon 10/10)	HW 7 Party 3PM - 5PM
Lab 3 Due				
			Work through Sections 12.1, 12.3	Work through Sections 12.1, 12.2, 12.3

Reading, Practice, and Class Meetings

Book	Topic	Lecture: Prof. A.	Section: GSIs	Optional Additional Practice
Ch 12	**Variance and Standard Deviation** - 12.1 has the basics of SD and variance; much of this should be an easy read - 12.2 connects variance and prediction - 12.3 shows how expectation and variance can be used to bound the tails of a distribution - 12.4 has examples of distributions with heavy tails, for students interested in economics, natural language processing, etc	Thursday 10/6 SD and variance: - Definition, alternative computational method, examples - Use in prediction - Tail bounds	Friday 10/7 Ch 12: - Ex 4, 5, 6	Ch 12 - All exercises not covered in section