PROB 140 Spring 2021



## **WEEK 14 STUDY GUIDE**

## **The Big Picture**

Probability for Data Science

Simple linear regression predicts Y as a linear function of a single X. No matter what the joint distribution of X and Y, there is always a least squares line. If X and Y are bivariate normal, this line turns out to be the best among all predictors.

- A straightforward least-squares calculation results in the Data 8 formula for the equation of the regression line.
- Standard bivariate normal X and Y can be constructed so that Y is the sum of a linear function of X and independent normal noise.
- For the bivariate normal, the conditional expectation is a linear function of the given variable, and hence is the same as the best linear predictor.
- The regression line can be written in multiple forms, one of which extends to the case of multiple regression.

### Week At a Glance

Mon 4/26	Tue 4/27	Wed 4/28	Thu 4/29	Fri 4/30
	Instructor's Session		Instructor's Session	
		GSIs' Sessions		GSIs' Sessions
No new checkpoint				
HW 11 Party 7PM HW 11 Due HW 12 (Due Mon 5/3)				
Lab 7 Due No new lab				
Skim Sections 24.1-24.2	Read Chapter 24			

# **Reading, Practice, and Live Sessions**

Sectio ns	Topic	Live Sessions: Prof. A.	Live Sessions: GSIs	Recommend ed Practice
Ch 24	Simple Regression - 24.1 derives the equation of the regression line - 24.2 constructs bivariate normal random variables so that the relation between can be expressed in terms of "linear signal plus noise" - 24.3 looks at least-squares prediction in the context of the bivariate normal, and the connection with linear regression - 23.4 writes the regression equation in multiple different ways, each one illuminating a different property	Tuesday 4/27  - Simple regression: general case - Bivariate normal - Regression and the bivariate normal  No checkpoint  Thursday 4/29	Wednesday 4/28 - Ch 24 Ex 2, 3, 5  Friday 4/30 - Ch 24 Ex 1, 7 - Overview of Ch 25.4	Take a break from exercises, but read Ch 24 carefully.