COLLEGE OF CHEMISTRY COURSE GUIDE (../INDEX.HTML)

STAT 140 - PROBABILITY FOR DATA SCIENCE (4 UNITS)

(Taken from the UC Berkeley Course Guide (http://guide.berkeley.edu))

COURSE OVERVIEW

SUMMARY

An introduction to probability, emphasizing the combined use of mathematics and programming to solve problems. Random variables, discrete and continuous families of distributions. Bounds and approximations. Dependence, conditioning, Bayes methods. Convergence, Markov chains. Least squares prediction. Random permutations, symmetry, order statistics. Use of numerical computation, graphics, simulation, and computer algebra.

PREREQUISITES

STAT/CS/Information C8, or STAT/CS C100, or both STAT 20 and CS 61A; one year of calculs at the level of MATH 1A (math1a.html)-MATH 1B (math1b.html) or higher. Corequisite: MATH 54 (math54.html), EE 16A (ee16a.html), STAT 89A, MATH 110 (math110.html) or equivalent linear algebra.

Students who have earned credit for STAT 134 (stat134.html) will not receive credit for STAT 140.

TOPICS COVERED

The emphasis on simulation and the bootstrap in Data 8 gives students a concrete sense of randomness and sampling variability. Stat 140 will capitalize on this, abstraction and computation complementing each other throughout.

The syllabus has been designed to maintain a mathematical level at least equal to that in Stat 134. So Stat 140 will start faster than Stat 134 (due to the Data 8 prerequisite), avoid approximations that are unnecessary when SciPy is at hand, and replace some of the routine calculus by symbolic math done in SymPy. This will create time for a unit on the convergence and reversibility of Markov Chains as well as added focus on conditioning and Bayes methods.

With about a thousand students a year taking Foundations of Data Science (Stat/CS/Info C8, a.k.a. Data 8), there is considerable demand for follow-on courses that build on the skills acquired in that class. Stat 140 is a probability course for Data 8 graduates who have also had a year of calculus and wish to go deeper into data science.

Understand the difference between math and simulation, and appreciate the power of both; Use a variety of approaches to problem solving; Work with probability concepts algebraically, numerically, and graphically

WORKLOAD

TIME COMMITMENT

3 hours of lecture, 2 hours of discussion, and 1 hour of supplement per week.

UC Berkeley Course Guide (http://guide.berkeley.edu)

Made by Angela Lee, c/o 2019

f   𝕏   🔗

(https://w(http://twitter.com/cal_study.berkeley.edu/ugrad/curr

lang=en) students/peer-

advising