Stat 218 Probability for Computer Scientists Fall 2018
Course Description and Syllabus

Vital Info

Class Schedule:	T/Th 10:30-11:50 in Wean 7500
Instructors:	Christopher Genovese 232E Baker Hall genovese@cmu.edu (Please put "[218]" in all email subject lines.) Mikael Kuusela 232C Baker Hall mkuusela@andrew.cmu.edu
Office Hours:	TBD empirically and by appointment
TA's:	Nil-Jana Akpinar Alek Podkopaev Shengming Luo Heejong Bong Email: 218-staff@stat.cmu.edu
Office Hours:	TBA and by appointment
Home Page:	http://www.cmu.edu/canvas
Text:	Course Notes

Course Objectives

To teach you to develop mathematical and computational models of uncertain or random systems so as to make better predictions and decisions.

By the end of this course, you should be able to

- build effective models of realistic random systems;
- use the models to make decisions, especially in computer-science and statistical-learning applications;
- plan an approach, generate strategies, and solve problems that require probabilistic/statistical thinking;
- explain fundamental concepts in probability theory, statistics, and statistical learning.

You will achieve these goals with deliberate practice of a specific set of skills, as described below.

Course Scope

This is a first course in probability theory, and no prior background in probability or statistics is assumed or required. Probability theory is at the core of statistics, machine learning, decision theory, and many other areas. Throughout the semester, we will develop the concepts and methods of probability and apply them to problems in statistics, statistical machine learning, and computer science, among others.

A probability model is a testable framework for describing a random system with which we can make better decisions and predictions. Building a model is an iterative process of refining our assumptions in light of the questions we wish to answer and the adequacy of the model's predictions. We will practice building models for systems great and small, through case studies and activities in class, through mathematical analysis and computational simulation, and through group discussion.

Mathematics is the language in which we describe and analyze our models. Many find the notation of the mathematics daunting, but we will see that it is very much like a programming language in how we use it. Like a programming language, mathematics has a very specific syntax and semantics that reinforce each other and help us to simplify complex and abstract ideas. We will use a consistent notation through the semester, and we will try to show you how the notation (syntax) provides mnemonics for the underlying meaning (semantics). Also like a programming language, mathematics has its own idioms – patterns of expression that are used again and again to solve analogous problems.

Unlike most probability courses at this level, we will emphasize the foundations of a probability model in some detail. The reason is that once this is understood, all of the work we do later with random variables, random experiments, probabilities, and expectation will make more sense. We will explore the concepts of distributions and conditioning. We will work with mathematical models for random systems evolving in time or space, which are called stochastic processes. And we will develop statistical methods using probabilistic concepts.

Resources

NOTES. The lecture slides from each class, complete with annotations made during class, will be available online. You can also access and annotate these slides online yourself during class (see next subsection) and save them for later reference.

In addition, notes describing concepts and examples in more detail will be available on the Canvas site. These collectively comprise the text for the course. You should plan to read and/or review these notes before class.

ISLE. As part of our in-class workflow, we will be using a technology developed here at Carnegie Mellon, called **ISLE**. We will be presenting and annotating lecture notes through ISLE. You can sign in during class to see these notes and add your own annotations. You can also sign in at any point to download PDFs of the annotated lecture notes. In addition, we will use ISLE for a variety of interactive, in-class questions and activities.

For these reasons, students are strongly encouraged to bring a web-enabled device to class if you have one.

To sign up the first time, visit

`https://isle.heinz.cmu.edu/36-218/welcome/`

click on "Sign Up" button in the pop-up dialog on the left. Use your andrew email address as a username.

For each class, there will be a separate link for the day's lecture notes. These will be posted on Canvas and presented in class on that day.

CANVAS. The course home page, hosted through Canvas, is a repository for all documents and other information for the course, including assignments, solutions, and course announcements. We will also maintain an updated course calendar and office hour list.

GRADESCOPE. In this course, we will be using Gradescope (as a plug-in tool within Canvas) to grade and provide feedback on assignments and exams. This will allow your graders to provide more timely and effective feedback. It also promotes fairer grading practices by facilitating anonymous grading and question-by-question (rather than student-by-student) grading. In addition, Gradescope makes it easy for you to

access and review all your (graded) work. During the semester, students will use Gradescope to (a) submit work online, (b) view feedback and scores on graded work, and (c) make a re-grade request within prescribed guidelines. To access Gradescope, simply log on to our course's Canvas site and click on Gradescope in the left-side navigation bar.

A key step in submitting your work to Gradescope is getting a high-quality, digitized scan of your work. There are numerous scanners around campus, and scans can be made with iOS and Android devices. We have prepared a website for students

`https://www.cmu.edu/teaching/gradescope/`

that provides more information on using Gradescope, including how to scan assignments via iOS and Android devices, where to find scanners around campus, and how to submit assignments once scanned.

Be sure to take the following important policies and procedures into account whenever you are submitting work to Gradescope:

- If you are writing your assignment by hand (on paper), be sure to use a dark pencil or pen, and write as clearly as possible.
- Provide a clear separation and demarcation between different exercises on the assignment.
- When you upload your work to Gradescope, be sure to (a) indicate where each question is located within your submission via the click-and-select interface and (b) after you submit, review each page of your uploaded submission to make sure everything is clear and legible.
- Give yourself some extra time to prepare and submit your assignment online to Gradescope, especially for the first few assignments when you are still getting familiar with it.
- Keep a soft copy of each scanned assignment for your records.
- Regrading requests will be handled through Gradescope; see *Exam and Homework Regrading* below.
- If you need help with technical issues related to Gradescope, email: `canvas-help@andrew.cmu.edu`.
- Important grading policy: If the grader cannot read your submission, there is no way to award points.

Office Hours. The instructors and the TAs will hold office hours during the week. These offer a terrific opportunity for you to deepen

your understanding of the material, by asking questions and going over problems. It is common and understandable (and good) for people to visit office hours when they are having problems with an assignment. Also, keep in mind that office hours are useful for getting questions answered more generally.

Assignments

All assignments and assessments in this class are designed to encourage you to engage with the material and *practice* the skills necessary to learn it well. That's really what we care about, and we are less concerned with answers than process.

Reading will be assigned for every class (except the first one), usually just a few pages, and you are expected to do the assigned reading before class. A path to success in this course includes reading carefully, thoughtfully, and actively. Write down questions as they occur to you; try the examples as you read; make connections to what we have done in class and homework.

Homework exercises will be a blend of mathematical work, modeling, and occasionally, simulation. Each assignment will be a mixture of questions from three groups, distinguished by their point value. The first group (2 points each) will offer you opportunities for practice with the basic definitions, concepts, and mathematical methods. These are typically short and focused, often without a detailed cover story. The second group (4 points each) consists of longer problems that synthesize multiple concepts or require several distinct steps. The final group (6 points each) deal with more complex concepts and reasoning, and often involve using the concepts in context (e.g., of an application). A few of these will be open-ended questions for you to answer based on a mathematical model or computer simulation of a random system. In addition, several exercises on each assignment, drawn from all three groups, will be marked "Solo," and have special rules on the sources you can use in solving them. See the *Academic Integrity* section for details.

The purpose of the homework is to give you practice with the techniques and ideas we cover in reading and class. As such, they can be of no help to you if you do not try them. Grading for the homeworks in the class is founded on this idea; see the rubric in the *Grading* section.

Your lowest two homework scores will be dropped from the calculation of your grade. We may also occasionally hand out supplementary exercises for additional practice and study. You are encouraged to try these and ask us questions about them, but they are not required.

A homework assignment will be posted on the home page on most Tuesdays and **due one week later**. However, there is a 24-hour grace period: **we will accept homeworks, without penalty, until 6pm Wednesday**. Homework is to be submitted electronically through Gradescope as described in the *Resources* section. The 6pm Wednesday deadline is strict to allow the TA's to grade the homework in reasonable time. In unusual circumstances, we are willing to accept homework beyond the grace period, but this must be discussed with one of the instructors *at least 36 hours before the assignment is due* and will be decided on a case-by-case basis. Otherwise, homework is not accepted beyond the grace period. Having another project or assignment due around the same time is usually not a sufficient excuse for late turn-in.

It is hard to over-stress the following advice: Start your homework early, and take it a little at a time. Talk to the TA's and instructors as questions arise, and give the exercises your best shot. It will take less of your time overall to work regularly on homework for a short time each session (asking questions in between) than it will to jam out an entire assignment the night (or afternoon) before it is due.

Solutions to the homework exercises will be posted on the course home page as soon as possible after 6pm on Wednesday. Your graded homeworks will be returned via Gradescope as soon as possible, typically a week after the due date (when class is on).

Exams

We will have two in-class midterm exams and a three-hour, cumulative final exam on the date to be announced by the registrar. If needed, we may also schedule a few short quizzes. See the *Grading* section for details on how these exams contribute to your final grade.

Early travel plans are not a valid reason for missing the final exam, so please do not schedule travel until after the final exam date has been announced.

Grading

Homework. Homework is assessed by the following criteria:

- *Effort* – Level of engagement with the problem ;
- *Approach* – Strategy for solving the problem, basic outline of the argument, and setup;
- *Development* – Implementation of strategy, quality of mathematical steps in the argument, and integration of ideas in the solution.
- *Results/Interpretation* – Outcome of the calculation and interpretation of the answer.

Exercises are categorized in three groups with the following point values and scoring rubric:

Basic (0–2 points):

0: Missing, not credible effort, or wholly incorrect approach.
1: A credible effort with a basically correct approach. May be missing parts and may have non-trivial errors.
2: Correct approach with a solid development. Basically correct result and/or sound interpretation.

Intermediate (0–4 points):

0: Missing, not credible effort, or wholly incorrect approach.
1: Credible effort but incorrect approach.
2: A credible effort with a basically correct approach. May be missing parts and may have non-trivial errors.
3: Correct approach with a mostly correct development that captures the key concepts and nuances/constraints in the problem.
4: Strong approach and development, result/interpretation correct up to minor typos or calculation snafus. No missing parts.

Advanced (0–6 points):

0: Missing, not credible effort, or wholly incorrect approach.
1: Credible effort but incorrect approach.
2: Partly incomplete, solid approach with some problems
3: A credible effort with a basically correct approach. May have missing parts and may have non-trivial errors.
4: Correct approach with a solid development that captures most of the main ideas, nuances, and constraints in the problem.
5: Correct approach and development, basically correct result and/or sound interpretation. May have a few minor errors or minor missing

pieces.

6: Strong approach and development, result/interpretation correct up to minor typos or calculation snafus. No missing parts.

Exams. There will be two in-class exams and a final exam. Exams will consist of a variety of question types consistent with the style of homework questions but fitting within the time constraints of the exam.

Final Grade. Your final grade will be calculated by a weighted average with the following proportions:

25% Homework
20% In-class exam with the lower score
25% In-class exam with the higher score
30% Final Exam (and quizzes, if any)

Target boundaries for letter grades are 90% for A, 80% for B, and so forth. We may adjust these boundaries if needed, but such adjustments will only be in your favor.

Policies

Communication. Please include "[218]" in the subject line of all email to the instructors and TAs. This enables us to process your message as efficiently as possible, and without this, messages tend to get lost. Also, please be considerate in all your communications and interactions with instructors, course staff, and your fellow students.

Missed Exams. If for some emergency circumstance you need to miss class on an exam day, you should ask for approval at least 24 hours in advance. If you have no choice but to schedule a job interview, for example, you should check with us as soon as possible. If you have a medical emergency with no advanced notice, you should bring us appropriate documentation that you visited a medical facility as soon as possible. (You are not expected to share any details of the visit nor any personal or private information.) We will consider such requests on a case by case basis.

Exam and Homework Regrading. If for any reason, you have a concern about the grading of an exam or homework, you may submit a regrade request (through Gradescope) in the time window *from 24 hours through 72 hours* after you receive the graded work. The procedure is as follows:

1. Log into Canvas and click on Gradescope in the left-side navigation bar.
2. Navigate to the specific problem at issue. This will display the scoring for that problem.
3. Click the "Request Regrade" button in the bottom action bar.
4. Put your *specific* concern in writing in the text box provided.
5. Repeat steps 2–4 for each problem at issue.

Note that for each problem regraded, the graders will reconsider the entire problem.

Collaboration and Academic Integrity. In this course, for all exams and quizzes and for homework exercises specifically designated as "Solo" all work is to be done completely on your own. That is, you may not consult with any one other than the instructors or TAs, and you may not use any external sources (other than the course notes for homework).

For homework exercises not designated "Solo," an appropriate amount of collaboration is allowed, but it is important that every student gets practice working on these problems, if for no other reason than you will have to do them alone on the exams. Many students have questions as to what constitutes reasonable collaboration and use of external sources. The following policies on collaboration and use of external sources provide guidance; if in doubt, you should feel free to ask the instructors.

Acceptable collaboration or use of external sources includes:

- Clarifying ambiguities, errata, or vague points in class materials or assignments.
- Discussing or explaining the general class material.
- Providing assistance with system facilities, computing tools, or online interfaces.
- Discussing the assignments to better understand what is being asked.
- Looking up background material (online or in books) on general concepts discussed in class.
- Discussing general approaches to solving specific problems, though see below.

Unacceptable collaboration or use of external sources includes:

- Copying of another student's solution to a problem (in part or whole) or obtaining a solution from an outside source (including a similar or related problem in part or whole).
- Allowing someone else to copy your solution in part or whole.
- Receiving help from students who have taken this or a related course in previous years.
- Communicating or having communicated (e.g., by seeing, speaking, pantomime) to you the specific steps of a solution.
- Reading the posted solution if you will be submitting your assignment late.
- Reviewing any course materials from this or related courses in previous years.

In general, all work must be written up individually, and no student should ask for assistance from any other student or offer assistance to any other student until that student has made a serious effort to solve the problem.

Cheating, inappropriate collaboration, or improper use of external sources can be grounds for course failure. We may be obliged in these situations to report the incident to the appropriate University authorities. Please refer to university policies at

`https://www.cmu.edu/policies/student-and-student-life/academic-integrity.html`

Feel free to come talk to us if you have any questions about this.

Accommodations for Students with Disabilities. If you have a disability and have an accommodations letter from the Disability Resources office, we encourage you to discuss your accommodations and needs with us as early in the semester as possible. We will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, we encourage you to contact them at `access@andrew.cmu.edu`.

Use of Mobile Devices. Research on learning shows that unexpected noises and movement automatically divert and capture people's attention. Thus, if your cell phone, pager, laptop, etc. makes noise or

is visually distracting during class, you are affecting everyones learning experience. For this reason, we allow you to use your laptop for taking notes and doing class-related activities (e.g., through ISLE), but we insist that you turn the sound off so that you do not disrupt other students' learning. Moreover, we strongly discourage doing anything other than class-related activities on your laptop. If you do so, please sit in the back row so that other students are not distracted by your screen.

Student Recording of Class. No student may record any classroom activity without express written consent from the instructors. If you have (or think you may have) a disability such that you need to record or tape classroom activities, you should contact the Office of Disability Resources to request an appropriate accommodation.

Wellness. We fully recognize the pressures and work loads of a Carnegie Mellon academic schedule. And while it can seem like devoting all your time to work is optimizing success, we believe – and research supports – that maintaining some balance is better for both your health and your academic performance. To that end, we encourage you take care of yourself and to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax.

All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful. If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at `http://www.cmu.edu/counseling/`. Consider reaching out to a friend, faculty, or family member you trust for help getting connected to the support that can help.

Policy Updates. Updates to policies and course information will be posted in updated versions of this syllabus and announced on the course homepage.

Last Modified: 26 Aug 2018