Probability Theory for Computer Scientists (36-218)

Fall 2020

Lectures: Tuesdays & Thursdays, 09:50–11:10 am ET, Zoom
Recitation: Thursdays, 09:10–10:00 pm ET, Zoom

Instructor: Mikael Kuusela (`mkuusela@andrew.cmu.edu`)
Instructor office hours: Announced on Canvas

Teaching assistants:

- Mike Stanley (Head TA, `mcstanle@andrew.cmu.edu`)

- Siddhaarth Sarkar

- Paul Twijukye

- Huiyi Guo

- Shihua Pei

- Kexiong Shen

- Caitlin Chou

TA office hours: Announced on Canvas

Website: https://canvas.cmu.edu

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 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 large and small, through case studies and activities in class, as well as through mathematical analysis and computational simulation.

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 expectations 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.

Course outline

This course covers the following topics:

- Foundations of probability

- Expectation

- Independence

- Distributions (univariate, multivariate)

- Elementary stochastic processes

- Tail bounds

- Limiting distributions

- Prediction

- Conditioning

- Graphical models

- Markov chains

- Statistical estimation and inference

- Regression

- Classification

A tentative schedule (subject to change) is provided on the course Canvas page.

Resources

Course text

The lecture slides for each class, complete with annotations made during class, will be available through ISLE (see below). You can also access and annotate these slides online yourself during class and save them for later reference. In addition, notes describing concepts and examples in more detail will be available on the Canvas site.

The annotated lecture slides, supplemented with the notes, collectively comprise the text for the course.

There is no required textbook. If you wish to consult additional materials, we recommend applicable sections from the following optional references:

- S. Ross, *A First Course in Probability*, 10th edition, Pearson, 2018 (or other editions from recent years)

- P. Whittle, *Probability via Expectation*, 4th edition, Springer, 2000 (electronic version available for free through the CMU library)

- K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press, 2012 (electronic version accessible through the CMU library)

Canvas

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

Zoom

We will be using Zoom for lectures, recitations and office hours. The Zoom links will be available on Canvas. *Do not share the links with anybody outside our class.* Please make sure that your Internet connection and equipment are set up to use Zoom and able to share audio during class meetings. You are not required use video in this class. Let the instructor know if there is a gap in your technology setup as soon as possible so we can see about finding solutions.

During our class meetings, please keep your microphone muted unless you are sharing with the class or your breakout group. If you have a question or want to answer a question, please use the Zoom chat box or the "raise hand" feature (available when the participant list is pulled up). I will be monitoring these channels in order to call on students to contribute.

All lectures will be recorded via Zoom so that students in this course (and only students in this course!) can watch or re-watch past class sessions. Office hours and most recitations will not be recorded. We will make the lecture recordings available on Canvas as soon as possible after each class session. Please note that *you are not allowed to share these recordings with anybody outside our class.* This is to protect your FERPA rights and those of your fellow students.

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.

At the end of each ISLE session, there will be two short questions designed to help you digest the class materials. You will have 36 hours after the end of each lecture to answer these questions. Answering these questions satisfactorily will constitute the *Participation* component of your course grade (see the *Grading* section below).

You should receive an automated email about signing up to ISLE at the beginning of the semester. Once you receive the email, follow the instructions there to complete the sign-up process. If you join the class after the beginning of the semester, you may need to manually sign up and enroll in the class on ISLE. To do this, go to `https://isle.stat.cmu.edu/` to create an account and then enroll by clicking on the "Enroll" button and then selecting "36-218". When signing up, **use your Andrew email address as a username.**

For each class, there will be a separate link in Canvas for the day's lecture notes.

Technical questions about ISLE should be directed to Philipp Burckhardt (`pgb@andrew.cmu.edu`).

Piazza

Piazza will be used for questions and discussion on the class contents. The link to the Piazza group is given on the course Canvas page. You should expect a reply in between one to two

working days, but not faster.

Piazza can be a very successful medium for helpful, class-wide discussions, but without rules, discussions can also quickly get out of hand. Here are the rules for our Piazza group:

1. Be considerate to others (respectful language, no sarcasm).

2. Before posting a question, check that it (or a related question) has a not already been posted. If it has, then use the existing thread for further questions/discussion. *Read Piazza first, then post!*

3. Along with your posted question, explain what you tried to answer your own question (without posting solutions).

Content deemed inappropriate—by the above rules or otherwise—may be taken down by the instructor or the TAs at their discretion.

Gradescope

In this course, we will be using Gradescope (as a plug-in tool within Canvas) to submit, 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 regrade 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. In addition to traditional scanners, scans can be made with iOS and Android devices. We have prepared a website for students

<p align="center"><code>https://www.cmu.edu/teaching/gradescope/</code></p>

that provides more information on using Gradescope, including how to scan assignments via iOS and Android devices and how to submit assignments once scanned. Let the instructor know as soon as possible if you do not have access to a device that allows you to scan your work for submission to Gradescope.

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 the system.

- 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, send email to `canvas-help@andrew.cmu.edu`.

- Important grading policy: *If the grader cannot read your submission, there is no way to award points.* So please make sure that you submit a high-quality scan of your work.

Recitations

The recitation sessions will be lead by the head TA. They provide additional examples and materials as well as an opportunity to go over selected parts of the course materials. The aim of these sessions is to specifically clarify difficult, confusing or unclear parts of the materials so feel free to ask questions and let the course staff know if there are specific parts of the course where extra clarifications would be helpful. Recitations will also be used to review materials before exams.

Office Hours

The instructor and TAs will hold office hours during the week using Zoom. These offer a terrific opportunity for you to deepen your understanding of the material by asking questions and going over problems. You are encouraged to visit office hours if you are having problems with an assignment. Also, keep in mind that office hours are useful for getting questions answered more generally. An up-to-date office hour schedule is available on the course Canvas page.

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 is really what we care about, and we are more concerned with the thought process than the final answers.

Homework exercises will be a blend of mathematical work, modeling, and occasionally, simulation. Problem sheets are assigned weekly. There will usually be 4 exercises in a problem

sheet and each exercise will be worth 6 points. A few exercises will be open-ended questions for you to answer based on a mathematical model or computer simulation of a random system.

You are allowed and encouraged to discuss homework assignments within certain limits with other students in this class unless otherwise noted; however, *you must write up your homework solutions on your own.* See the section on Academic Integrity below for more details, and feel free to ask the instructor if you have any questions.

The purpose of the homework is to give you practice with the techniques and ideas we cover in 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 below. **Your lowest two homework scores will be dropped from the calculation of your final 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 course homepage on most Fridays and will be due one week later. The submission deadline will be Fridays at midnight Pittsburgh time. Homework is to be submitted electronically through Gradescope as described in the *Resources* section. The Friday midnight deadline is strict to allow the TAs to grade the homework in reasonable time. In exceptional circumstances, we may consider accepting late homework submissions, but this must be discussed with the instructor *at least 72 hours before the assignment is due* and will be decided on a case-by-case basis. Otherwise, homework is not accepted beyond the submission deadline. Having a busy week or 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. Use Piazza and talk to the TAs and the instructor 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 the submission deadline. Your graded homeworks will be returned via Gradescope as soon as possible, typically approximately one week after the due date.

Exams

We will have four short mini-exams during the semester and a longer cumulative final exam during the final exam period. The length and scope of a mini-exam is approximately half of the length and scope of a traditional in-class midterm exam. All exams will be administered virtually in way that will enable participation from various times zones. See the *Grading* section for details on how these exams contribute to your final grade.

For each exam, you will be asked to sign a statement affirming that you will not cheat, plagiarize, or receive unpermitted assistance on the work that you turn in. Specific details

on these policies will be communicated before the first mini-exam; see also the section on Academic Integrity below.

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.

The homework exercises are graded on the scale of 0–6 points. The scoring rubric is as follows:

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.

This rubric may be adapted as needed for exercises with multiple parts or other complexities.

Exams. There will be four mini-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:

35% Homework (ignoring the lowest two homework scores)

25% Final exam

30% Mini-exams (ignoring the lowest mini-exam score)

10% Participation

Participation is measured using two short questions as the end of each lecture; see section on ISLE above.

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 and Email

Questions about homework and other class materials should be posted in Piazza or asked during office hours and recitations. *Email should be primarily reserved for administrative or logistic questions.* If you have an administrative question, please send it to both the instructor and the head TA.

You should not assume that emails will be responded right away. Complicated questions via email or Piazza may be deferred to office hours. We will strive to answers emails within a reasonable time, but please be advised that sending email to the instructor or the TAs does not create a responsibility or obligation to respond.

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

Missed Exams

If for some emergency circumstance you are not able to take an exam during the scheduled time window, you should contact the instructor at least 24 hours before the beginning of the exam time window to ask for alternative arrangements. 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. The regrade request must be submitted *within 5 days* of receiving 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, all work is to be done completely on your own. That is, you may not consult with anyone other than the instructor and the TAs, and you may not use external sources, other than those explicitly specified before each exam.

For homework exercises, an appropriate amount of collaboration is allowed and encouraged, but it is also important that every student gets practice working on these problems. You may 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 instructor.

Acceptable collaboration or use of external sources includes:

- Forming (socially distant) study groups to discuss homework or prepare for an exam.

- 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 a specific problem.

Unacceptable collaboration or use of external sources includes:

- Copying 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 or course materials from students who have taken this or a related course in previous years. (You are, of course, allowed to ask questions from those students who are TAs this year during their office hours or via Piazza.)

- Reading the posted solution if you will be submitting your assignment late.

- Reviewing any course materials from this or related courses in previous years.

As a general rule, *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 contact the instructor 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 the instructor 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`.

Student Recording of Class

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

Diversity, Equity and Inclusion

We must treat every individual with respect. We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community. Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Each of these diverse identities, along with many others not mentioned here, shape the perspectives our students, faculty, and staff bring to our campus. We will work to promote diversity, equity and inclusion not only because diversity fuels excellence and innovation, but because we want to pursue justice. We acknowledge our imperfections while we also fully commit to the work, inside and outside of our classrooms, of building and sustaining a campus community that increasingly embraces these core values.

Each of us is responsible for creating a safer, more inclusive environment.

Unfortunately, incidents of bias or discrimination do occur, whether intentional or unintentional. They contribute to creating an unwelcoming environment for individuals and groups at the university. Therefore, the university encourages anyone who experiences or observes unfair or hostile treatment on the basis of identity to speak out for justice and support, within the moment of the incident or after the incident has passed. Anyone can share these experiences using the following resources:

- Center for Student Diversity and Inclusion: `csdi@andrew.cmu.edu`, (412) 268-2150

- Report-It online anonymous reporting platform: `reportit.net` (username: tartans, password: plaid)

All reports will be documented and deliberated to determine if there should be any following actions. Regardless of incident type, the university will use all shared experiences to transform our campus climate to be more equitable and just.

Health and Wellness

Take care of yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available at CMU and an important part of the university experience is learning how to ask for help. Asking for support sooner rather than later is often 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.

If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:

CaPS: 412-268-2922
Re:solve Crisis Network: 888-796-8226

If the situation is life threatening, call the police:

On campus: CMU Police 412-268-2323
Off campus: 911

Policy Updates

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

Last updated: August 31, 2020