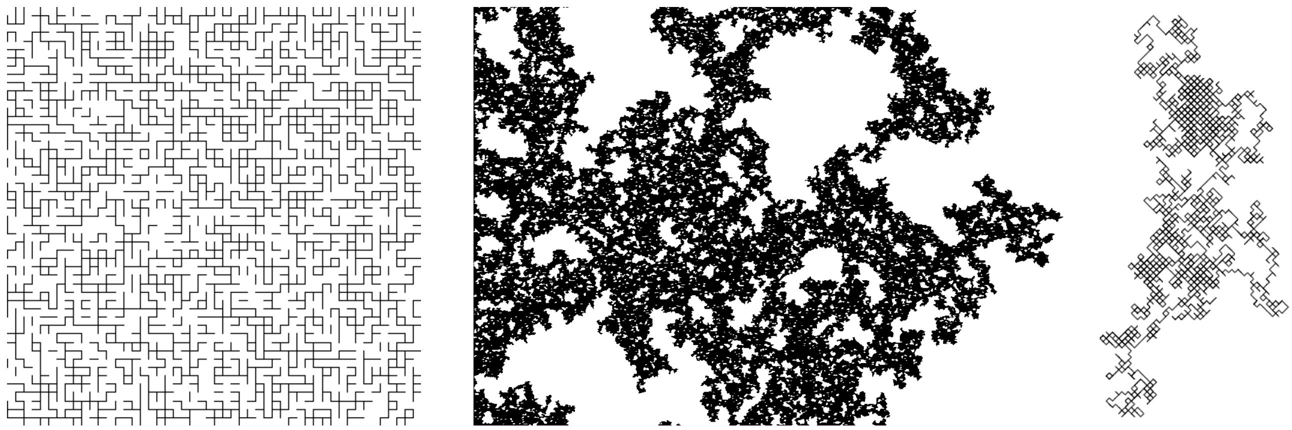
MATH 3100: INTRODUCTION TO PROBABILITY

LEONID PETROV
FALL 2023
SECTIONS 001 AND 002

1. ***A mathematical study of randomness***

How random is everything around us, and what chance do we have of understanding it? What to do when unsure, and how to do it right? How many falling stars will you see as you walk outside one beautiful night?

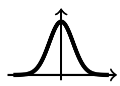
Probability theory is a mathematical study of uncertainty. It is a rigorous foundation of statistics — and many areas of human knowledge operate in a language of statistics nowadays (yes, and robots use it, too!). The course introduces fundamental concepts, ideas, and techniques of probability theory. It will provide the foundational mathematical knowledge needed to address the questions above and help you develop intuition about randomness.



Examples of random structures: bond percolation `close-up` (left), at a `larger scale` (center), and a `random walk` (see also a `simulation` of a random walk). *Note: this PDF has green clickable links, like in the previous sentence. This feature only works if you download the PDF first — it will not work in a browser on GitHub.*

What you will get from this course.

1. Mastery of basic probability concepts:
   (a) What is a probability space and how to translate commonly-sounding problems into this language;
   (b) How to count (in an advanced way) to compute probabilities;
   (c) What is a random variable, a probability distribution, and what are their main quantitative properties;
   (d) How commonly encountered probability distributions (binomial, Poisson, exponential, Gaussian) look like and behave, what are their properties, and in which situations they typically arise.

2. How large random systems behave, and what the bell-shaped curve  has to do with this.

Date: Compiled on Monday 21st August, 2023, 22:57.
An up to date syllabus is always on `GitHub` at `https://github.com/lenis2000/Syllabi/blob/master/Syllabus_3100_f23.pdf`. For direct PDF download, use `this link`. LaTeX source with *changes* to the syllabus is `here` (click "History").

3. How to describe and quantify the mutual dependence of random events, and how to use such a description to infer properties of "hidden" random events.
4. How to apply probability theory to model real-life processes. For example, how to use Bayes theorem to understand various medical tests.
5. How to collaborate on solving probability problems in pairs, small groups, and online, and present solutions clearly and efficiently.
6. In what ways probability theory is connected to science, engineering, and other branches of knowledge.

Prerequisite. You should have taken at least one semester of calculus (MATH 1320 level): a mathematical study of random variables often requires single and double integrals and infinite series.

What this course is and what it is not. This course in probability *theory* belongs to pure mathematics, with rigorous definitions, calculations, and proofs. However, the objects we study are motivated by real-life applications, so pure mathematical arguments often appeal to our common sense understanding of these objects. There will be opportunities to explore (and discover new) connections between the theory studied in the course with the real world.

Also, this course does not thoroughly discuss *applications to statistics*. Probability theory focuses on developing the mathematical side, and statistics apply these mathematical theories to real data (coming from observations). In this course, we will not discuss how to analyze data from observations — there are courses in statistics for that.

2. *Necessary information*

2.1. Meeting times.

	Section 001	Section 002
Class times August 22 — December 4	 MoWe, New Cabell 232 5:00PM - 6:15PM	 MoWe, New Cabell 323 3:30PM - 4:45PM
Midterm 1	September 25, class time	September 25, class time
Midterm 2	October 25, class time	October 25, class time
Final exam	Friday, December 15 2:00PM - 5:00PM	Tuesday, December 12 2:00PM - 5:00PM

Note on scheduling the final exams: Anyone from both sections can use any final exam times. If December 12 works for all, we will have a combined final exam that day in a bigger room. If not, we will have two final exams on December 12 and 15. I will ask you about your preferences in the middle of the semester.

2.2. Lectures+quizzes on Mondays, and problem solving on Wednesdays.

Typically, on Mondays we focus on introducing new material through lectures, while on Wednesdays, we emphasize problem-solving practice during "discussion" sessions. On Mondays, we also typically have quizzes. Refer to Section 5 for the schedule.

During discussion days, we work as a class and in groups to hone our probabilistic thinking skills. Monday lectures for both sections are recorded, posted on Canvas with notes, and have a "watch by" deadline. Wednesday sessions are not recorded.

Lectures for two different sections will be different to cover all the necessary material, so watch or review notes for both lectures each week, either in person or in recording. But do not worry if did not watch a particular lecture — come to class on Wednesday and catch up during problem solving.

2.3. Instructor.

Leonid Petrov, Kerchof 209

Questions to the instructor: `leniapetrov+f23@gmail.com`

Instructor office hours: To Be Arranged; and by appointment ↓

You can automatically schedule an office hours appointment at this page (you do not need an appointment for regular office hours). You can make as many appointments as you need throughout the semester. Each appointment scheduled online must be made at least 6 hours prior to the time of the appointment. Office hours by appointment are usually on zoom (link above); in-person option is also possible.

Teaching Assistant: To Be Arranged
TA office hours: To Be Arranged

2.4. **About the instructor.** I am an Associate Professor in the Department of Mathematics at UVA, and I've been here since 2014. My research area is probability theory (very appropriate for this course!). More precisely, I am using exact formulas to study large random systems. I also like computer simulations of random systems like `this one`. I'm happy to tell you more if you're interested.

2.5. **Textbook.** Anderson, Seppäläinen, Valkó, *Introduction to Probability*, 1st Edition.
ISBN-13: 978-1108415859; ISBN-10: 9781108415859.
See also Section 4 below for discussion of how we'll use the textbook, and for other helpful resources.

3. *Assessing your learning*

this is under construction and can change before the semester begins

Learning mathematics means *doing* mathematics: during class meetings, on your own, and in groups. In this course, doing mathematics mainly amounts to solving problems. The following aspects are assessed in this course:

3.1. **Course engagement (25%).** I am putting a lot of emphasis into the various engagement components which ask you to interact with your peers (and the instructor) while learning the material. This includes:

- **In-class participation**. Being present and involved in discussions at most of the in-class meetings is essential to getting a hands-on experience in problem solving. I understand that many people cannot participate in all class meetings due to various circumstances. I expect to give full credit for class participation if you come to at least 75% of the class meetings (not counting midterms).
- **Quizzes**. The quizzes are based on the material similar to the previous homework. Lowest 2 quiz grades our of 12 total quizzes are dropped. See the course schedule (Section 5) for quiz dates.
- **Office hours**. Come to office hours (if needed, sing up for appointments, see Section 2.1) with questions about the material. I am always happy to discuss your ideas and explain math. As a rule, I expect every student to come to office hours or schedule a one on one appointment at least once throughout the semester.

I expect that most people who are paying attention to the class, come to meetings, and interact with peers, will get close to full credit for the course engagement.

3.2. **Problem sets (20%).**

- Weekly problem sets consist of textbook and other problems aligned with lectures and in-class discussions, to help you practice new concepts and techniques. The written solutions are typically due on Tuesdays at 10pm, and sometimes on Thursdays (see the schedule in Section 5). Problem sets are posted to Canvas assignments about one week in advance.
- You are encouraged to work together on problem sets. Group work allows to take advantage of challenge-defend discussions which help understand things better. However, each student needs to submit her/his own written work, and should write this up individually. This helps better retain the material and prepare for tests.
- When solving homework problems, use your math and common sense understanding to check for your own mistakes, see Section 4.4 for details.
- The work *must be submitted only on Canvas* — take pictures or scan your work, make sure it's readable, put it into a *single PDF file with correct orientation*, and upload it to the Canvas assignments before the deadline. Failure to make a single PDF might result in zero points for a work (first, a warning will be given).

The homeworks are graded *for participation only.* That is, I will typically not read all of your work, but will look for typical mistakes to address in class. In your solution submitted to Canvas you need to attempt most of the problems — otherwise, credit for the homework might not be given (there is no partial credit for a single homework). However, it is rare that a homework will not be accepted for credit.

To get full 10% for the homework grades, you need to submit 10 solutions (for which I give credit) out of 12 total problem sets.

3.3. **Midterm tests (2 tests each worth 15%; 30% in total).** There are two midterm tests held during regular class times, on **September 27** and **October 27**. They have similar taste as homework and quizzes, and test basic knowledge of the material.

"Cheat sheet". A two-sided letter size formula sheet, hand-written by yourself, is allowed on each midterm test and on the final exam. Preparing this formula sheet will help you review the material, and paint a systematic picture of the material in your mind. In general, formula sheets cannot contain any photocopied or printed material — do everything by hand (of course, you can include any theorems, formulas, pictures, examples, etc). One exclusion: if you write the formula sheet out on a tablet and print, this is allowed, if helps — but don't copy formula sheets from other people.

I encourage you to collaborate on preparing for the tests, but needless to say that during the test and the final exam each student must work individually.

When solving midterm problems, use your math and common sense understanding to check for your own mistakes, see Section 4.4 for details.

3.4. **Final exam (25%).** The final exam will be cumulative, but will put more focus on topics covered after the second midterm. Formula sheet is also allowed, same as for midterms.

On the final exam, use your math and common sense understanding to check for your own mistakes, see Section 4.4 for details.

Letter grades. The scale by which course percent grades are turned into course letter grades will most likely be the following:

Grade	$A+$	A	$A-$	$B+$	B	$B-$	$C+$	C	$C-$	$D+$	D	$D-$
Minimum %	100	93	89	86	82	79	76	72	69	66	62	59

I reserve the right to slightly change this grade scale after the final exam. This may be needed to better incorporate into the letter grade possible fluctuations in the difficulty level of midterms and the final.

4. *How to succeed in the course*

If you have read the long syllabus through here, you are on the right track to succeed!

4.1. **General things.** The best way to learn in the course is to come to all classes, read the textbook, and do all the homework problems on your own or in collaboration. This will prepare you well for quizzes and tests.

4.2. **Main textbook.** The textbook *Introduction to Probability* by Anderson, Seppäläinen, and Valkó is an excellent resource to gain understanding of the course material. Some notes about it:

- I strongly encourage you to read the textbook in parallel with the lectures, as lectures are largely based on the textbook. The textbook includes many examples and extra exercises which augment the concepts discussed in class.
- The textbook contains much more material than will be covered in classes, so it makes sense to come to classes to note which parts are omitted (and so won't be in tests and quizzes).

4.3. **Additional textbooks.**

(1) *"Probability" by Jim Pitman* is a reasonable alternative textbook.
(2) *Free textbook "Introduction to Probability" by Grinstead and Snell.* Download: `https://math.dartmouth.edu/~prob/prob/prob.pdf`; Accompanying web page: `https://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html`.

These textbooks contain additional problems and material. They may be helpful if you want a deeper understanding of some concepts, or if you want to read an exposition of the familial material in a different style, which might be very helpful for better learning.

(It absolutely not required that you buy or read these books.)

4.4. **Problem solving and self-checking.** Solving problems, one can easily make arithmetic mistakes, and this is understandable. However, we are doing probability theory, so some answers to problems may be easily dismissed as wrong based on common sense. The most obvious examples are getting a *negative probability*; *probability strictly greater than 1*; getting a *negative variance*; and so on. It is expected that you use this type of common sense to filter out obviously wrong answers. Solid partial credit will be given even in the case when you get an obviously wrong answer, and note near it:

> "Well, this answer is clearly incorrect because {expanation}".

Obviously wrong answers without such a note will result in much less partial credit for the whole problem.

4.5. **Extra reading.** The popular book *"How Not to Be Wrong: The Power of Mathematical Thinking" by Jordan Ellenberg* discusses how math touches every aspect of real life, and has numerous examples related to probability and statistics. I can recommend this nice book as a parallel reading. Some examples I learned from this book might be mentioned in class. (It absolutely not required that you buy or read this book.)

4.6. **Office hours.** I am available during office hours to answer questions on the content of the course, clarify various points, and I can also help you with homework assignments. Besides regular office hours, you can schedule appointments online, see Section 2.

4.7. **Math Collaborative Learning Center.** The Math Department Collaborative Learning Center is available for helping students in this course: see `https://math.virginia.edu/undergraduate/MCLC/` for more information and schedule. (Usually it takes a week or so after the semester starts for MCLC to start working.)

4.8. **Other resources.** There is a number of online resources which may help you while doing the homework: Khan Academy, Wikipedia, and many other places contain lots of basic material on probability theory. Google Search in general is also a valuable resource.

4.9. **Collaboration on homework assignments.** Group work on homework problems is allowed and strongly encouraged. Discussions are in general very helpful and inspiring. Class meetings will also contain ample time for group work on homework problems. Nevertheless, before talking to others, get well started on the problems, and contribute your fair share to the process.

When completing the written homework assignments, everyone must write up his or her own solutions in their own words, and cite any reference (other than the textbook and class notes) that you use. Quotations and citations are part of the Honor Code for both UVa and the whole academic community.

It is very important that you truly understand the homework solutions you hand in, otherwise you may be unpleasantly surprised by your quiz and test results.

5. *Approximate course schedule*

Add/drop information: `see here`.

> The course has 3 "pillars": central limit theorem for Gaussian approximation, Poisson processes, and conditional expectations. Plus there are several technical things to learn: random variables, expectations as integrals, joint distributions, etc.

All sections are from the main textbook (see Section 2.5).

[**week 1**] 8/23. Introduction. Sample space, axioms of probability, random sampling, review of counting, infinitely many outcomes. (`1.1-1.3`).
[**week 2**] 8/28[Q1], 8/30. No problem set due. Hypergeometric sampling. Infinitely many outcomes. Geometric series. Rules of probability, Venn diagrams. Random variables (first look). (`1.3-1.5`).
[**week 3**] 9/4[Q2], 9/6. *PS1* due on Tuesday at 10pm. Conditional probability, Bayes' formula, independence. (`2.1-2.3`).
[**week 4**] 9/11[Q3], 9/13. *PS2* due on Tuesday at 10pm. Independent trials, birthday problem, conditional independence, probability distribution of a random variable (`2.4-2.5, 3.1`).
[**week 5**] 9/18[Q4], 9/20. *PS3* due on Tuesday at 10pm. Probability distribution of a random variable, cumulative distribution function (`3.1-3.2`).
[**week 6**] 9/25, 9/27. **Midterm 1, September 25.** No problem set due, practice problems posted a week before the midterm. Cumulative distribution function, expectation. (`3.3-3.4`).
[**week 7**] 10/4[Q5]. *PS4* due on Thursday at 10pm. Expectation and variance. (`3.3-3.4`).
[**week 8**] 10/9[Q6], 10/11. *PS5* due on Tuesday at 10pm. Gaussian distribution, normal approximation, law of large numbers, applications of normal approximation (`3.5, 4.1-4.3`).
[**week 9**] 10/16[Q7], 10/18. *PS6* due on Tuesday at 10pm. Poisson approximation, exponential distribution, Poisson process (`4.4-4.6`).
[**week 10**] 10/23, 10/25. **Midterm 2, October 25.** No problem set due, practice problems posted a week before the midterm. Poisson process, gamma distribution (`4.6`).
[**week 11**] 10/30[Q8], 11/1. *PS7* due on Tuesday at 10pm. Joint distributions (`6.1-6.3`). Sums of independent random variables and related topics (survey of selected material of chapters `7` and `8`).
[**week 12**] 11/6[Q9], 11/8. *PS8* due on Tuesday at 10pm.
[**week 13**] 11/13[Q10], 11/15. *PS9* due on Tuesday at 10pm. Law of large numbers, central limit theorem (`9.1-9.3`).
[**week 14**] 11/20. *PS10* due on Tuesday at 10pm. Conditional distributions (`10.1-10.3`).
[**week 15**] 11/27[Q11], 11/29. No problem set due. Conditional distributions (`10.1-10.3`).
[**week 16**] 12/4. *PS11* due on Tuesday at 10pm. Repair credit problem set due on Thursday at 10pm. Conditional distributions (`10.1-10.3`), final exam practice problems.

[Qx] means quiz number x and *PS*y means problem set number y, where $x, y = 1, 2, \dots, 11$.

6. *Policies*

6.1. **Late/make up work.** Each homework assignment will have due date and time by which it must be submitted to Canvas. After the 1-hour grace period, late assignments are not accepted. There will also be no make up for the quizzes or midterm tests. However, if you have special needs, emergency, or unavoidable conflicts, please let me know as soon as possible, so we can arrange a workaround for midterms and the final exam.

6.2. **Special needs.** All students with special needs requiring accommodations should present the appropriate paperwork from the Student Disability Access Center (SDAC). It is the student's responsibility to present this paperwork in a timely fashion and follow up with the instructor about the accommodations being offered. Accommodations for midterms or final exams (e.g., extended time) should be arranged at least 5 days before an exam.

6.3. **Honor Code.** The University of Virginia Honor Code applies to this class and is taken seriously (in particular, see Section 4.9 on homework collaboration). Any honor code violations, especially in written tests (midterms and the final exam) will be referred to the Honor Committee.