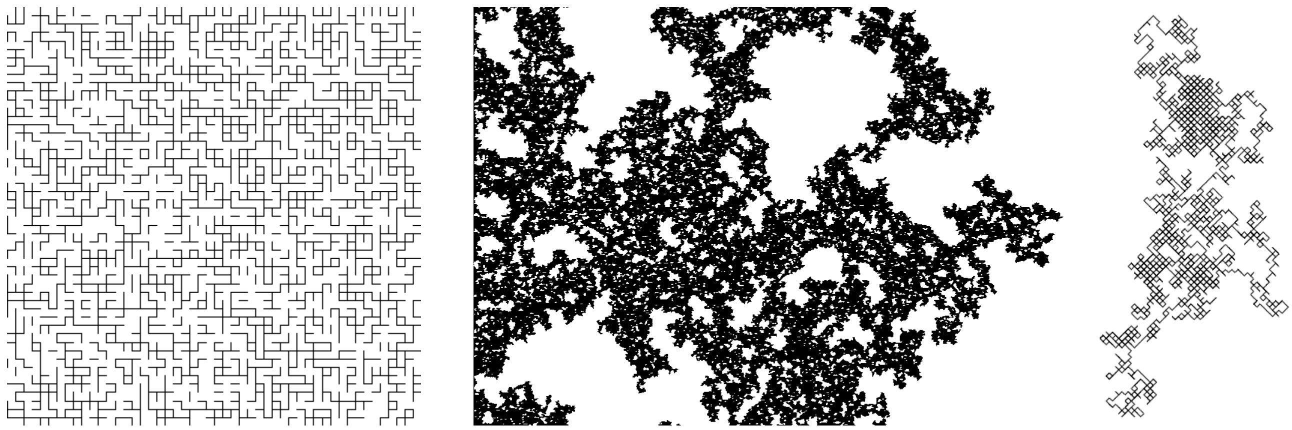
MATH 3100: INTRODUCTION TO PROBABILITY

LEONID PETROV
FALL 2022
SECTION 001

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 you're not certain, and how to do it right? How many falling stars will you see as you walk outside one beautiful night?

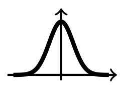
Probability theory is a mathematical study of uncertainty. It is a rigorous foundation of statistics — and many areas of human knowledge operate in a language of statistics nowadays (yes, and robots use it, too!). The course introduces fundamental concepts, ideas, and techniques of probability theory. It will provide you with the foundational mathematical knowledge needed to address the questions above and will 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 won't work in 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 Sunday 21st August, 2022, 22:13.
An up to date syllabus is always on `GitHub` at `https://github.com/lenis2000/Syllabi/blob/master/Syllabus_3100_f22.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): 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 which we study are motivated by real-life applications, and so pure mathematical arguments often appeal to our common sense understanding of these objects. There will be opportunities to explore (and discover new) connections of 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 applies these mathematical theories to real data (coming from observations). In this course we will not discuss how to analyze data coming from observations — there are courses in statistics for that.

2. *Necessary information*

2.1. Meeting times.

	Section 001
Class times August 23 — December 6	Tue-Thu, New Cabell 323 12:30PM - 1:45PM
Midterm 1	TBD
Midterm 2	TBD
Final exam	Monday, December 12 2:00PM - 5:00PM

Instructor: Leonid Petrov, Kerchof 209
Questions to the instructor: We use `Slack`, see Section 4
Office hours: Mondays at 2:30PM-3:30PM on zoom `https://virginia.zoom.us/j/97731277583?pwd=UFNvZ0NHNWRRaHpPTGYrTnJiZ3Rpdz09`; Wednesdays at 2:30PM-3:30PM in-person in Kerchof 209; and by appointment ↓

You can automatically schedule an office hours appointment at `this page` (you don't 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.

2.2. 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.

2.3. Textbook.

Anderson, Seppäläinen, Valkó, *Introduction to Probability*, 1st Edition.
ISBN-13: 978-1108415859; ISBN-10: 9781108415859.
See also Section 5 below for discussion of how we'll use the textbook, and for other helpful resources.

3. *Assessing your learning*

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 (15%).

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:

- **Slack**. Ask questions about the class, and answer others' questions.
- **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. If needed, you can occasionally come to meetings of another section (see Section 7.1 for restrictions on that). I expect to give full credit for class participation if you come to at least 75% of the class meetings (not counting midterms).
- **Quizzes**. Prior to each class, you are expected to have watched the corresponding recorded lecture. (Each lecture on Panopto is dated with the date you should watch it by.) There will be occasional pop quizzes in class, to make sure you are paying attention to the material. Lowest x quiz grades is dropped from the calculation (here $x = 1$ or 2, depending on the total number of quizzes we end up having).
- **Office hours**. Come to office hours (if needed, sing up for appointments, see Section 2.1) with questions about problem sets and recorded lectures. 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 (30%; lowest problem set grade is dropped).

> **Note added 2022-03-24**: Problem set 8 (due March 28) is counted towards participation grade and not the homework grade. If you skip Problem set 8, no grade is harmed.

- Weekly problem sets consist of textbook and other problems aligned with lectures, to help you practice new concepts and techniques. The written solutions are due on Mondays at 10pm (except weeks with midterms). Problem sets are posted to Collab assignments about one week in advance, except for the first HW1 (which is mostly introductory). You will be working on them during the discussion times in class meetings.
- You are encouraged to work together on problem sets in the off-times, too, in slack (you may create private groups there; ask me if you need more permissions), and by any other means of communication, or in person. 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.

- The written solutions are graded "coarsely", that is, each problem set will be assigned one of four grades:

Grade	VG (very good)	G (good)	OK	N
	All problems solved correctly with minor issues like arithmetic mistakes, and solutions explained in full detail	Most problems solved correctly, and solutions explained in reasonable (close to full) detail	More than 3/4 of problems attempted, many solutions are incorrect, incomplete, or not explained in detail, but the work displays adequate understanding of most of the material	Work not submitted on time, or less than 3/4 of problems attempted, or most solutions are incomplete, or work clearly displays lack of understanding of most of the material
%	100%	90%	75%	0%

  It is expected that most students who put reasonable effort into the work will get VG or G grades.
- The work *must be submitted only on Collab* — 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 Collab assignments before the deadline. Failure to make a single PDF might result in zero points for a work (first, a warning will be given).

When solving homework problems, use your math and common sense understanding to check for your own mistakes, see Section 5.4 for details.

3.3. **Midterm tests (2 tests each worth 15%; 30% in total).** There are two midterm tests held during regular class times, on **February 18** and **March 30**. They have similar taste as homework, and test basic knowledge of the material.

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 5.4 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 5.4 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. *Communication*

4.1. **Email.** My email address is petrov@virginia.edu. This is a backup channel for communication in case slack or collab fails. In general, you have to ask your mathematical questions (and questions about the course) on slack.

4.2. **Slack.** This is the space to ask (and answer!) questions about problem sets, lectures, and anything related to the course. The aim of slack is to build a community of learners of probability theory.

Slack is an industrial standard of work messengers, with a web version and apps for all platforms. This will make me more accessible if you have questions, and also will let you answer questions of your fellow students.

The link to sign up to the slack space for the course is found on collab. Please let me know if you have issues with access. It is also expected that you will check announcements, and will participate in (or at least read) most of the discussions of the course material.

Some things to note:

- **Population**: The `Slack` team will contain of students of all 3 sections of MATH3100 this semester, which is around 120 students.
- **Direct messages**: Direct messages are an excellent tool in `Slack` to communicate with me. When sending your very first direct message to me, please write your full name and which section you're from (if this is not obvious from your nickname).
- **Direct messages to other students**: There are direct messages where you can talk to other students one-on-one. You can also create private groups with up to 9 people, which is good for homework collaboration (but read Section 5.10 on collaboration).
- **Privacy**: Although `Slack` is a messaging app, it should be used professionally, especially in public discussions. The app supports private direct and group messages. But please note that in principle the admin (i.e., myself) can obtain access to **all** direct messages between members of the team. I assure you that this may be done only in extreme circumstances.
- **Notifications**: With `Slack`, it is easy to not miss important announcements and direct messages, even if you don't check it all the time — under the default settings, you will get notified (by email, too) about direct messages and mentions of your name. All announcements will have mentions of everybody to draw attention. Also, you can change notification settings, but please remember that there will be sometimes very important activity in `Slack` that you don't want to miss.
- **Many other features**: `Slack` supports reminders out of the box. You can ask it to remind yourself about a particular message, or just set a general reminder. With this, it is easy to keep track of important messages (like class announcements or problems I'll post there).

The slack space is separated into channels:

- #hw-01, . . . , #hw-10 for homeworks
- #midterm-1, #midterm-2 for midterms
- #final for the final exam
- #lectures for lecture announcements and questions,
- #questions for any other questions you might have, including general discussion.
- #general is for general announcements and I will keep only the most necessary things here.
- #introductions if you'd like to write a couple of words about yourself - I'm curious to meet and get to know you all! (Except #introductions, all other channels include everybody by default.)

4.3. **Collab anonymous feedback.** If you have anonymous comments on anything related to the course, you can make them via Collab.

5. *How to succeed in the course*

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

5.1. **General things.** The best way to learn in the course is to watch all recorded lectures and take notes to retain the material in memory; come to all classes; and do all the homework problems on your own or in collaboration. This will prepare you well for tests.

5.2. **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 watching 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 watch lectures and come to classes to note which parts are omitted (and so won't be in tests).

5.3. **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.)

5.4. **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.

5.5. **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.)

5.6. **Slack.** Ask mathematical questions about homework, lectures, and anything else related to the class on Slack (see Section 4.2). Answer other students' questions.

5.7. **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 automatically schedule appointments, see Section 2.

5.8. **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.)

5.9. **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.

5.10. **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 test results.

6. *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.

A summary of the lectures content is posted at `https://github.com/lenis2000/Math3100_S22_LectureNotes/blob/main/lecture_notes.pdf` (`link to automatically download`).

All sections are from the main textbook (Section 2.3):

- [**week 1**] 1/19, 1/21. Introduction. Sample space, axioms of probability, random sampling, review of counting (`1.1-1.2`).
- [**week 2**] 1/24, 1/26, 1/28. Problem set 1 (shortened) due on Monday. Infinitely many outcomes. Geometric series. Rules of probability, Venn diagrams, random variables (`1.3-1.5`).
- [**week 3**] 1/31, 2/2, 2/4. Problem set 2 due on Monday. Conditional probability, Bayes' formula, independence (`2.1-2.3`).
- [**week 4**] 2/7, 2/9, 2/11. Problem set 3 due on Monday. Independent trials, birthday problem, conditional independence, probability distribution of a random variable (`2.4-2.5, 3.1`).
- [**week 5**] 2/14, 2/16, **2/18** (**Midterm 1**). Problem set 4 due on Monday. Probability distribution of a random variable, cumulative distribution function (`3.1-3.2`).
- [**week 6**] 2/21, 2/23, 2/25. No problem set due on Monday. Cumulative distribution function, expectation, variance, Gaussian distribution (`3.3-3.5`).
- [**week 7**] 2/28, 3/2, 3/4. Problem set 5 due on Monday. Gaussian distribution, normal approximation, law of large numbers (`3.5, 4.1-4.2`). Applications of normal approximation (`4.3`).
- [**week 8**] 3/14, 3/16, 3/18. Problem set 6 due on Monday. Poisson approximation, exponential distribution, Poisson process (`4.4-4.6`).
- [**week 9**] 3/21, 3/23, 3/25. Problem set 7 due on Monday. Poisson process, gamma distribution (`4.6`). Joint distributions (`6.1-6.3`).
- [**week 10**] 3/28, **3/30** (**Midterm 2**), 4/1. Problem set 8 (optional; for participation grade) due on Monday. Joint distributions (`6.1-6.3`).
- [**week 11**] 4/4, 4/6, 4/8. No problem set due. Sums of independent random variables and related topics (survey of selected material of chapters `7` and `8`).
- [**week 12**] 4/11, 4/13, 4/15. Problem set 9 due on Monday. Law of large numbers, central limit theorem (`9.1-9.3`).
- [**week 13**] 4/18, 4/20, 4/22. Problem set 10 due on Monday. Conditional distributions (`10.1-10.3`).
- [**week 14**] 4/25, 4/27, 4/29. Problem set 11 due on Monday. Conditional distributions (`10.1-10.3`).
- [**week 15**] 5/2. Problem set 12 due on Monday. Overview of cutting edge research in Probability Theory.

7. *Policies*

7.1. **Going to another section's meeting.** I teach three sections of MATH3100 this semester, with identical content except for quizzes, tests, and exams. In general, **you should come to class at your assigned section's time** (that is, switching sections on a regular basis is not allowed). However, occasionally when you have a schedule conflict, you can come at another section's time. Occasionally means 3-5 times during the semester. There are also the following caveats:

- Coming to a different section time must not exceed classroom's capacity (capacity does not matter for zoom meetings which are till February 4).
- You must come at **your assigned section time** for all midterms and the final exam (unless you arrange with me otherwise before the midterm / exam).

7.2. **Late/make up work.** Each assignment will have due date and time. Late assignments are not accepted. There will also be no make up for the midterm test. However, if you have special needs, emergency, or unavoidable conflicts, please let me know as soon as possible, so we can arrange a workaround. A workaround may include an extra homework which would repair some of the grade.

7.3. **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 test-taking (e.g., extended time) should be arranged at least 5 business days before an exam.

7.4. **Honor Code.** The University of Virginia Honor Code applies to this class and is taken seriously (in particular, see Section 5.10 on homework collaboration). Any honor code violations, especially in written tests (midterms and the final exam) will be referred to the Honor Committee.

7.5. **Recording.** The lectures are recorded and posted to Panopto (collab → lecture capture). Per University of Virginia policies, these recordings cannot be shared outside of our course.