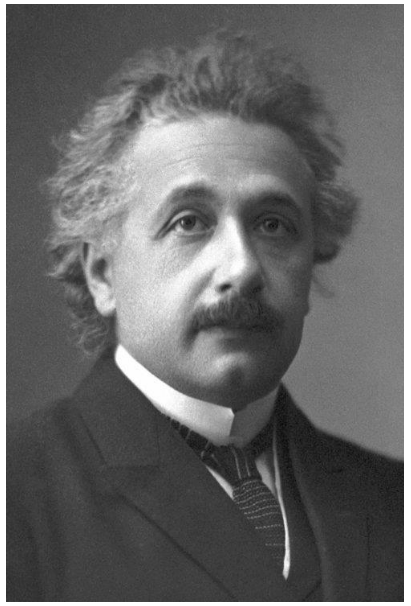
CS 237: PROBABILITY IN COMPUTING

INSTRUCTORS: TIAGO JANUARIO, SOFYA RASKHODNIKOVA

BOSTON UNIVERSITY



Albert Einstein
[1879-1955]

“God doesn’t play dice
with the universe."

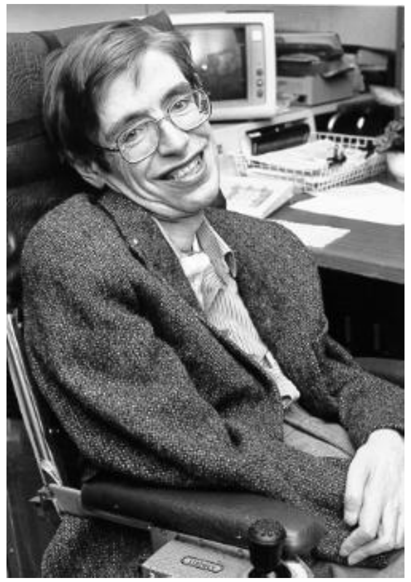
LECTURE 1

- Course information
- Introduction to probability theory
- Sample Spaces and Events
- Examples

CS 237: PROBABILITY IN COMPUTING

INSTRUCTORS: TIAGO JANUARIO, SOFYA RASKHODNIKOVA

BOSTON UNIVERSITY



Stephen Hawking
[1942-2018]

"Einstein was doubly wrong ... not only does God definitely play dice, but He sometimes confuses us by throwing them where they can't be seen."

LECTURE 1

- Course information
- Introduction to probability theory
- Sample Spaces and Events
- Examples

INSTRUCTORS

Sofya Raskhodnikova

https://cs-people.bu.edu/sofya/

Tiago Januario

https://cs-people.bu.edu/januario/

STAFF MEMBERS

- Teaching Fellows:

[Erick Jimenez]

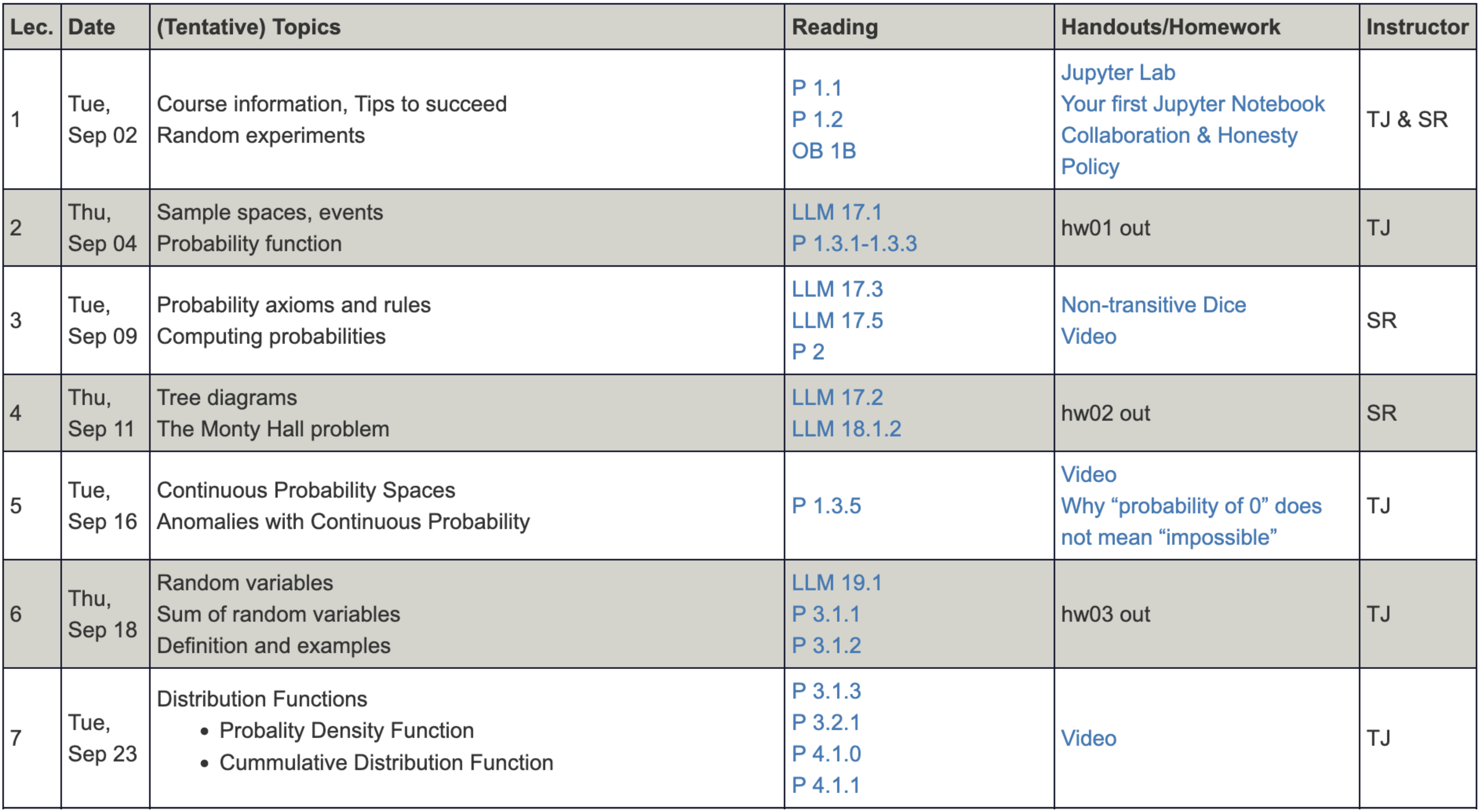
[Anatoly Zavyalov]

- Course Assistants:
  - Daniel Matuzka
  - Letitia Caspersen
  - Sarah Yuhan
  - Steve Choi
  - Vi Tjiong
  - Yoon Oh

Check the class website for OH and more

https://cs-people.bu.edu/januario/teaching/cs237/fa25/

SCHEDULE



Lec.	Date	(Tentative) Topics	Reading	Handouts/Homework	Instructor
1	Tue, Sep 02	Course information, Tips to succeed Random experiments	P 1.1 P 1.2 OB 1B	Jupyter Lab Your first Jupyter Notebook Collaboration & Honesty Policy	TJ & SR
2	Thu, Sep 04	Sample spaces, events Probability function	LLM 17.1 P 1.3.1-1.3.3	hw01 out	TJ
3	Tue, Sep 09	Probability axioms and rules Computing probabilities	LLM 17.3 LLM 17.5 P 2	Non-transitive Dice Video	SR
4	Thu, Sep 11	Tree diagrams The Monty Hall problem	LLM 17.2 LLM 18.1.2	hw02 out	SR
5	Tue, Sep 16	Continuous Probability Spaces Anomalies with Continuous Probability	P 1.3.5	Video Why “probability of 0” does not mean “impossible”	TJ
6	Thu, Sep 18	Random variables Sum of random variables Definition and examples	LLM 19.1 P 3.1.1 P 3.1.2	hw03 out	TJ
7	Tue, Sep 23	Distribution Functions • Probality Density Function • Cummulative Distribution Function	P 3.1.3 P 3.2.1 P 4.1.0 P 4.1.1	Video	TJ

https://cs-people.bu.edu/januario/teaching/cs237/fa25

Textbook: Mathematics for CS (probability part)

https://cs-people.bu.edu/aene/cs237fa21/mcs.pdf

PREREQUISITES

- CS 131: Combinatoric Structures
- MA 123: Calculus I
- CS 111: Introduction to Computer Science 1
- Good working knowledge of:
  - Sets, counting, proofs (induction, contradiction, etc.)
  - Differentiation and integration
  - Python programming

PIAZZA

- https://piazza.com/bu/fall2025/cascs237
- Use the available “folders” for each question
- Before asking your question, think about how you can frame it to benefit you and other students in the course.
- Bad questions exist and will hurt our ability to answer you efficiently.

COURSE STRUCTURE

- Lectures: Tuesdays and Thursdays
- Discussions: Fridays
- Homework: problem sets posted every week (25%)
    - Math proofs and programming problems (using Python 3)
- In-class midterm exam (30%)
    - On Thursday, October 23rd, 2025 - Location: CGS 129
- Final exam (35%)
    - On Thursday, December 18th, 2025 - Location: TBA
- Attendance (5%) and Participation (10%)

ATTENDANCE AND PARTICIPATION

Required in-class software: **Top Hat Pro platform**

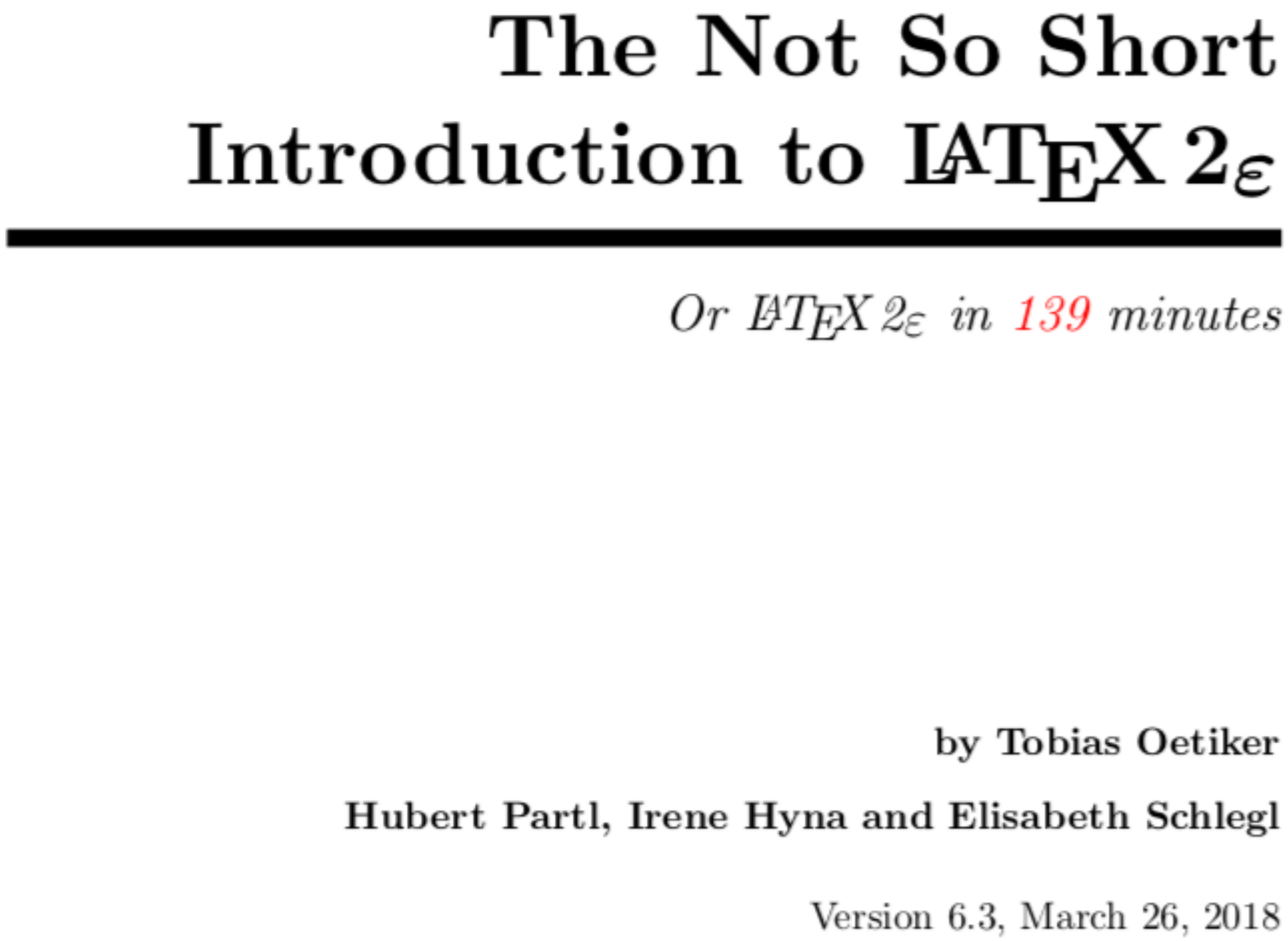
- Used for pre-lecture quizzes and in-lecture exercises for location-based attendance
- Join codes:
- A1 Section, from 2:00 pm to 3:15 pm
    - 037447
- A2 Section, from 3:30 pm to 4:45 pm
    - 606363

Join codes for discussion sections will be provided later.

HOMEWORK SUBMISSION

- Deadline: Wednesdays, at 09:00 PM
- Register at https://www.gradescope.com/courses/1066812
- Use entry code VWJD8X
- Sign up using your BU email, and include your Student ID in the format U00000000
- Use Gradescope for regrades as well

LATEX



The Not So Short Introduction to LaTeX 2_{ε}

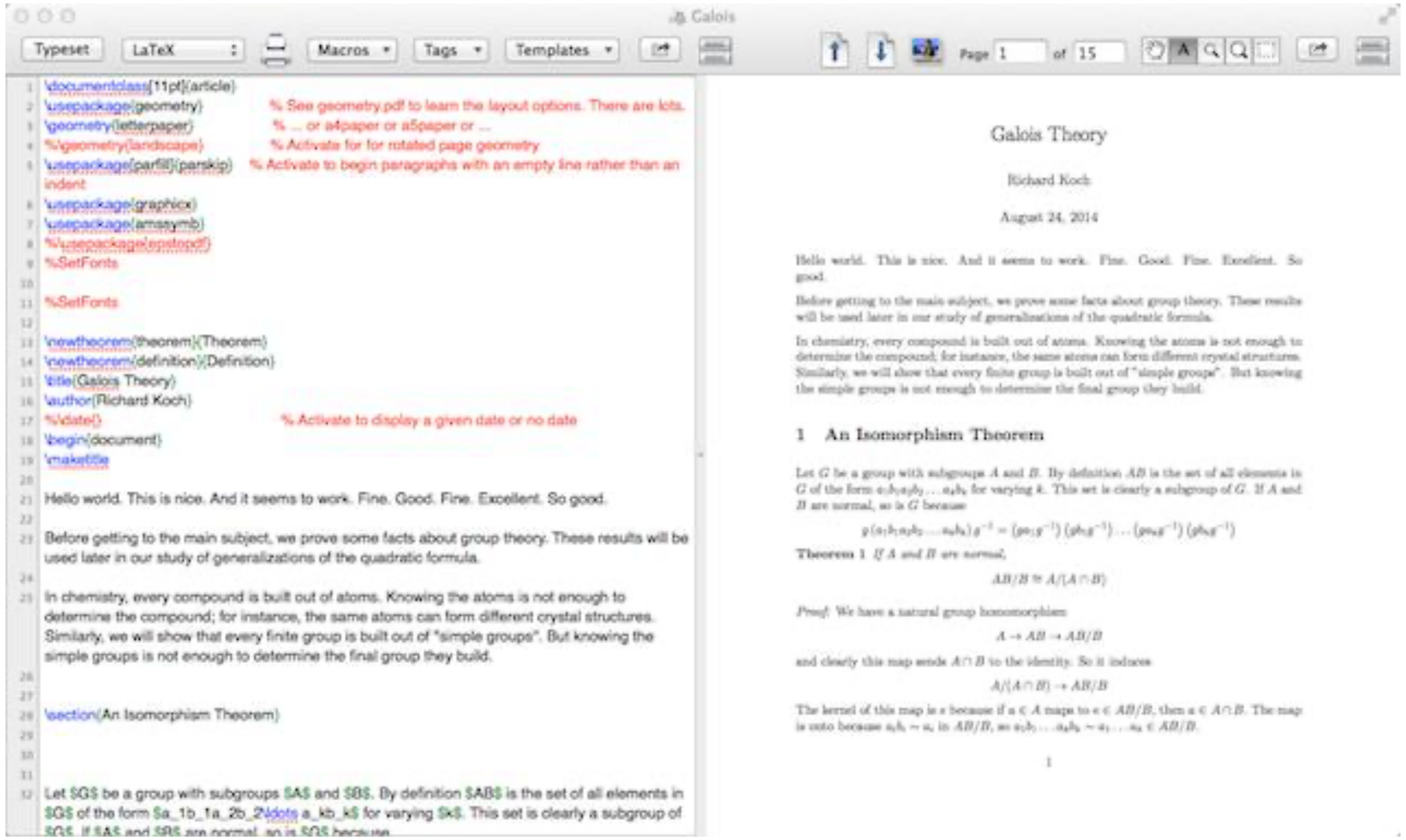
Or LaTeX 2_{ε} in 139 minutes

by Tobias Oetiker
Hubert Partl, Irene Hyna and Elisabeth Schlegl

Version 6.3, March 26, 2018

LATEX

- Many editors: TeXShop, TeXStudio, overleaf.com, …



```
\documentclass[11pt]{article}
\usepackage{geometry}                % See geometry.pdf to learn the layout options. There are lots.
\geometry{letterpaper}               % ... or a4paper or a5paper or ...
%\geometry{landscape}                % Activate for for rotated page geometry
\usepackage[parfill]{parskip}    % Activate to begin paragraphs with an empty line rather than an indent
\usepackage{graphicx}
\usepackage{amssymb}
%\usepackage{epstopdf}
%SetFonts

%SetFonts

\newtheorem{theorem}{Theorem}
\newtheorem{definition}{Definition}
\title{Galois Theory}
\author{Richard Koch}
%\date{}                              % Activate to display a given date or no date
\begin{document}
\maketitle

Hello world. This is nice. And it seems to work. Fine. Good. Fine. Excellent. So good.

Before getting to the main subject, we prove some facts about group theory. These results will be used later in our study of generalizations of the quadratic formula.

In chemistry, every compound is built out of atoms. Knowing the atoms is not enough to determine the compound; for instance, the same atoms can form different crystal structures. Similarly, we will show that every finite group is built out of "simple groups". But knowing the simple groups is not enough to determine the final group they build.


\section{An Isomorphism Theorem}



Let $G$ be a group with subgroups $A$ and $B$. By definition $AB$ is the set of all elements in $G$ of the form $a_1b_1a_2b_2\ldots a_kb_k$ for varying $k$. This set is clearly a subgroup of $G$. If $A$ and $B$ are normal, so is $G$ because
```

Galois Theory

Richard Koch

August 24, 2014

1 An Isomorphism Theorem

SUBMIT YOUR HOMEWORK THE RIGHT WAY

- You are responsible for submitting **one single PDF file** with high-quality images of your solutions.
- Illegible submissions will receive a 0 grade
- We highly recommend Dropbox to scan your homework before uploading it.
- Select the correct pages on Gradescope for each problem (solved or not) to avoid a 10% homework penalty.
- In cases where you do not have a solution to submit for a specific problem, write a brief note such as "No solution provided".

TIPS FOR THE COURSE

- Concepts in this course take some time to sink in: be careful not to fall behind.
- Prepare for each lecture by reviewing material from the previous lecture and doing assigned reading.
- Be active in lectures, discussions, and on Piazza.
- Take advantage of office hours.
- Solved exercises in Pishro-Nik’s textbook

TIPS FOR THE COURSE

- Study with a friend: do exercises and quiz each other.
- Allocate lots of time for the course: comparable to a project course, but spread more evenly.
- Start working on HW early and solve it over multiple days.
- You can work in groups (up to 3 people), but spend at least 30 minutes thinking about it on your own before your group meeting.

COLLABORATION AND HONESTY POLICY

- https://cs-people.bu.edu/januario/teaching/cs237/collaboration-policy.pdf
- Read, sign, and submit it to Gradescope by Sep 10th, 09:00 PM
- Discuss with classmates (strongly encouraged!)
- Write up in your own words, acknowledge people you worked with
- Do not share written work, and write your own code!
- Do not submit anything you cannot explain to the course staff

PYTHON

- We will use Python 3 and Jupyter Lab
- https://jupyter.org/try-jupyter/lab/

WHAT IS CS 237 ABOUT?

Fundamentals of Probability

- What is a random process and how can we model it and analyze it?
- Basic probability and events
- Conditional probability and independence
- Discrete and continuous random variables and distributions
- Expectation and variance

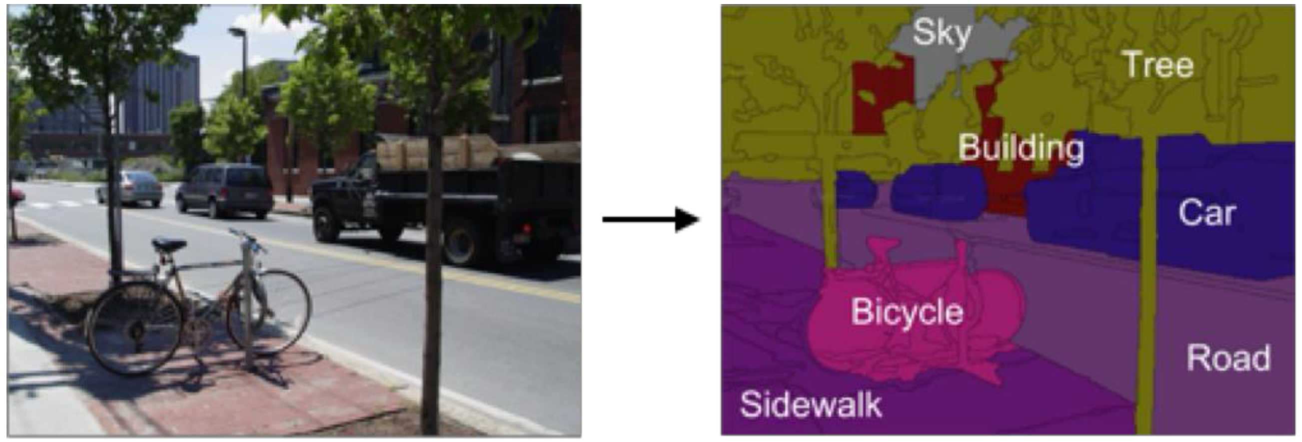
WHAT IS CS 237 ABOUT?

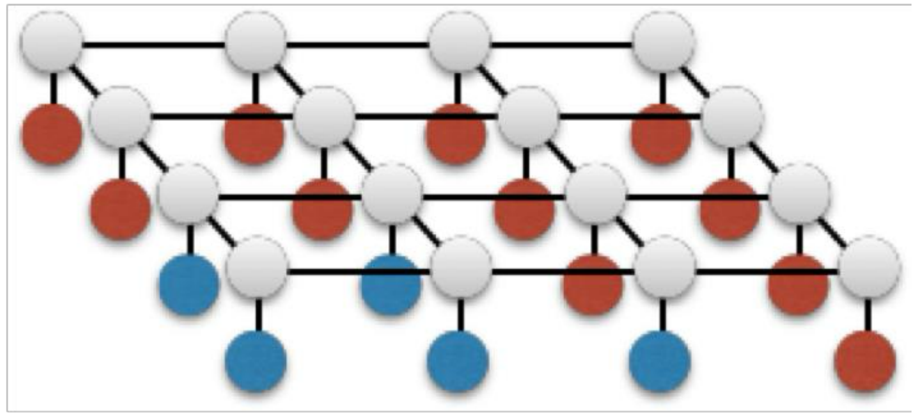
Fundamental Tools for CS

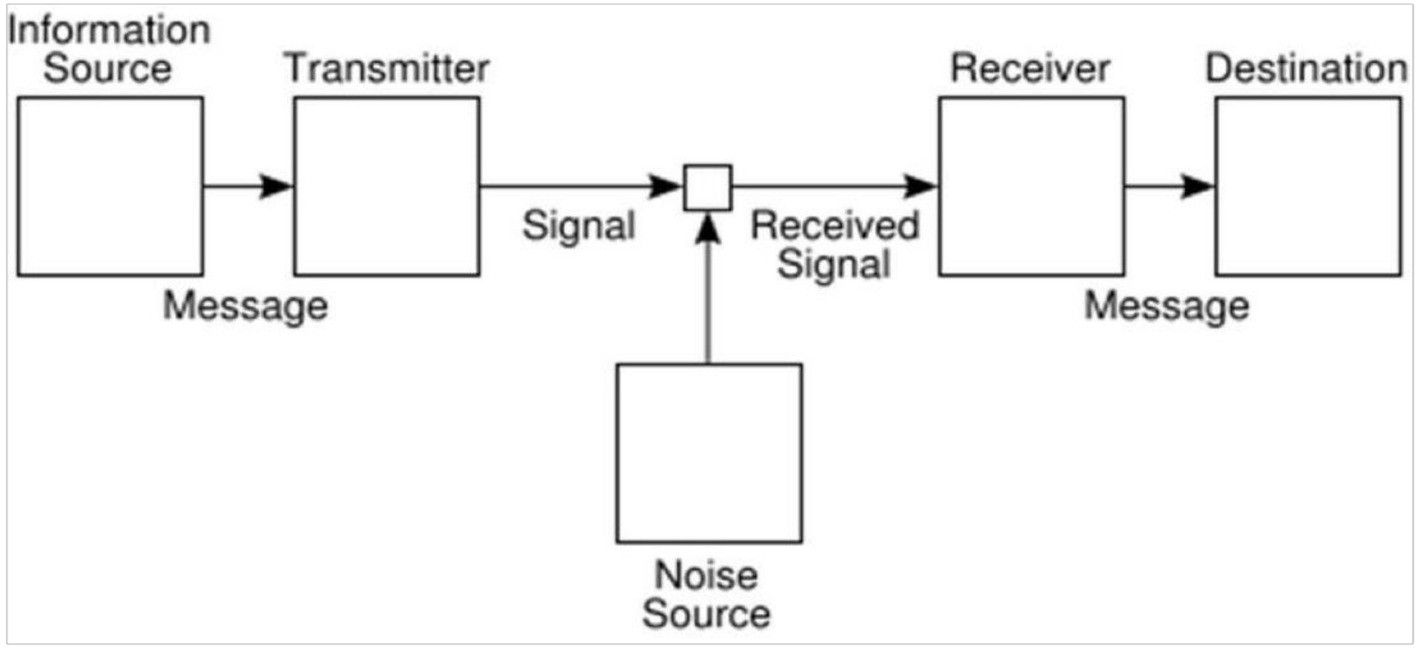
- Estimation by sampling
- Probabilistic analysis via concentration inequalities
- Probabilistic Data Structures
- Randomized algorithms

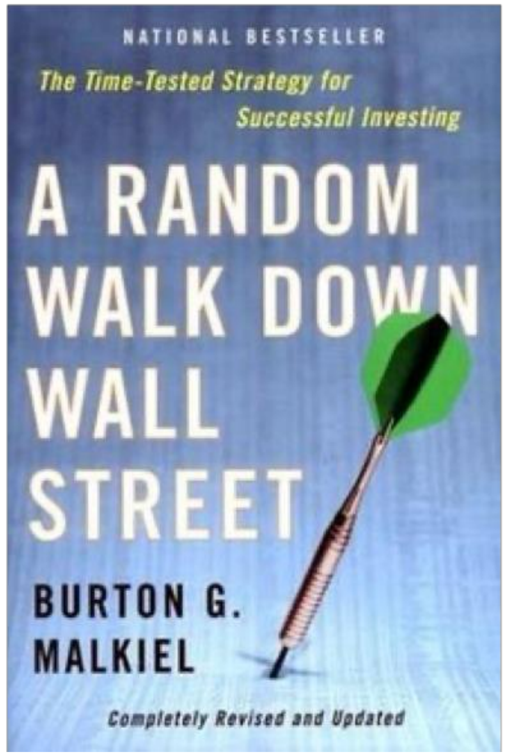
PROBABILITY IS UBIQUITOUS

Computing   Statistics   Engineering   Economics/Finance   Linguistics

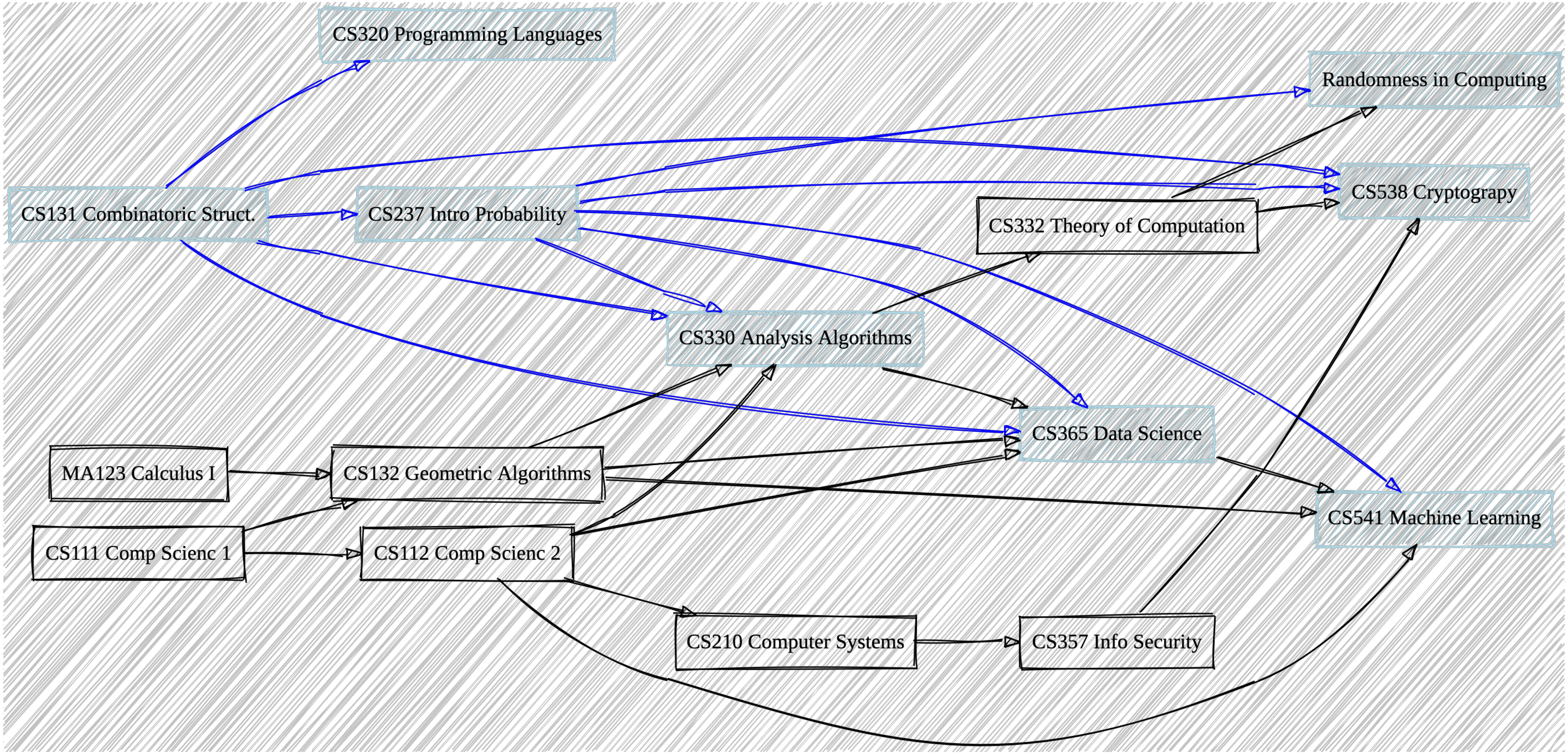








PROBABILITY IN THE CS CURRICULUM



WHAT ABOUT CS 237 MATERIAL ON JOB INTERVIEWS?

- InterviewBit
- Indeed
- Interview Query
- Nick Sigh
- ML Stack Café
- Strata Scratch
- Towards Data Science

SKILLS WE WILL WORK ON

- Expressing your ideas
  - abstractly (suppress inessential details)
  - precisely (rigorously)
- Mathematical, Probabilistic and Algorithmic thinking
- Computer simulations of probabilistic experiments
- Problem solving
- Having **FUN** with all of the above!!!

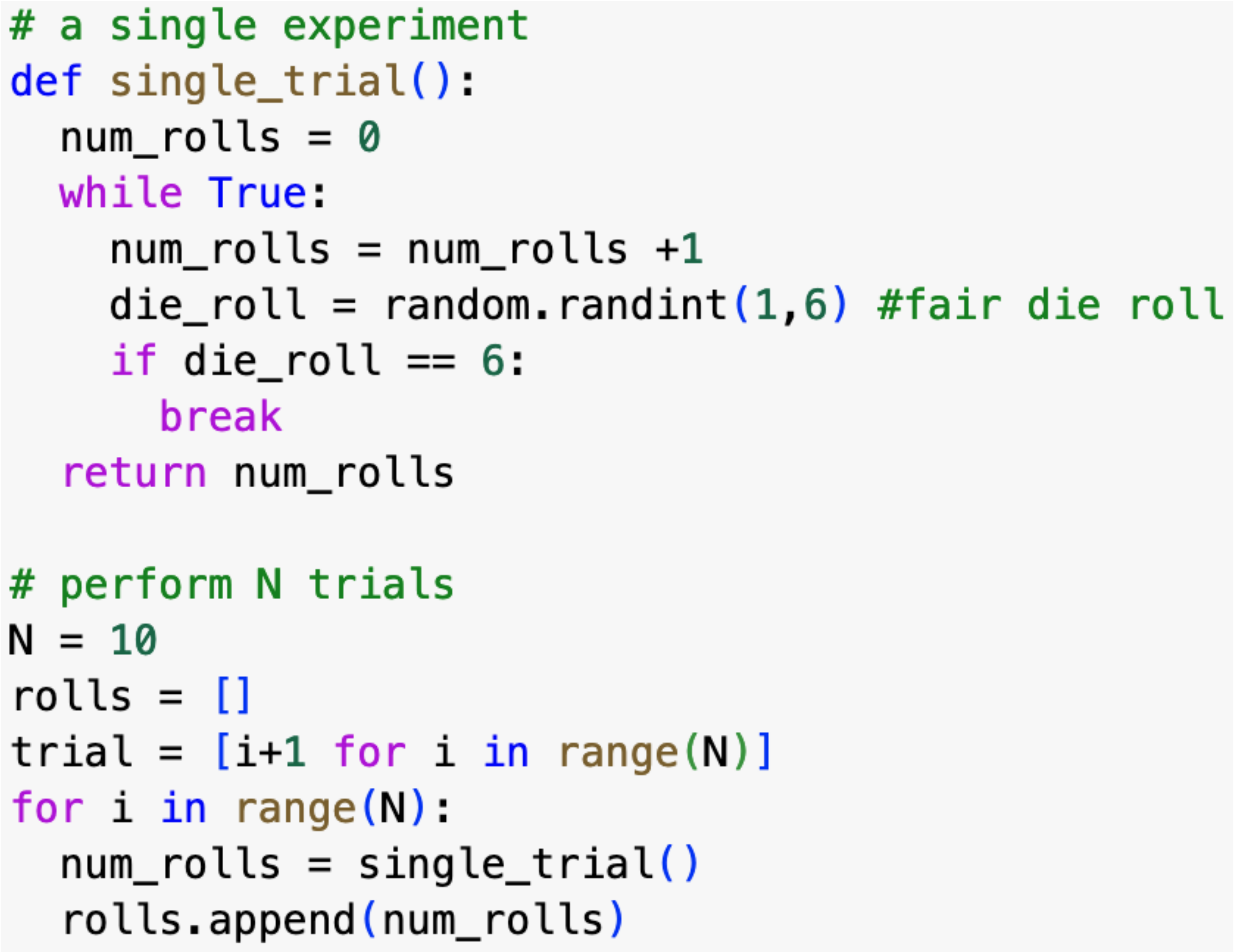
LET'S ROLL SOME DICE

- Take a fair 6-sided die

- Roll the die until we get a 6
- What is the expected number of rolls?

- Let's run an experiment using Jupyter Lab

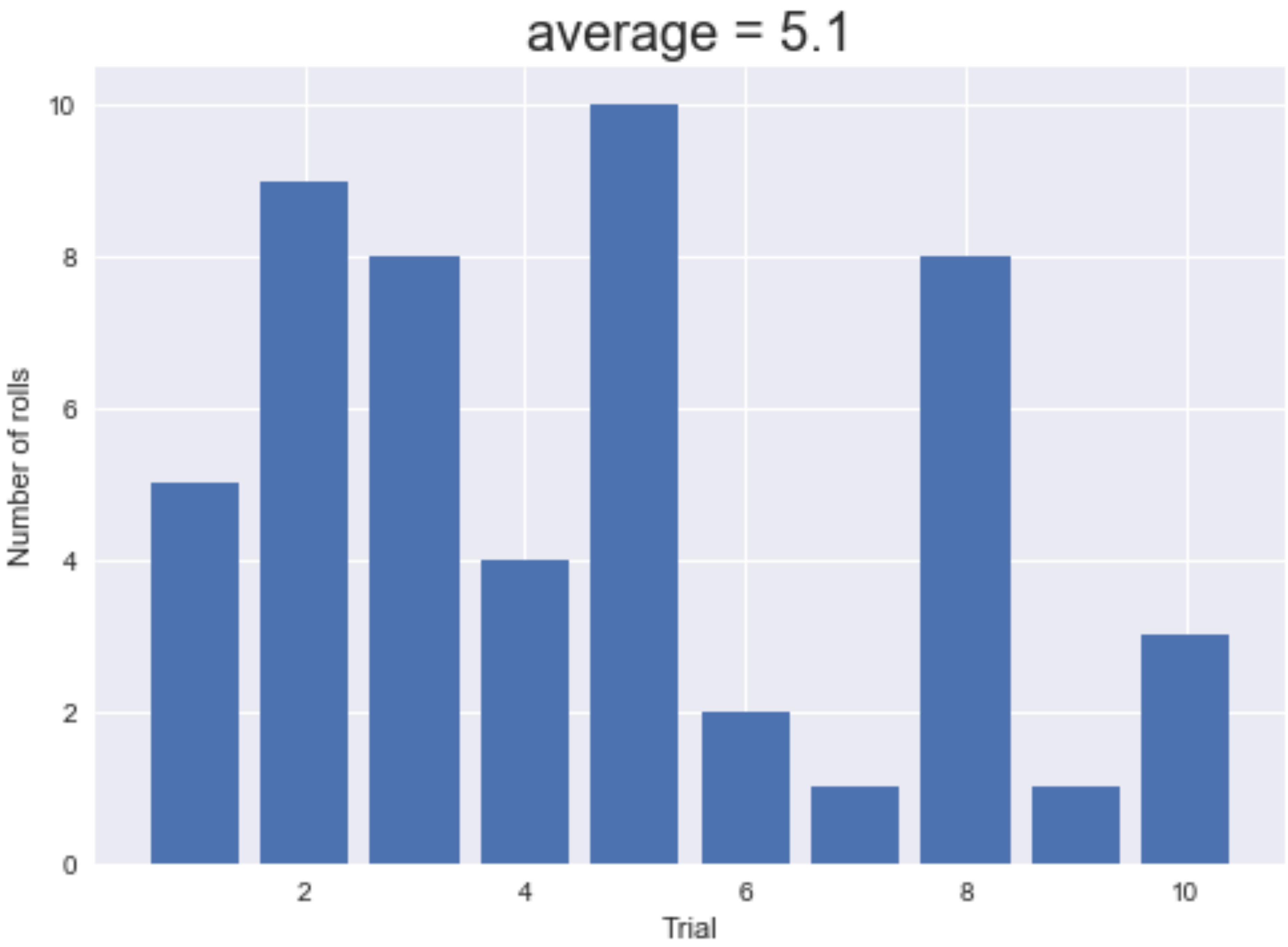
LET’S ROLL SOME DICE



```
# a single experiment
def single_trial():
  num_rolls = 0
  while True:
    num_rolls = num_rolls +1
    die_roll = random.randint(1,6) #fair die roll
    if die_roll == 6:
      break
  return num_rolls

# perform N trials
N = 10
rolls = []
trial = [i+1 for i in range(N)]
for i in range(N):
  num_rolls = single_trial()
  rolls.append(num_rolls)
```

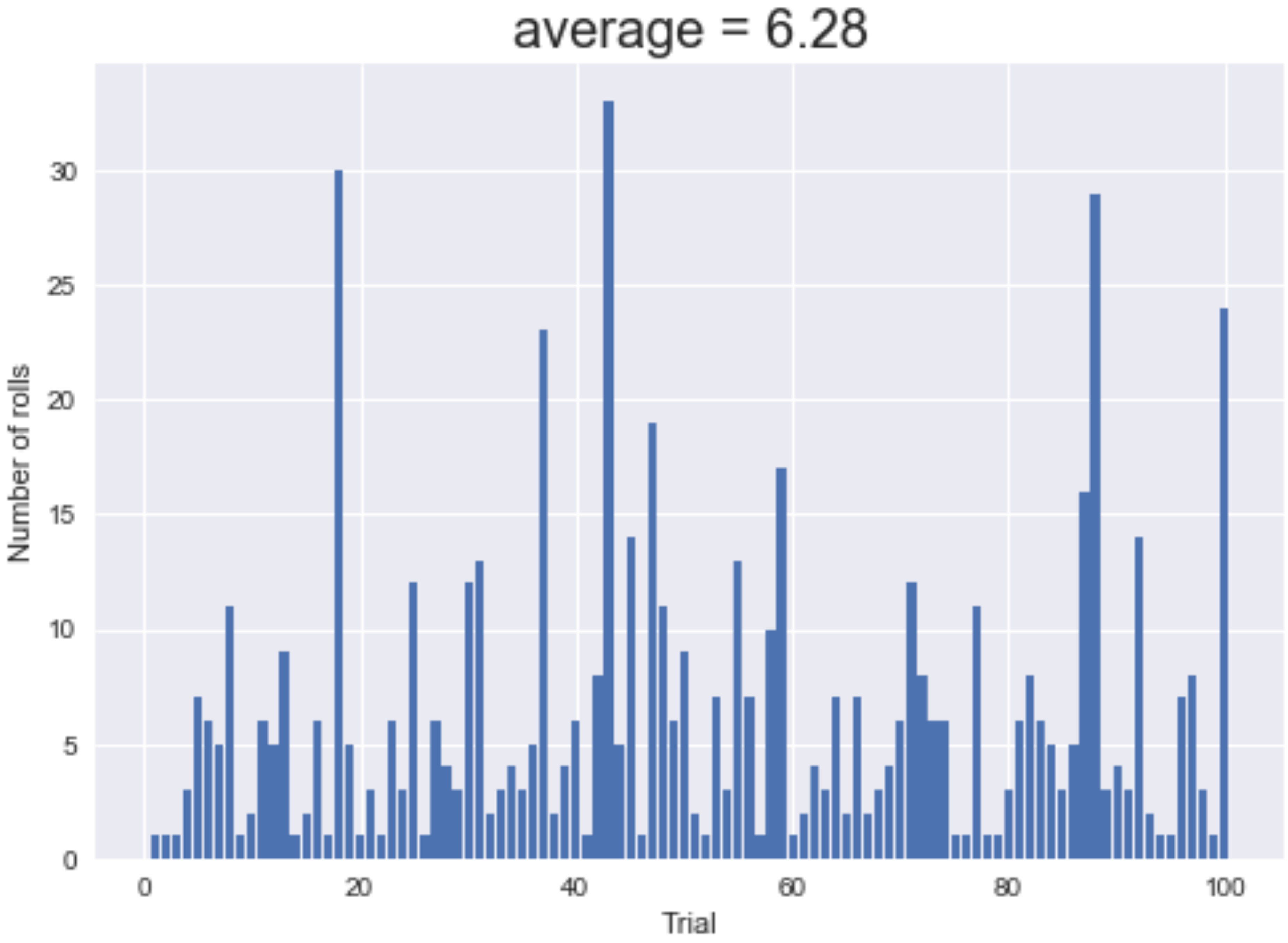
LET’S ROLL SOME DICE



10 trials

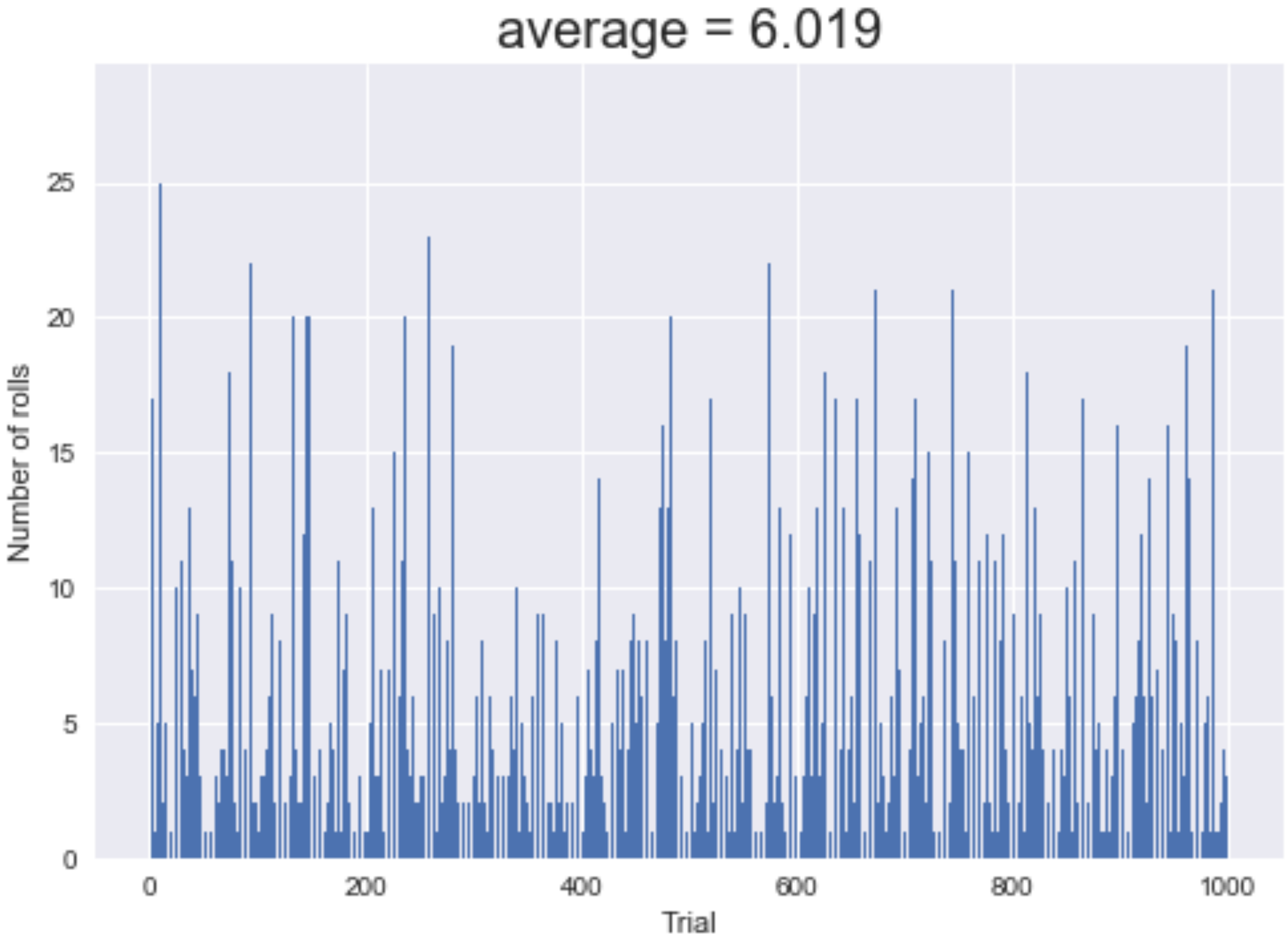
Tiago Januario, Sofya Raskhodnikova. Based on slides by Alina Ene.

LET'S ROLL SOME DICE



100 trials

LET’S ROLL SOME DICE



1000 trials

Tiago Januario, Sofya Raskhodnikova. Based on slides by Alina Ene.

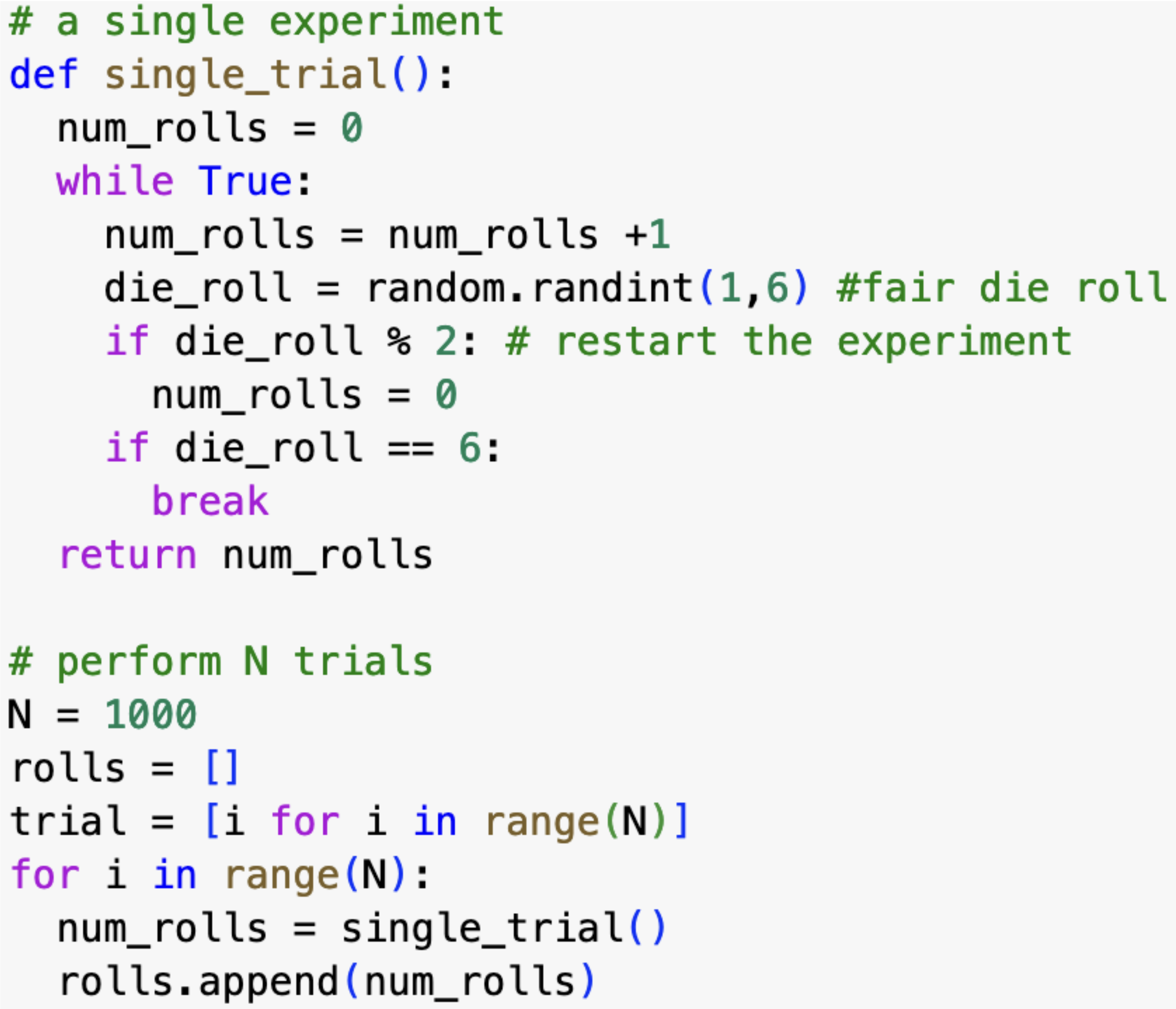
LET’S ROLL SOME DICE

- Fair 6-sided die
- Roll the die until we get a 6
- What is the expected number of rolls?
- Correct answer = 6

LET’S ROLL SOME DICE

- Fair 6-sided die
- Roll the die until we get a 6
- What is the expected number of rolls given that all rolls gave even numbers?

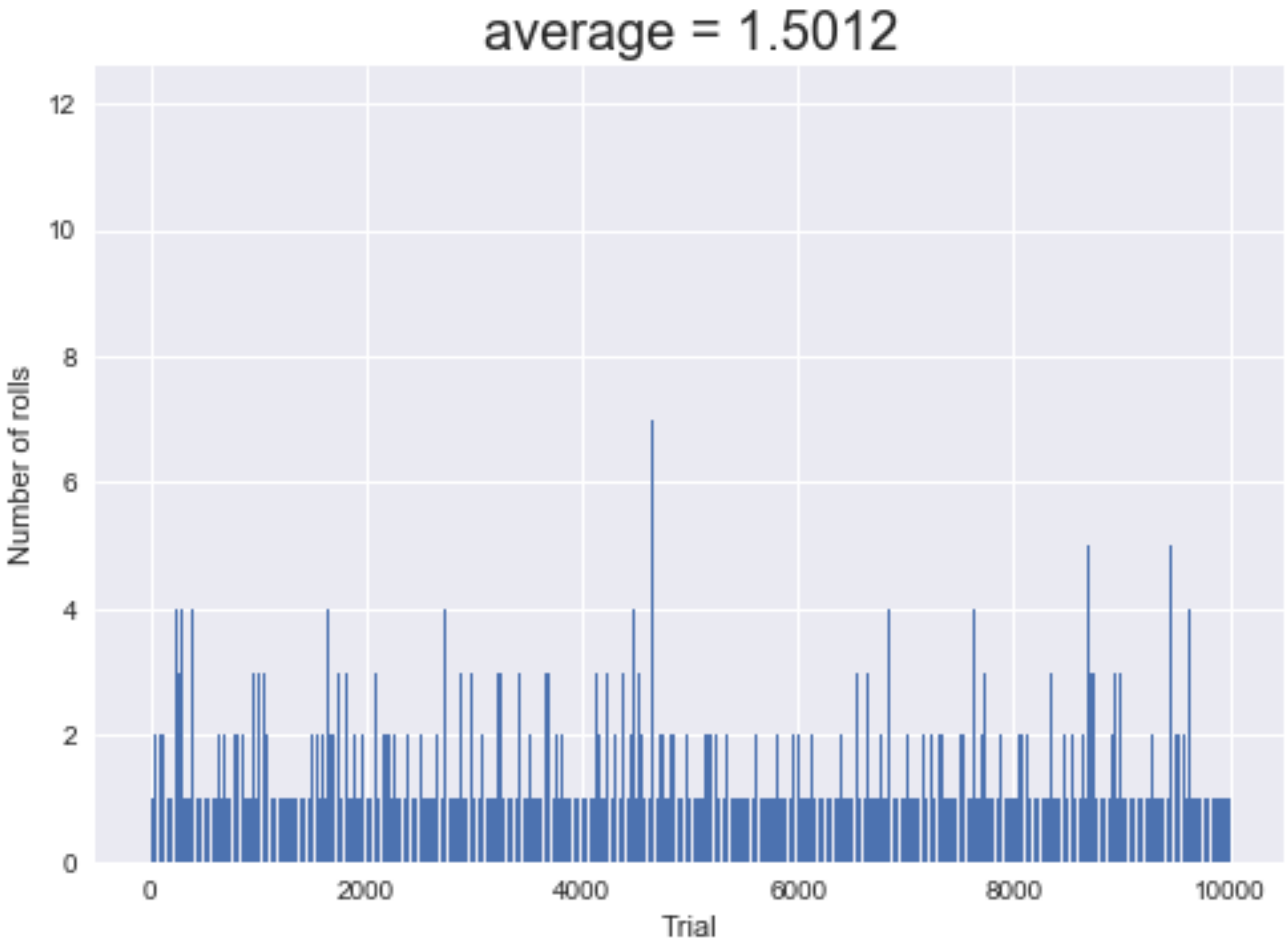
LET’S ROLL SOME DICE



```
# a single experiment
def single_trial():
  num_rolls = 0
  while True:
    num_rolls = num_rolls +1
    die_roll = random.randint(1,6) #fair die roll
    if die_roll % 2: # restart the experiment
      num_rolls = 0
    if die_roll == 6:
      break
  return num_rolls

# perform N trials
N = 1000
rolls = []
trial = [i for i in range(N)]
for i in range(N):
  num_rolls = single_trial()
  rolls.append(num_rolls)
```

LET’S ROLL SOME DICE



LET’S ROLL SOME DICE

- Fair 6-sided die
- Roll the die until we get a 6
- What is the expected number of rolls given that all rolls gave even numbers?
- Correct answer = 1.5

PROBABILITY THEORY

- Introduce self-evident and indisputable properties of probability (the axioms)
- Develop the mathematical theory of probability from these axioms

Andrey Kolmogorov

[1903 - 1987]

> “The theory of probability as a mathematical discipline can and should be developed from axioms in exactly the same way as geometry and algebra.”

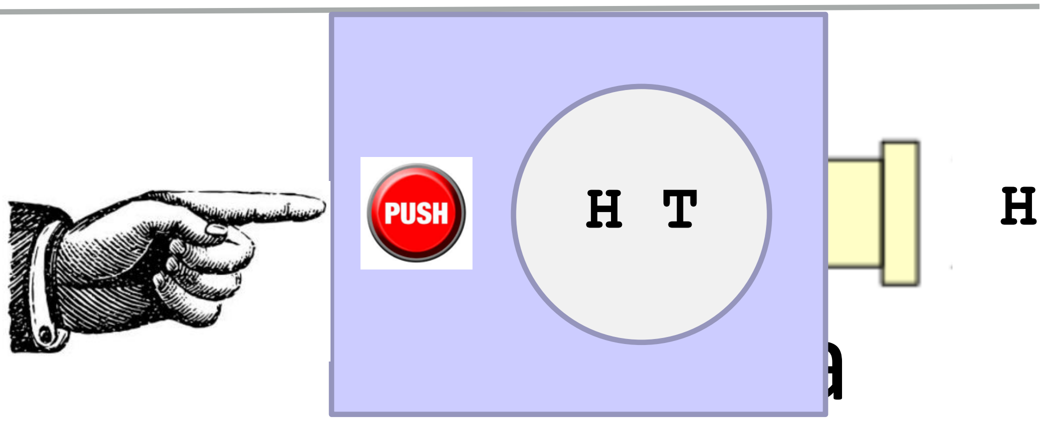
TYPICAL STATEMENTS ABOUT PROBABILITY

1. The probability that a randomized algorithm for checking polynomial identities accepts when the input is not an identity is at most 1/100.
2. The chance of getting a flush (that is, all cards of the same suit) in a 5-card poker hand is about 2 in 1000.
3. The chance of precipitation today in Boston is 20%.

Each such statement is implicitly talking about a random experiment

- either constructed by us, as in (1) and (2)
- or used by us to model the world, as in (3)

PROBABILITY: RANDOM EXPERIMENT



- Random experiment: a
  repeatable procedure
  - Toss a coin


  - Toss a coin 3 times

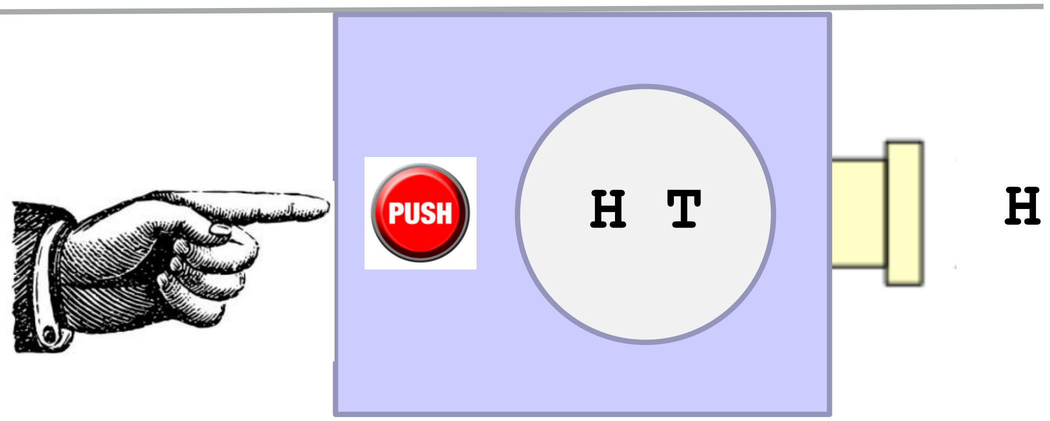

  - Roll two dice


  - Pick a 5-card hand out of a deck of cards


  - Observe the number of goals in a
    soccer match between robots


Tiago Januario, Sofya Raskhodnikova. Based on slides by Alina Ene.

PROBABILITY: SAMPLE SPACE



- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
  - Toss a coin

$$\Omega = \{H, T\} \quad , \quad \Omega = \{\text{Heads}, \text{Tails}\}$$

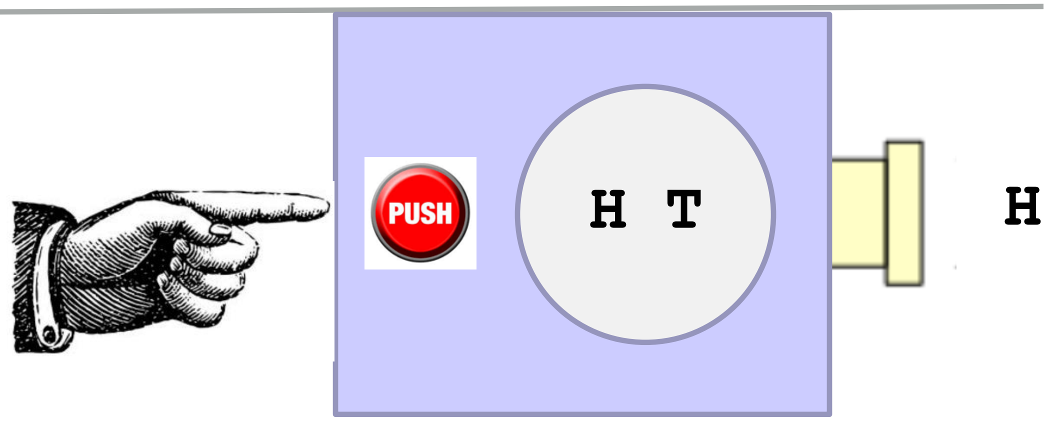
  - Toss a coin 3 times

$$\Omega = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

$$|\Omega| = 2 \cdot 2 \cdot 2 = 8$$

Tiago Januario, Sofya Raskhodnikova. Based on slides by Alina Ene.

PROBABILITY: SAMPLE SPACE



- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
  - Roll two dice

    $$\Omega = \{(i,j) \mid 1 \le i,j \le 6 \text{ and } i,j \in \mathbb{N}\}$$

  - Pick a 5-card hand out of a deck of cards (the order of cards doesn't matter, so {5♥, 5♦, 5♠, Q♥, Q♣} and {5♦, 5♥, Q♥, 5♠, Q♣} are the same hand)

    Ω is the set of all subsets of five cards, $|\Omega| = \frac{52\cdot 51\cdot 50\cdot 49\cdot 48\cdot}{5!} = \binom{52}{5} = \frac{52!}{5!\,(52-5)!}$

  - Observe the number of goals in a soccer match between robots

    $$\Omega = \{0, 1, 2, 3, 4, \ldots\}$$

PROBABILITY: EVENT

- Event: a subset of the sample space

  (that is, a set of outcomes)

  - Experiment: toss a coin 3 times

Event A: get at least 2 heads

$A = \{$ HHT, HTH, THH, HHH $\}$

  - Experiment: roll two 6-sided dice

Event B: the sum of the two numbers rolled is 11

$B = \{ (5,6), (6,5) \}$

LET’S VOTE

Experiment: toss a coin 3 times. Which of the following is the event ``exactly 2 heads”?

- $E_1 = \{HHT, HTH, THH, HHH\}$
- $E_2 = \{HHT, HTH, THH\}$
- $E_3 = \{HTH, THH\}$

A. E_1 ✗ has the outcome HHH that has 3 heads

B. [✓ E_2]

C. E_3 ✗ does not include HHT

D. Both E_2 and E_3 are correct ✗

LET’S VOTE

Experiment: toss a coin 3 times

- Event $E = \{HTH, HHT, THH\}$

Which of the following describes the event E?

A. “exactly one head” ✗ No outcome matches this description

B. [✓ “exactly one tail”]

C. “at most one tail” ✗ miss HHH

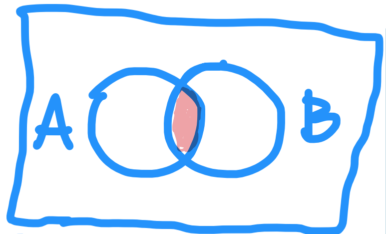
D. None of the above ✗

EVENTS

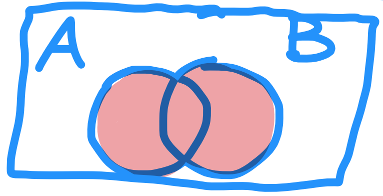
- Events are *sets* of outcomes

- We can combine events using set operations

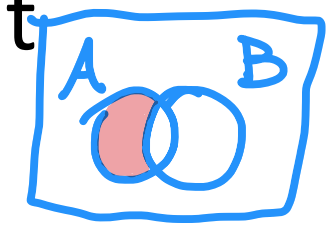
$A \cap B$: the event that both A and B occurred



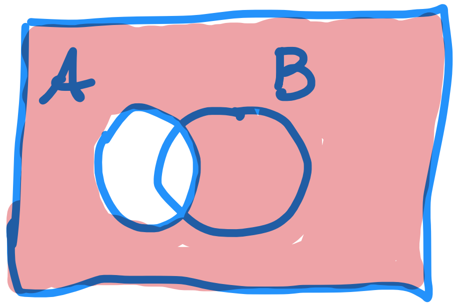
$A \cup B$: the event that A or B occurred

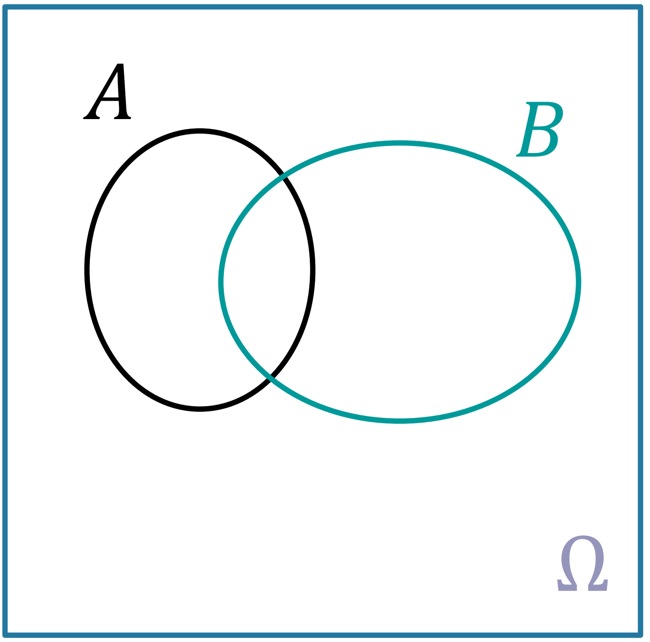


$A \setminus B$: the event that A occurred but B did not



$\overline{A}$: the event that A did not occur





Tiago Januario, Sofya Raskhodnikova. Based on slides by Alina Ene.

LET'S VOTE

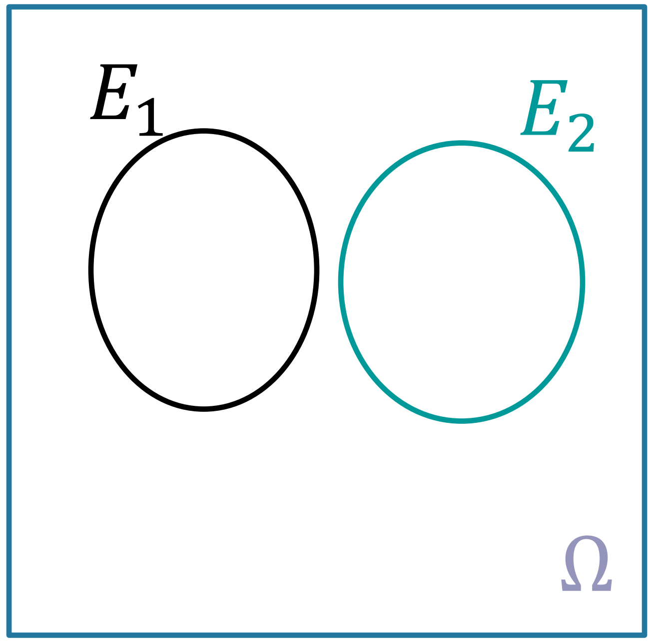
- Experiment: toss a coin 3 times

Are the following events disjoint?

E_1 = "exactly 2 heads"

E_2 = "exactly 2 tails"

$E_1 \cap E_2 = \emptyset$



A. [✓ YES]

B. NO

$\Omega = \{HHH, \underbrace{HHT, HTH, THH}_{E_1}, \underbrace{TTH, THT, HTT}_{E_2}, TTT\}$

Tiago Januario, Sofya Raskhodnikova. Based on slides by Alina Ene.

LET’S VOTE

- Experiment: toss a coin 3 times

Does the first event imply the second?

$E_1 \subseteq E_2$

E_1 = “at least 2 heads”

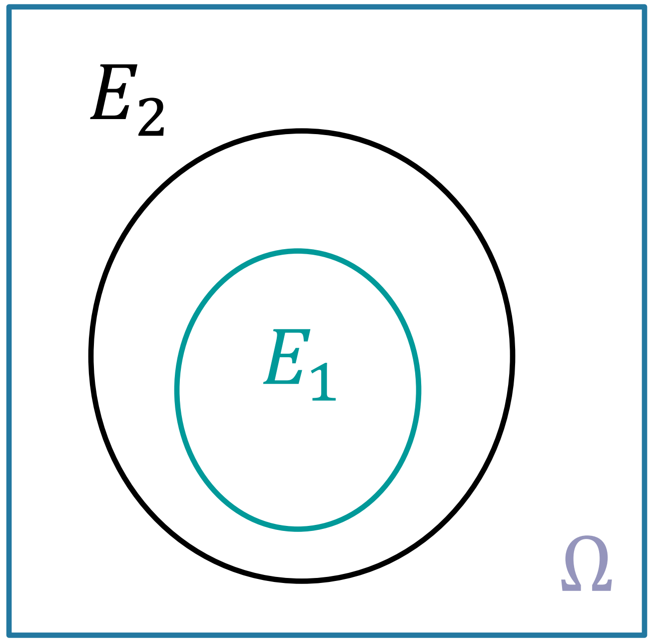
E_2 = “exactly 2 heads”

A. YES

B. [✓ NO]

$E_1 = \{HHT, HTH, THH, HHH\}$

$E_2 = \{HHT, HTH, THH\}$



PROBABILITY FUNCTION

- Each outcome in the sample space Ω is assigned a probability, which is a number greater or equal to 0.
- All probabilities of outcomes in Ω must add up to 1.

$\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$ $\frac{1}{4}$

- Probability of event E, denoted Pr(E), is the sum of probabilities of all outcomes in E.

RECAP: PROBABILITY CAST OF CHARACTERS

- Experiment: a repeatable procedure
- Outcome: result of the experiment
- Sample space Ω: set of all possible outcomes
- Event: a subset of the sample space
- Probability function $\Pr$: assigns a probability Pr(E) to each event E

CHECKLIST

- Go to the course website
- Browse course materials
- Make sure you understand the course policies
- Sign up on Piazza
- Sign up on Gradescope
- Sign up on Top Hat
- Install LaTeX and Python