Lecture 11 – Probability

DSC 10, Fall 2023

- possible remote lecture on Wed, stay tuned on ed.

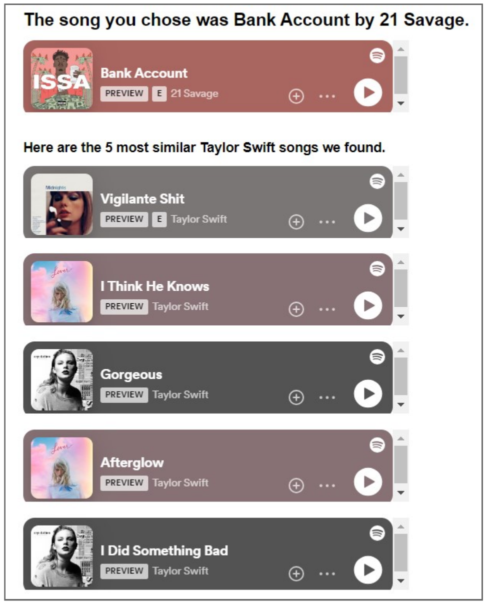
- Rod's office hours go remote effective Nov 6.

Announcements

- Quiz 2 is on **Wednesday in discussion section**.
    - The quiz covers Lectures 5 through 10.
    - Practice by solving problems from old exams at **practice.dsc10.com**.
- Lab 3 is due on **Thursday at 11:59PM**.
- Homework 3 is due on **Saturday at 11:59PM**.

Midterm Project released, due **Saturday, 11/4 at 11:59PM**

- In the project, you'll explore Taylor Swift's music and lyrics and implement some fun tools. You'll make a song recommender that suggests the Taylor Swift songs that are similar to your favorite song. Here's a sneak peek:



- **Start early!** You should be halfway done with the project by the end of this weekend.

Last time: `for`-loops

- Almost every `for`-loop in DSC 10 will use the **accumulator pattern**.
  - This means we initialize a variable, and repeatedly add on to it within a loop.
  - The variable could be an integer, an array, or even a string (as in Homework 3, Question 4: Triton Tweets).
  - Analogy: Start with a blank piece of paper and write something on it each time you run an experiment.
- Do **not** use `for`-loops to perform mathematical operations on every element of an array or Series.
  - Instead, use DataFrame manipulations and built-in array or Series methods.
- Helpful video 🎥: **For Loops (and when not to use them) in DSC 10**.

Agenda

We'll cover the basics of probability theory. This is a math lesson; take written notes ✍🏽.

Probability resources

Probability is a tricky subject. If it doesn't click during lecture or on the assignments, take a look at the following resources:

- **Computational and Inferential Thinking, Chapter 9.5**.
- **Theory Meets Data, Chapters 1 and 2**.
- **Khan Academy's unit on Probability**.

Probability theory

- Some things in life *seem* random.
    - e.g., flipping a coin or rolling a die 🎲.
- The **probability** of seeing "heads" when flipping a fair coin is $\frac{1}{2}$.
- One interpretation of probability says that if we flipped a coin infinitely many times, then $\frac{1}{2}$ of the outcomes would be heads.

Terminology

- **Experiment**: A process or action whose result is random.
    - e.g., rolling a die.
    - e.g., flipping a coin twice.
- **Outcome**: The result of an experiment.
    - e.g., the possible outcomes of rolling a six-sided die are 1, 2, 3, 4, 5, and 6.
    - e.g., the possible outcomes of flipping a coin twice are HH, HT, TH, and TT.
- **Event**: A set of outcomes.
    - e.g., the event that the die lands on a even number is the set of outcomes {2, 4, 6}.
    - e.g., the event that the die lands on a 5 is the set of outcomes {5}.
    - e.g., the event that there is at least 1 head in 2 flips is the set of outcomes {HH, HT, TH}.

Terminology

- **Probability**: A number between 0 and 1 (equivalently, between 0% and 100%) that describes the likelihood of an event.
    - 0: The event never happens.
    - 1: The event always happens.
- Notation: If A is an event, $P(A)$ is the probability of that event.

Equally-likely outcomes

- If all outcomes in event A are equally likely, then the probability of A is

$$P(A) = \frac{\#\text{ of outcomes satisfying } A}{\text{total } \#\text{ of outcomes}}$$

- **Example 1**: Suppose we flip a fair coin 3 times. What is the probability we see exactly 2 heads?

HHH THH
HHT THT
HTH TTH
HTT TTT

$P(A) = \frac{3}{8}$

Concept Check ✅ – Answer at **cc.dsc10.com**

I have three cards: red, blue, and green. What is the chance that I choose a card at random and it is green, then – **without putting it back** – I choose another card at random and it is red?

- A) $\frac{1}{9}$
- B) $\frac{1}{6}$
- C) $\frac{1}{3}$
- D) $\frac{2}{3}$
- E) None of the above.

R G
R B
G B
G R
B G
B R

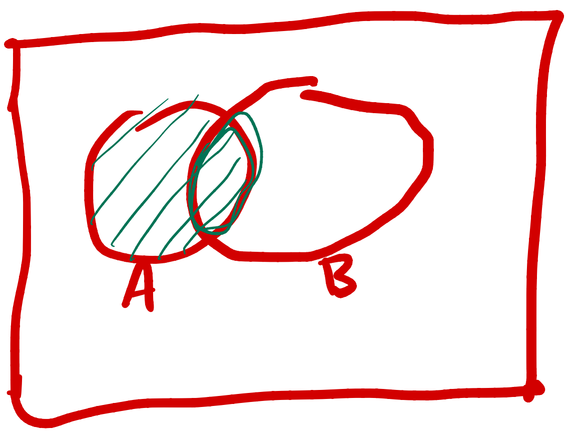
$\frac{1}{6}$

Conditional probabilities

- Two events A and B can both happen. Suppose that we know A has happened, but we don't know if B has.
- If all outcomes are equally likely, then the conditional probability of B given A is:

$$P(B \text{ given } A) = \frac{\#\text{ of outcomes satisfying both } A \text{ and } B}{\#\text{ of outcomes satisfying } A}$$

- Intuitively, this is similar to the definition of the regular probability of B, $P(B) = \frac{\#\text{ of outcomes satisfying } B}{\text{total } \#\text{ of outcomes}}$, if you restrict the set of possible outcomes to be just those in event A.



A = spilling coffee

B = rushing to a meeting

P(rushing to meeting given that I spilled coffee)

Concept Check ✅ – Answer at **cc.dsc10.com**

$$P(B \text{ given } A) = \frac{\# \text{ of outcomes satisfying both } A \text{ and } B}{\# \text{ of outcomes satisfying } A}$$

I roll a six-sided die and don't tell you what the result is, but I tell you that it is 3 or less. What is the probability that the result is even?

- A) $\frac{1}{2}$
- B) $\frac{1}{3}$
- C) $\frac{1}{4}$
- D) None of the above.

1 2

3 4

5 6

1/3

Probability that two events both happen

- Suppose again that A and B are two events, and that all outcomes are equally likely. Then, the probability that both A and B occur is

$$P(A \text{ and } B) = \frac{\# \text{ of outcomes satisfying both } A \text{ and } B}{\text{total } \# \text{ of outcomes}}$$

- **Example 2**: I roll a fair six-sided die. What is the probability that the roll is 3 or less **and** even?

1 2

3 4

5 6

$1/6$

The multiplication rule

- The multiplication rule specifies how to compute the probability of both A and B happening, even if all outcomes are not equally likely.

$$P(A \text{ and } B) = P(A) \cdot P(B \text{ given } A)$$

- **Example 2, again**: I roll a fair six-sided die. What is the probability that the roll is 3 or less and even?

$$P(\text{3 or less and even}) = \underbrace{P(\text{3 or less})}_{1/2} \cdot \underbrace{P(\text{even given 3 or less})}_{1/3}$$

1 2

3 4

5 6

$$= 1/6$$

What if A isn't affected by B? 🤔

- The multiplication rule states that, for any two events A and B,

$$P(A \text{ and } B) = P(A) \cdot P(B \text{ given } A)$$

- What if knowing that A happens doesn't tell you anything about the likelihood of B happening?
    - Suppose we flip a fair coin three times.
    - The probability that the second flip is heads doesn't depend on the result of the first flip.
- Then, what is $P(A \text{ and } B)$?

if A tells you nothing about probability of B,

$$P(B \text{ given } A) = P(B)$$

Independent events

- Two events A and B are independent if $P(B \text{ given } A) = P(B)$, or equivalently if

$$P(A \text{ and } B) = P(A) \cdot P(B)$$

- **Example 3**: Suppose we have a coin that is **biased**, and flips heads with probability 0.7. Each flip is independent of all other flips. We flip it 5 times. What's the probability we see 5 heads in a row?

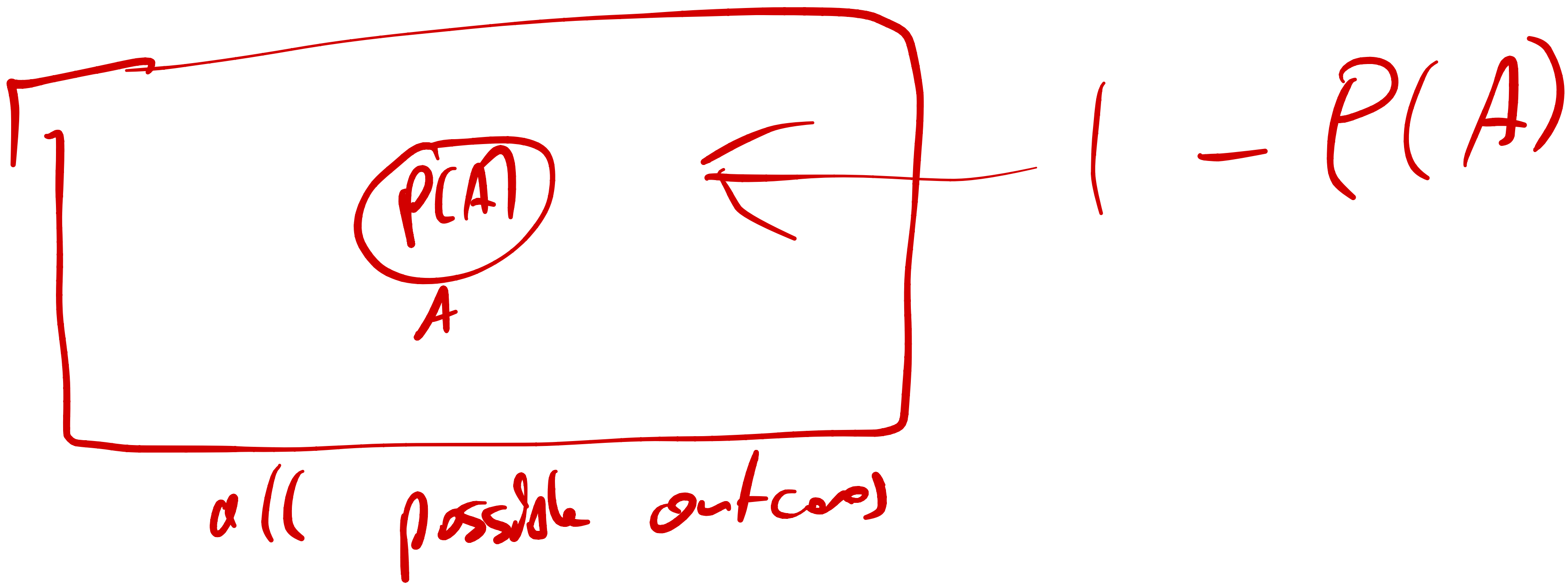
$$P(\text{all 5 heads}) = P(\text{1st heads}, \text{2nd heads}, \ldots \text{5th H})$$

$$= P(\text{1st heads}) \cdot P(\text{2nd heads}) \ldots$$

$$= 0.7^5$$

Probability that an event *doesn't* happen → complement

- The probability that A **doesn't** happen is $1 - P(A)$.
- For example, if the probability it is sunny tomorrow is 0.85, then the probability it is not sunny tomorrow is 0.15.



Concept Check ✅ – Answer at cc.dsc10.com

Every time I call my grandma 👵, the probability that she answers her phone is $\frac{1}{3}$, independently for each call. If I call my grandma three times today, what is the chance that I will talk to her at least once?

- A) $\frac{1}{3}$
- B) $\frac{2}{3}$
- C) $\frac{1}{2}$
- ~~D) 1~~
- E) None of the above.

$$P(\text{at least once}) = 1 - P(\text{never})$$

$$= 1 - P(\text{not 1st and not 2nd and not 3rd})$$

$$= 1 - P(\text{not 1}) \cdot P(\text{not 2}) \cdot P(\text{not 3})$$

$$= 1 - \left(\frac{2}{3}\right)^3 \qquad = \frac{19}{27}$$

Probability of either of two events happening

- Suppose again that A and B are two events, and that all outcomes are equally likely. Then, the probability that either A or B occur is

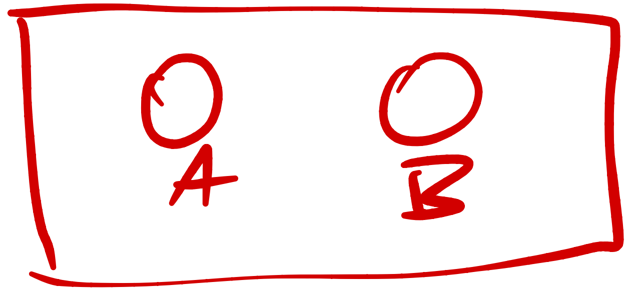
$$P(A \text{ or } B) = \frac{\#\text{ of outcomes satisfying either } A \text{ or } B}{\text{total } \#\text{ of outcomes}}$$

- **Example 4**: I roll a fair six-sided die. What is the probability that the roll is even or at least 5?

1 2
3 4
5 6

$P(\underline{\text{even}} \text{ or } \underline{\text{at least 5}})$

$= 2/3$



The addition rule

- Suppose that if A happens, then B doesn't, and if B happens, then A doesn't.
  - Such events are called **mutually exclusive** – they have **no overlap**.
- If A and B are any two mutually exclusive events, then

$$P(A \text{ or } B) = P(A) + P(B)$$

- **Example 5**: Suppose I have two biased coins, coin A and coin B. Coin A flips heads with probability 0.6, and coin B flips heads with probability 0.3. I flip both coins once. What's the probability I see two different faces?

$$\underbrace{P(A = \text{heads and } B = \text{tails})}_{\text{"event A"}} \text{ OR } \underbrace{P(A = \text{tails and } B = H)}_{\text{"event B"}}$$

Mutually exclusive

$$P(A = H \text{ and } B = T) + P(A = T \text{ and } B = H)$$

$$P(A = H) \cdot P(B = T) + P(A = T) \cdot P(B = H)$$

$$.6 \cdot 0.7 + 0.4 \cdot 0.3 = 0.54$$

Aside: Proof of the addition rule for equally-likely events

You are not required to know how to "prove" anything in this course; you may just find this interesting.
If A and B are events consisting of equally likely outcomes, and furthermore A and B are mutually exclusive (meaning they have no overlap), then

$$
\begin{aligned}
P(A \text{ or } B) &= \frac{\# \text{ of outcomes satisfying either } A \text{ or } B}{\text{total } \# \text{ of outcomes}} \\
&= \frac{(\# \text{ of outcomes satisfying } A) + (\# \text{ of outcomes satisfying } B)}{\text{total } \# \text{ of outcomes}} \\
&= \frac{(\# \text{ of outcomes satisfying } A)}{\text{total } \# \text{ of outcomes}} + \frac{(\# \text{ of outcomes satisfying } B)}{\text{total } \# \text{ of outcomes}} \\
&= P(A) + P(B)
\end{aligned}
$$

Summary, next time

- Probability describes the likelihood of an event occurring.
- There are several rules for computing probabilities. We looked at many special cases that involved equally-likely events.
- There are two general rules to be aware of:
    - The **multiplication rule**, which states that for any two events, $P(A \text{ and } B) = P(B \text{ given } A) \cdot P(A)$.
    - The **addition rule**, which states that for any two **mutually exclusive** events, $P(A \text{ or } B) = P(A) + P(B)$.
- **Next time:** Simulations.