Lecture 12 – Probability

DSC 10, Winter 2022

Announcements

- Lab 4 is due **tomorrow 2/1 at 11:59pm**.
- Homework 4 is due on **Saturday 2/5 at 11:59pm**.
- The Midterm Project is due on **Saturday 2/12 at 11:59pm**.
    - Start early !!
- The Midterm Exam is **next Wednesday 2/9 during lecture**.
    - You will take it remotely on Gradescope.
    - More details to come.
- As of today (1/31), lectures and discussions are offered in-person, though you can join via Zoom as well.
    - In addition, some office are now in-person. See the Calendar for details.

Agenda

- Review problem (video): conditional statements and iteration.
- Motivation for probability.
- Probability theory.

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:

- Note 18 in the course notes.
- Computational and Inferential Thinking, Chapter 9.5.
- Theory Meets Data, Chapters 1 and 2.
- Khan Academy's unit on Probability.
- Suraj's DSC 40A slides on probability from last quarter.
    - More in-depth than what you need to know for this course, but still may be helpful.

Conditional statements and iteration

Walkthrough video

Note: we will not get a chance to cover this problem in class. However, it's excellent practice for Lab 4, Homework 4, and the Midterm Project, so we recommend you work through it on your own. The solution to it can be found in the video below.

In [ ]:

```
YouTubeVideo('6HOAk0GAqKU')
```

Example: number of days in between two dates

Below, complete the implementation of the function `days_between`, which takes in two (month, day) pairs in the same year and returns the number of days in between those two dates. Example behavior is shown below.

```
>>> days_between(3, 12, 3, 31)
19 # Number of days between March 12th and March 31st

>>> days_between(6, 5, 11, 2)
150 # Number of days between June 5th and November 2nd
```

In [ ]:

```python
# Your solution here

days_per_month = [0, 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31]

def days_between(month1, day1, month2, day2):
    ...
```

In [ ]:
```
days_between(3, 12, 3, 31)
```

In [ ]:
```
days_between(6, 5, 11, 2)
```

In [ ]:
```
days_between(1, 31, 12, 25)
```

In [ ]:

```python
# SOLUTION

days_per_month = [0, 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, 31] # We can use this in

def days_between(month1, day1, month2, day2):
    # If the two dates are from the same month, return the difference in the days
    if month1 == month2:
        return day2 - day1
    else:
        # Create a variable to keep track of the total number of days
        total_days = 0

        # Add the number of days remaining in month1
        total_days = total_days + days_per_month[month1] - day1

        # Add the number of days in every full month between month1 and month2
        for full_month in np.arange(month1 + 1, month2):
            total_days = total_days + days_per_month[full_month]

        # Add the number of days in month2
        total_days = total_days + day2

        return total_days
```

Motivation for probability

Swain vs. Alabama, 1965

- Robert Swain was a black man convicted of crime in Talladega County, Alabama.
- He appealed the jury's decision all the way to the Supreme Court, on the grounds that Talladega County systematically excluded Black people from juries.
- At the time, only men 21 years or older were allowed to serve on juries. 26% of this population was Black.
- But of the 100 men on Robert Swain's jury panel, only 8 were Black.

The Supreme Court’s ruling

- About disparities between the percentages in the eligible population and the jury panel, the Supreme Court wrote:

> *"... the overall percentage disparity has been small..."*

- The Supreme Court denied Robert Swain’s appeal and he was sentenced.
- The fact that the jury panel had far fewer Black men proportionally than Talladega County is an example of racial bias.
- Over the next few weeks, we will give you tools to quantitatively highlight this bias.
  - We will try to answer the question, "what are the chances that this disparity was due to random chance?"
  - If this chance is small, we know something is wrong.
  - But first: we need to formalize what **probability** is.

Probability theory

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 50%.
- One interpretation of probability says that if we flipped a coin infinitely many times, then 50% of 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%) which 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?

Example 1 solved

- When we flip a fair coin 3 times, there are 8 possible outcomes: HHH, HHT, HTH, HTT, THH, THT, TTH, and TTT.
- These outcomes are all equally likely.
- 3 of these outcomes have exactly 2 heads: HHT, HTH, and THH.
- So, the probability of seeing exactly 2 heads in 3 flips of a fair coin is $\frac{3}{8}$.

Discussion Question

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.

To answer, go to menti.com and enter the code 1918 2877.

Discussion Question solved

- There are 6 possible outcomes: RG, RB, GR, GB, BR, and BG.
- These outcomes are equally likely.
- There is only 1 outcome which makes the event happen: GR.
- Hence the probability is $\frac{1}{6}$.

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.
- **Example 2:** Suppose I roll a fair six-sided die, and suppose A is the event "roll is 3 or less" and B is the event "roll is even". What is $P(B \text{ given } A)$?

Example 2 solved

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

Another way of phrasing the problem: I roll a fair 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?

- There are three outcomes where the roll is 3 or less: 1, 2, and 3.
- There is only one outcome where the roll is 3 or less and even: 2.
- So the probability that the roll is even given that it is 3 or less is $P(B \text{ given } A) = \frac{1}{3}$.

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

Example 3 solved

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

- Only one outcome is both 3 or less and even: 2.
- There are 6 total outcomes.
- Thus, $P(A \text{ and } B) = \frac{1}{6}$.

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

Example 3 solved, again

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

- The probability that the roll is 3 or less is $P(A) = \frac{1}{2}$.
- From before, the probability that the roll is even given that the roll is 3 or less is $P(B \text{ given } A) = \frac{1}{3}$.
- Thus, the probability the roll is both 3 or less and even is $P(A \text{ and } B) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$.
- Note that an equivalent formula is $P(A \text{ and } B) = P(B) \cdot P(A \text{ given } B)$.

Generally, situations involving an ”and” involve multiplication.

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 $P(A \text{ and } B)$?

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 4:** 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?

Example 4 solved

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?

- The probability of seeing heads on a single flip is 0.7.
- Each flip is independent.
- So, the probability of seeing 5 heads in a row is

$$0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 \cdot 0.7 = 0.7^5$$

Probability that an event *doesn't* happen

- 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.

Discussion Question

Every time I call my grandma 👵, the probability that she answers her phone is $\frac{1}{3}$. 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.

To answer, go to menti.com and enter the code 1918 2877.

Discussion Question solved

- Let's first calculate the probability that she **doesn't** answer her phone in three tries.
    - The probability she doesn't answer her phone on any one attempt is $\frac{2}{3}$.
    - So the probability she doesn't answer her phone in three tries is $\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \frac{8}{27}$.
- But we want the probability of her answering **at least** once. So we subtract the above result from 1.
    - $1 - \frac{8}{27} = \frac{19}{27}$; none of the above!

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

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

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

Example 5 solved

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

- There are three outcomes that are even: 2, 4, 6.
- There are two outcomes that are at least 5: 5, 6.
- There are four total outcomes that satisfy at least one of the two conditions: 2, 4, 5, 6.
- Thus, the probability that the roll is even or at least 5 is $\frac{4}{6} = \frac{2}{3}$.
  - Note that this is not $P(A) + P(B)$, which would be $\frac{3}{6} + \frac{2}{6} = \frac{5}{6}$, because there is **overlap** between events A and B.

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 6:** 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?

Example 6 solved

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. The two coins are independent of one another. I flip both coins once. What's the probability I see two different faces?

- The event we see two different faces corresponds to either seeing a head then a tail, **or** a tail then a head (i.e. not both heads and not both tails).
- The probability of seeing a head then a tail is $0.6 \cdot (1 - 0.3)$, because the two coins are independent of one another.
- The probability of seeing a tail then a head is $(1 - 0.6) \cdot 0.3$.
- So, the probability of seeing two different faces is

$$0.6 \cdot (1 - 0.3) + (1 - 0.6) \cdot 0.3 = 0.54$$

Generally, situations involving an ”or” involve addition.

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

Summary

- 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.