Lecture 5-in-class edits, morning

September 9, 2020

1 Lecture 5 - Introduction to Conditional Probability

```
[1]: import matplotlib.pyplot as plt
     %matplotlib inline
     plt.style.use('seaborn-colorblind')

     import random
     import numpy as np
     import numpy.random as npr
```

Last class we:

1. Worked through some analytical and simulation exercises to describe the sample space, events and probability of outcomes from discrete sequential models.
   - When characterizing an experiment, the best practice is to:
     - Define the sample space
     - Define the event class
     - Define probabilities of outcomes

2. Learned how to represent discrete sequential models in tree-based representations.
   - This representation can be useful in determining probabilities of outcomes

1.1 Introduction to Conditional Probability

Consider the following scenarios:

Example 1: A magician has in her pocket a fair coin and a two-headed coin. She chooses one at random and flips it. What is the probability that the result is heads?

- Is this experiment fair? No, because the probability of each outcome (H, T) are not equal. P(H)=3/4 and P(T)=1/4

Let's compute this probability on the virtual whiteboard.

Let's build a simulation to answer this:

```
[5]: def one_flip(num_sims = 100000):
         coins = ['fair','2head']
         head_count = 0
         for sim in range(num_sims):
             coin = random.choice(coins)
```

1

```
        if coin == 'fair':
            S = ['H', 'T']
        else:
            S = ['H', 'H']
        value = random.choice(S)
        if value == 'H':
            head_count+=1
    print('Probability of heads is ',head_count/num_sims)
```

[6]: `one_flip()`

```
Probability of heads is  0.751
```

Example 2: Suppose that she chooses a coin at random. Using that coin, she flips it once and observes heads. What is the probability of observing heads in the second flip (using the same coin) if we observed heads in the first flip?

- How can we visualize and compute the analytical probability of this event?

Let's compute this probability on the virtual whiteboard.

Let's build a simulation to answer this:

[9]:
```python
def double_flip(num_sims=100000):
    coins = ['fair', '2head']
    heads1_count = 0
    heads2_count = 0

    for sim in range(num_sims):
        coin = random.choice(coins)
        if coin == 'fair':
            S = ['H', 'T']
        else:
            S = ['H', 'H']
        values = random.choices(S, k=2)
        if values[0] == 'H':
            heads1_count +=1
            if values[1]=='H':
                heads2_count +=1

    print('Probability of observing heads in the 2nd ',
        'flip if heads in the 1st flip is ',heads2_count/heads1_count)
```

[10]: `double_flip()`

```
Probability of observing heads in the 2nd  flip if heads in the 1st flip is
0.8324728928054785
```

This probability is called **conditional probability** as it provides us a way to reason about the outcome of an experiment, based on **partial information**.
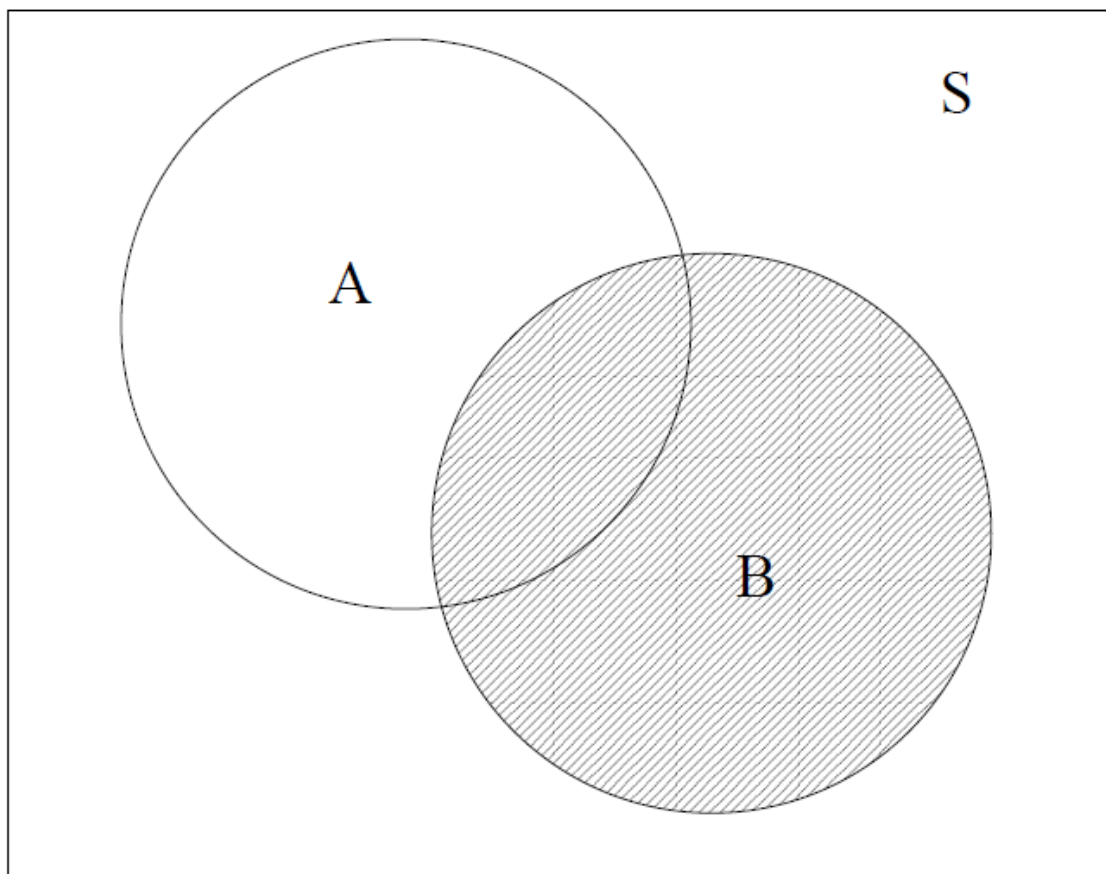
For scenario 2, consider the event H_i =heads on flip i. We are asking what is the **probability of** H_2 **given** H_1 **occurred**, that is,

$$P(H_2|H_1) = \frac{5}{6}$$

Consider the Venn diagram:

```
[11]: from IPython.display import Image
      Image("figures/condProb1.png")
```
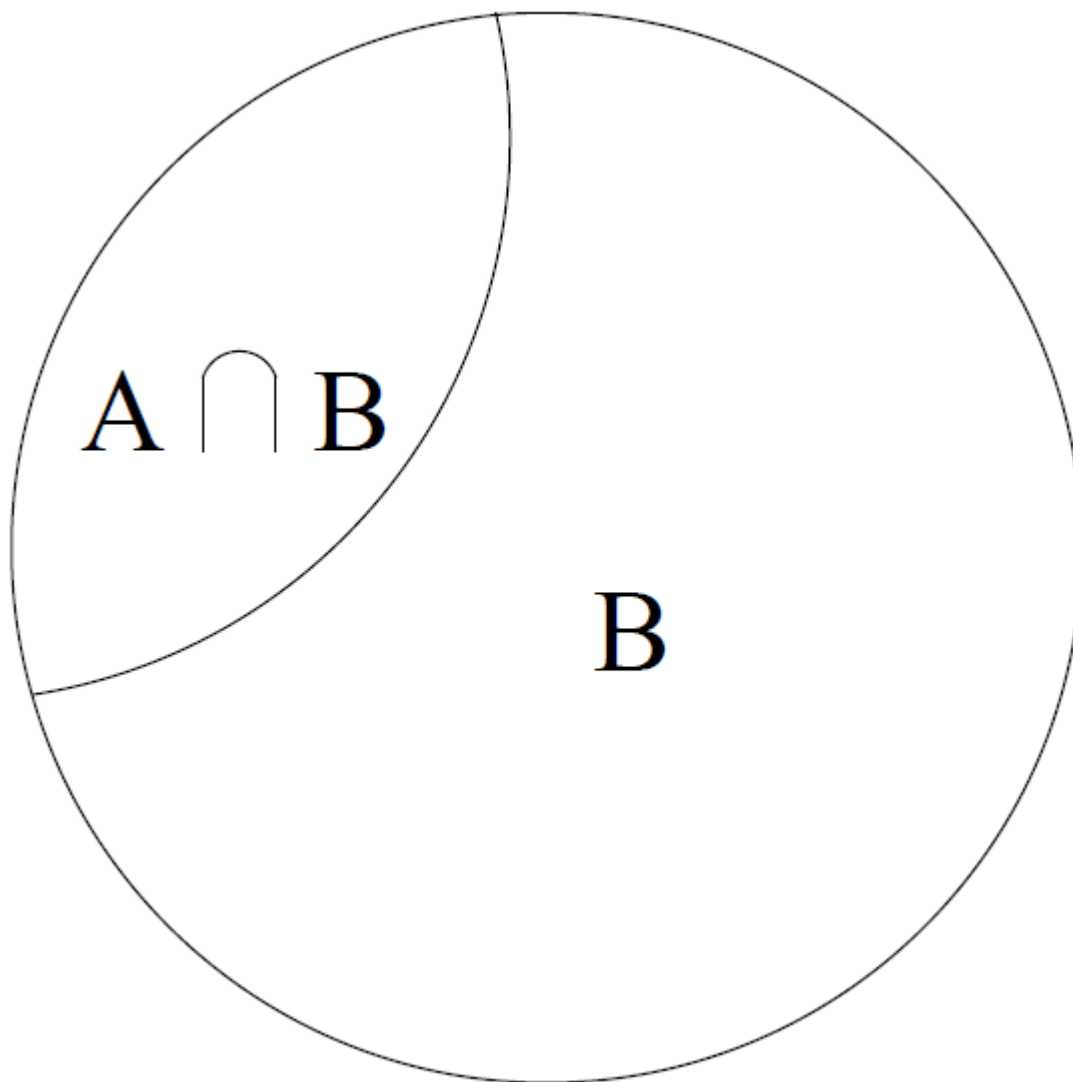
[11]:

If we **condition** on B having occurred, then we can form the new Venn diagram:

```
[12]: from IPython.display import Image
      Image("figures/condProb2.png")
```

[12]:

3

This diagram suggests that if $A \cap B = \emptyset$ then if B occurs, A could not have occurred.

Similarly if $B \subset A$, then if B occurs, the diagram suggests that A must have occurred.

A definition of conditional probability that agrees with these and other observations is:

**Conditional Probability

For $A \in \mathcal{F}$, $B \in \mathcal{F}$, the **conditional probability** of event A *given* event B occurred is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ for } P(B) > 0$$

Claim: If $P(B) > 0$, the conditional probability $P(\bullet|B)$ satisfies the axioms on the original sample space $(\Omega, \mathcal{F}, P(\bullet|B))$.

For a fixed event $B \neq \emptyset$, the conditional probabilities $P(A|B)$ form a legitimate probability law that satisfies the three axioms!

Example 3: A computer lab contains

- two computers from manufacturer A, one of which is defective
- three computers from manufacturer B, two of which are defective

A user sits down at a computer at random. Let the properties of the computer she sits down at be denoted by a two letter code, where the first letter is the manufacturer and the second letter is D for a defective computer or N for a non-defective computer. (We add a subscript to differentiate computers with the same two-letter code.)

- What is the sample space?

$$\Omega = \{AD, AN, BD_1, BD_2, BN\}$$

Let * E_A be the event that the selected computer is from manufacturer A * E_B be the event that the selected computer is from manufacturer B * E_D be the event that the selected computer is defective

Let's find

$$P(E_A) = \frac{2}{5}$$

$$P(E_B) = \frac{3}{5}$$

$$P(E_D) = \frac{3}{5}$$

Now, suppose that I select a computer and tell you its manufacturer. Does that influence the probability that the computer is defective?

- For example, I tell you the computer is from manufacturer A. Then what is the probability that it is defective?

$$P(E_D|E_A) = \frac{P(E_D \cap E_A)}{P(E_A)} = \frac{1}{2}$$

- Let's find:

$$P(E_D|E_B) = \frac{2}{3}$$

$$P(E_A|E_D) = \frac{1}{3}$$

$$P(E_B|E_D) = \frac{2}{3}$$

Relating Conditional and Unconditional Probabilities

Which of the following statement is true?

1. $P(A|B) \geq P(A)$
2. $P(A|B) \leq P(A)$
3. Not necessarily 1 or 2

1.2 Conditional Probability for Discrete Sample Spaces with Equal Probabilities

Consider again our simulation of the magician's coins.

We directly estimated of those outcomes where the first flip was heads what proportion was the second flip heads. I.e., we did not use the definition of conditional probability, which involves a ratio of probabilities.

How does that work out?

Let H_i be the event that the outcome of the ith flip was heads. We were trying to estimate $P(H_2|H_1)$.

If we were to use the definition of conditional probability, then we would find this as

$$P(H_2|H_1) = \frac{P(H_2 \cap H_1)}{P(H_1)}$$

If we didn't know how to solve these analytically, we could estimate them by their relative frequencies. Let:

- N be the number of simulations,
- N_1 be the number of simulations in which the first flip is heads, and
- N_{12} be the number of simulations in which both flips are heads.

Then

$$P(H_2|H_1) = \frac{N_{12}/N}{N_1/N} \tag{1}$$

$$= \frac{N_{12}}{N_1} \tag{2}$$

Example 4: XOR of two independent binary values.

Flip a fair coin with sides labeled `0' and `1' two times. * Let E_i denote a `1' on the top face on flip $i$. * Let $F_i$ denote a `0' on the top face on flip i. * Let G denote the event that the XOR of the values observed on the top faces on the two flips is `1'.

Compute:

$$P(E_1) = \frac{1}{2}, P(E_2) = \frac{1}{2}, P(F) = \frac{1}{2}$$

$$P(E_1|E_2) = \frac{1}{2}$$

$$P(E_2|E_1) = \frac{1}{2}$$

$$P(G|E_1) = \frac{1}{2}$$

and

$$P(G|E_1 \cap E_2) = 0$$

1.3   Using Conditional Probability to Decompose Events - Chain Rules

Let's use the `virtual whiteboard` to depict how we can decompose conditional probability using a tree-based sequential representation.

In general, note that:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$\Rightarrow P(A \cap B) = P(A|B)P(B)$$

and

$$P(B|A) = \frac{P(A \cap B)}{P(B)}$$

$$\Rightarrow P(A \cap B) = P(B|A)P(A)$$

Equations (19) and (21) **are chain rules** for expanding the probability of the intersection of two events.

- The chain rule can be easily generalized to more than two events:

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B) \tag{3}$$

$$= P(A) \cdot \frac{P(A \cap B)}{P(A)} \cdot \frac{P(A \cap B \cap C)}{P(A \cap B)} \tag{4}$$

Similarly,

$$P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C)$$

7

2 Short Assignment 1 - due before Friday's class (@10:39 AM for everyone)

Suppose that you have 2 coins, one fair and one 2-headed coin. Consider the experiment where you select a coin at random and flip it once. What is the probability that the coin is the 2-headed coin *given that* the observed result is *heads*?

1. Write down, in equation form, the conditional probability that this problem is asking.
2. Use the tree-based description, the definition of conditional probability and the chain rule to find out the analytical probability.
3. Build a simulation to answer this question, and compare it with the analytical solution.

2.1 Tomorrow is Recitation day

You will solve problems in groups, and I will be answering questions and helping solving these problems.

[ ]: