Lecture 3: Conditional Probability
Recap: Probability

- Sample space $\Omega$: all possible outcomes of an experiment
- Probability of an event $A$ (a set of possible outcomes) denoted as $P(A) \in [0, 1]$, $A \subset \Omega$
- Axioms of probability
  - Nonnegativity: $P(A) \geq 0$
  - Additivity: For any disjoint events $A_1, A_2, A_3, \ldots$,
    $P(\bigcup_i A_i) = \sum_i P(A_i)$
  - Normalization: $P(\Omega) = 1$
- Conditional probabilities $P(A|B)$: probability of event $A$ given event $B$ happens
- New probability space:

$$P(A|B) = \frac{\text{number of elements of } A \cap B}{\text{number of elements of } B} = \frac{P(A \cap B)}{P(B)}.$$
“Real-world” example of conditional probability
New probability space $P(\cdot | B)$

Verify that the axioms of probability are satisfied!

- **Nonnegativity**: $P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0$ since $P(A \cap B) \geq 0$

- **Additivity**: For any two disjoint sets $A$ and $C$, show that $P(A \cup C|B) = P(A|B) + P(C|B)$.

\[
P(A \cup C|B) = \frac{P((A \cup C) \cap B)}{P(B)}
\]

\[
= \frac{P((A \cap B) \cup (C \cap B)))}{P(B)} = \frac{P(A \cap B) + P(C \cap B)}{P(B)}
\]

\[
= P(A|B) + P(C|B).
\]

- **Normalization**: New sample space is $B$.

\[
P(B|B) = \frac{P(B \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1
\]
Example

Let us have two unfair coin tosses where the joint probability has

\[ P(\{HH\}) = \frac{1}{2} \]
\[ P(\{HT\}) = \frac{1}{4} \]
\[ P(\{TH\}) = \frac{1}{8} \]
\[ P(\{TT\}) = \frac{1}{8} \]

What is the probability that we have exactly one \( H \) given that the second toss shows \( H \)?

Define \( A \) and \( B \) first.

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{TH\})}{P(\{HH, TH\})} = \frac{\frac{1}{8}}{\frac{1}{2} + \frac{1}{8}} = \frac{1}{5}.
\]
Challenging Exercise

Example: Throw of two dice. Each of the 36 outcomes are equally likely

- $A = \text{max of two dice is less than 5}$
- $B = \text{min of the two dice is greater than 1}$

What is $P(A|B)$? (Please try this at home)

- $P(A) = \frac{16}{36}$
- $P(B) = \frac{25}{36}$
- $P(A \cap B) = \frac{9}{36}$
- $P(A|B) = \frac{9}{25}$
Sequential Model for Conditional Probabilities

Many experiments have a sequential characteristic: the future outcomes depending on the past.

For example, consider an example involving three coin tosses.

- The first toss is unbiased (fair): \( P(H) = 0.5 \) and \( P(T) = 0.5 \).

- Based on the outcome of the first toss, the second toss is biased towards that outcome by 60%.
  - For example, if the outcome of the first toss is \( H \), then the second toss has \( P(H) = 0.6 \) and \( P(T) = 0.4 \).

- Based on the outcome of the second toss, the third toss is biased towards that outcome by 70%.

Let us draw a tree-based sequential description.
Sequential Model for Conditional Probabilities

First Toss  |  Second Toss  |  Third Toss
---|---|---
H  |  H  |  HHH  
P(H₁) = 0.5  
P(H₂ | H₁) = 0.6  
P(H₃ | H₂H₁) = 0.7
T  |  H  |  HHT  
P(T₁) = 0.5  
P(T₂ | T₁) = 0.6  
P(H₃ | T₂T₁) = 0.7
H  |  T  |  HTH  
P(T₁) = 0.4  
P(T₂ | T₁) = 0.3  
P(H₃ | T₂T₁) = 0.3
T  |  H  |  HTH  
P(T₁) = 0.5  
P(T₂ | T₁) = 0.4  
P(T₃ | H₂T₁) = 0.3
H  |  T  |  TTH  
P(T₁) = 0.4  
P(T₂ | T₁) = 0.6  
P(T₃ | H₂T₁) = 0.7
T  |  T  |  TTH  
P(T₁) = 0.5  
P(T₂ | T₁) = 0.4  
P(T₃ | H₂T₁) = 0.7
H  |  T  |  TTT  
P(T₁) = 0.4  
P(T₂ | T₁) = 0.6  
P(T₃ | H₂T₁) = 0.3
Sequential Model for Conditional Probabilities

How to setup a tree-based sequential description and use it?

1. Leaves represent events of interest, which occur in a sequential manner
2. Branches represent the conditional probability
3. The probability of the end-leaf can be computed by multiplying conditional probabilities from the root.
Sequential Model for Conditional Probabilities

<table>
<thead>
<tr>
<th>First Toss</th>
<th>Second Toss</th>
<th>Third Toss</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- \( P(H_1) = 0.5 \)
- \( P(T_1) = 0.5 \)
- \( P(H_2 | H_1) = 0.6 \)
- \( P(H_2 | T_1) = 0.4 \)
- \( P(P_3 | H_2 H_1) = 0.7 \)
- \( P(P_3 | T_2 H_1) = 0.3 \)
- \( P(H_3 | H_2 T_1) = 0.7 \)
- \( P(H_3 | T_2 T_1) = 0.3 \)
- \( P(T_3 | H_2 H_1) = 0.3 \)
- \( P(T_3 | T_2 T_1) = 0.7 \)
- \( P(T_1 T_2 H_3) = 0.09 \)
- \( P(T_1 T_2 T_3) = 0.21 \)
- \( P(H_1 H_2 H_3) = 0.21 \)
- \( P(H_1 H_2 T_3) = 0.09 \)
- \( P(H_1 T_2 H_3) = 0.06 \)
- \( P(H_1 T_2 T_3) = 0.14 \)
- \( P(T_1 H_2 H_3) = 0.14 \)
- \( P(T_1 H_2 T_3) = 0.06 \)
- \( P(T_1 T_2 H_3) = 0.09 \)
- \( P(T_1 T_2 T_3) = 0.21 \)

Root
Multiplication Rule

- We learned conditional probability
  \[ P(A|B) = \frac{P(A \cap B)}{P(B)} , \]
  which can be re-written as
  \[ P(A \cap B) = P(B)P(A|B) = P(A)(B|A) \]

- Now, what about
  \[ P(A \cap B \cap C) = P((A \cap B) \cap C) \]
  \[ = P(D \cap C), \text{ where } D = (A \cap B) \]
  \[ = P(D)P(C|D) \]
  \[ = P(A \cap B)P(C|A \cap B) \]
  \[ = P(A)P(B|A)P(C|A \cap B) \]

These are other equivalent results for \( P(A \cap B \cap C) \).
Multiplication Rule

In general,

\[ P(\cap_{i=1}^{n} A_i) \equiv P(A_1 \cap A_2 \cap \ldots A_n) \]

\[ = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots P(A_n|\cap_{i=1}^{n-1} A_i) \]
Sequential Model

Probability of a leaf of this tree?

Root

H

P(H₁)=0.5

P(H₂|H₁)=0.6

P(H₃|H₂H₁)=0.7

P(H₁H₂H₃)=0.21

T

P(T₁)=0.5

P(H₂|T₁)=0.4

P(H₃|H₂T₁)=0.7

P(T₁H₂H₃)=0.14

P(T₂|H₁)=0.4

P(T₃|H₂H₁)=0.3

P(H₁T₂H₃)=0.06

P(T₂|T₁)=0.6

P(T₃|H₂T₁)=0.3

P(T₁T₂H₃)=0.09

P(H₂|T₁)=0.4

P(T₃|T₂T₁)=0.7

P(T₁T₂T₃)=0.21

HH

P(H₁H₂T₃)=0.09

TH

P(H₁T₂H₃)=0.06

HT

P(H₁T₂T₃)=0.14

P(H₁)=0.5

P(H₂|H₁)=0.6

P(H₃|H₂H₁)=0.7

P(H₁H₂H₃)=0.21

P(H₁)=0.5

P(T₁)=0.5

P(H₂|T₁)=0.4

P(H₃|H₂T₁)=0.7

P(T₁H₂H₃)=0.14

P(T₂|H₁)=0.4

P(T₃|H₂H₁)=0.3

P(H₁T₂H₃)=0.06

P(T₂|T₁)=0.6

P(T₃|H₂T₁)=0.3

P(T₁T₂H₃)=0.09

P(T₃|T₂T₁)=0.7

P(T₁T₂T₃)=0.21
More exercise

Three cards are drawn from an ordinary 52-card deck without replacement (drawn cards are not placed back in the deck). We wish to find the probability that none of the three cards is a heart. We assume that at each step, each one of the remaining cards is equally likely to be picked. (see, Textbook)
More exercise

- $A_i =$ first draw is not a heart, $i = 1, 2, 3$
- $P(A_1) = \frac{39}{52}$
- $P(A_2|A_1) = \frac{38}{51}$. 
- $P(A_3|A_2 \cap A_1) = \frac{37}{50}$
- $P(A_1 \cap A_2 \cap A_3) = \frac{39 \cdot 38 \cdot 37}{52 \cdot 51 \cdot 50}$
The sequential model