Lecture 2

- **Read:** Chapter 1.5-1.7

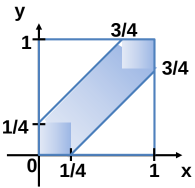
- Conditional Probability
- Sequential Calculation of Probability
- Total Probability and Bayes' Theorem
- Independence and Disjointness
- Pairwise Independence and Mutual Independence

Continuous Models

- Sample space does need not be finite; probability of single outcomes is not enough to describe what is going on, so we use continuous models
- **Example:** Romeo and Juliet have a date at 12 pm, but each will arrive at the meeting place with a delay between 0 and 1 hour, with all pairs of delays "equally likely." The first to arrive will wait for 15 mins and leave if the other does not arrive. What is the probability that they will meet?

Continuous Models

- **Model and solution:** $S = [0,1] \times [0,1]$. Equally likely means the probability of a subset of S is equal to its area. This probability law satisfies our axioms. The event that they meet is $M = \{(x,y) : |x-y| \leq 1/4, (x,y) \in S\}$.



- The area of this event is $7/16$, i.e., 1 minus the area of the two triangles, i.e., square with area $(3/4)^2$.

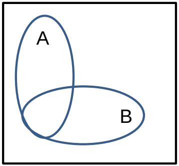
Conditional Probability

Reasoning about events when we have partial information: Key to dealing with many real world problems, e.g., signal detection, stock market, etc.

1. A spot shows up on the radar screen. How likely is it that it corresponds to an aircraft?
2. A fair die is rolled and you are told that the outcome was even. How likely is it that the outcome was a 6?

Example (2) above: fair die $\rightarrow$ all outcomes equally likely; we know the outcome is even, and expect each even outcome to be equally likely, i.e., $P[\text{outcome } 6 \mid \text{outcome even}] = 1/3$

Conditioning



- Consider two events A and B
- $P[A]$ = our knowledge of the likelihood of A
- $P[A]$ = "a priori" probability
- Suppose we cannot completely observe an experiment
  - We learn that event B occurred
  - We do not learn the precise outcome
- Learning B occurred changes $P[A]$
- $P[A|B]$ = probability of A given B occurred (B is our new universe)

Conditional Probability Definition

- Assuming $P[B] \neq 0$, the **conditional probability** A given the occurrence of B is

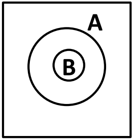
$$P[A|B] = \frac{P[AB]}{P[B]} = \frac{P[A \cap B]}{P[B]}$$

- **Note:** $A \cap B = A \cdot B = AB$
- It follows that $P[A \cap B] = P[B] \cdot P[A|B] = P[A] \cdot P[B|A]$

Conditional Probability: Subsets

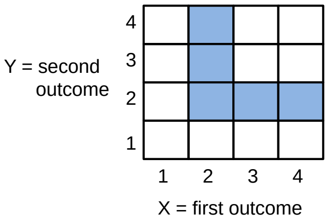
- Suppose that $B \subset A$, then

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B]}{P[B]} = 1$$



Die Roll Example

Consider rolling a four-sided die twice; denote the outcomes X and Y.

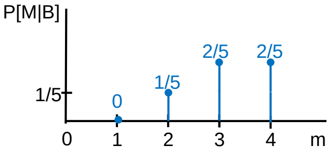


- Let B be the event: $min(X, Y) = 2$
- Let $M = max(X, Y)$
- Compute $\mathrm{P}[\mathrm{M} = \mathrm{m}|\mathrm{B}]$ for $m = 1, 2, 3, 4$

Die Roll Example (cont.)

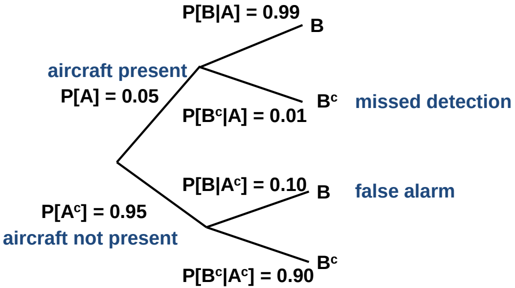
- $P[B] = 5/16$

$$P[M = 1|B] = \frac{P[A \cap B]}{P[B]} = 0$$
$$P[M = 2|B] = \frac{P[A \cap B]}{P[B]} = \frac{1/16}{5/16} = \frac{1}{5}$$
$$P[M = 3|B] = \frac{P[A \cap B]}{P[B]} = \frac{2/16}{5/16} = \frac{2}{5}$$
$$P[M = 4|B] = \frac{P[A \cap B]}{P[B]} = \frac{2/16}{5/16} = \frac{2}{5}$$



Models Based on Conditional Probabilities: Estimation Problems

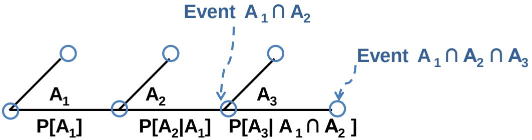
Radar Example



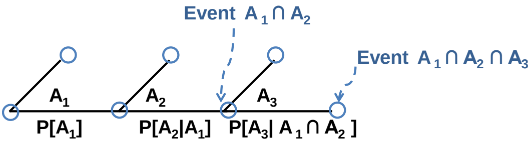
- Event $A =$ airplane is flying above
- Event $B =$ something registers on radar screen
- Suppose the information shown in the tree model shown above
- What are A^c and B^c?

Models Based on Conditional Probabilities: Tree-Based Sequential Description

1. Set up tree with event of interest as a leaf
2. Record the conditional probabilities associated with branches
3. Multiply the probability to get to the leaf



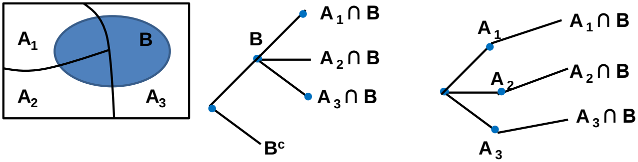
Models Based on Conditional Probabilities: Multiplication Rule



Assuming all the conditioning events have positive probability, we have that

$$P[\cap_{i=1}^{n} A_i] = P[A_1]P[A_2|A_1]P[A_3|A_1 \cap A_2]...P[A_n| \cap_{i=1}^{n-1} A_i]$$

Visualization of the Total Probability Theorem



- The events $A_1, \ldots, A_n$ form a partition of the sample space, so the event B can be decomposed into the disjoint union of its intersections $A_i \cap B$ with the sets A_i.
- In the equivalent sequential model shown on the top right, the probability of leaf $A_i \cap B$ is the product $P[A_i]P[B|A_i]$ of the probabilities along the path leading to that leaf. The event B consists of the three leaves shown in the figure, and $P[B]$ is obtained by adding their probabilities.

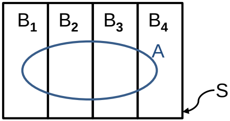
Intuition Behind the Total Probability Theorem

- Intuitively, we are partitioning the sample space into a number of scenarios (events) A_i.
- Then, the probability that B occurs is a weighted average of its conditional probability under each scenario, where each scenario is weighted according to its (unconditional) probability.
- One of the uses of the theorem is to compute the probability of various events B for which the conditional probabilities $P[B|A_i]$ are known or easy to derive.
- The key is to choose appropriately the partition $A_1, ..., A_n$, and this choice is often suggested by the problem structure.

Law of Total Probability (divide and conquer approach to computing the probability of an event)

- If $B_1, B_2, ..., B_m$ is an event space (in other words, given a partition $B_1, B_2, ..., B_m$) and $P[B_i] > 0$ for $i = 1, ..., m$, then

$$P[A] = \sum_{i=1}^{m} P[A|B_i]P[B_i]$$



Bayes' Theorem

- In many situations, we have advance information about $P[A|B]$ and need to calculate $P[B|A]$
- To do so, we have **Bayes' Theorem**:

$$P[B|A] = \frac{P[A|B]P[B]}{P[A]}$$

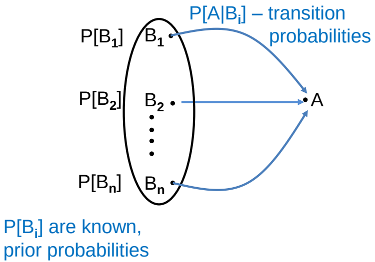
- **Proof:**

$$\frac{P[A|B]P[B]}{P[A]} = \frac{\frac{P[A\cap B]}{P[B]} \times P[B]}{P[A]} = \frac{P[A \cap B]}{P[A]}$$

Bayes' Theorem: Rules for Combining Evidence

- Suppose we are given
  1. "Prior" probabilities: $P[B_i]$
  2. "Transition" probabilities: $P[A|B_i]$ for each i
- Then, for an event space (or partition) $B_1, B_2, ..., B_n$,

$$P[B_i|A] = \frac{P[A|B_i]P[B_i]}{P[A]} = \frac{P[A|B_i]P[B_i]}{\sum_{i=1}^{n} P[A|B_i]P[B_i]}$$



Radar Example Revisited

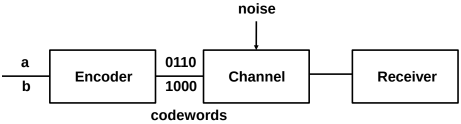
- Event $A=$ airplane is flying above
- Event $B=$ something registers on radar screen

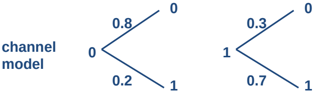
$$\begin{aligned} P[\text{airplane}|\text{register}] &= P[A|B] \\ &= \frac{P[A]P[B|A]}{P[B]} \\ &= 0.34 \end{aligned}$$

$$\begin{aligned} P[B] = P[B|A]P[A] + P[B|A^c]P[A^c] &= 0.99 \times 0.05 + 0.10 \times 0.95 \\ &= 0.1445 \end{aligned}$$

Why so small when P[B|A]= 0.99? (probability of correct detection) *Probability of A^c and false alarm are quite high!*

Example: Decoding of Noise Corrupted Messages





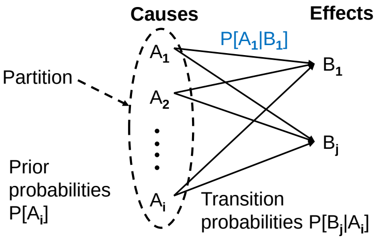
- Prior probabilities: P[a]$=1/3$, P[b]$=2/3$
- Received 0001. What was transmitted?

$$P[a|0001]=\frac{P[a]P[0001|a]}{P[a]P[0001|a]+P[b]P[0001|b]}$$

$$P[b|0001]=\frac{P[b]P[0001|b]}{P[a]P[0001|a]+P[b]P[0001|b]}$$

- Take the most likely event

Summary: One Picture, Two Formulas

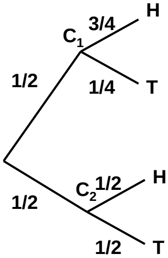


- Total probability theorem
- Bayes' theorem
  - "Reverse the direction" of conditioning (contrast with logical statements)
  - Often used to make inferences "most likely cause for a given effect", commonly applied in drugs testing, detection, ...

Sequential Experiments - Example

- Two coins, one biased, one fair, but you do not know which is which.
- Coin 1: P[H]$=3/4$. Coin 2: P[H]$=1/2$
- Pick a coin at random and flip it. Let C_i denote the event that coin i is picked. What is $P[C_1|H]$?

Solution: Tree Diagram



$$P[C_1|H] = \frac{P[C_1 H]}{P[C_1 H] + P[C_2 H]} = \frac{3/8}{3/8 + 1/4} = 3/5$$

Randomness and Probability

- Recall: We call a phenomenon random if individual outcomes are uncertain but there is a regular distribution of outcomes in a large number of repetitions
- The probability of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions

Probabilities in a Finite Sample Space

- If the sample space is finite, each distinct event is assigned a probability
- The probability of an event is the sum of the probabilities of the distinct outcomes making up the event
- If a random phenomenon has k equally likely outcomes, each individual outcome has probability $1/k$
- For any event A,

$$P[A] = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S}$$

Example

- Roll a fair die and look at the face value
- Sample space: $S = \{1, 2, 3, 4, 5, 6\}$
- This is a finite sample space, and each outcome is equally likely.
- That is,

  $P[X = j] = 1/6, \; \forall j \in S$

  where X is the face value of the die after rolling.
- $P[X \geq 5] = P[X = 5] + P[X = 6] = 1/6 + 1/6 = 1/3$
- $P[X \leq 2] = ?$

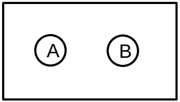
Independent Events

- Events A and B are *independent* if and only if

$$P[AB] = P[A]P[B]$$

- Always check this condition if you are asked about independence!

- **Example:** Suppose $A \cap B = \emptyset$
Is $P[A \cap B] = P[A] \cdot P[B]$? No! Then, not independent!



Independence and Disjointness

- Two events A and B are independent if knowing that one occurs does not change the probability that the other occurs. If A and B are independent, then $P[A \cap B] = P[A]P[B]$

- Events A and B are disjoint if they have no outcomes in common ($A \cap B = \emptyset$). If A and B are disjoint, then $P[A \cup B] = P[A] + P[B]$

- If A and B are disjoint, then the fact A occurs tells us that B cannot occur. So, disjoint events are not independent

- Independence cannot be shown in a Venn diagram because it involves the probabilities of the events rather than just the outcomes that make up the events

- If A and B are independent, so are A and B^c (Clearly, if B does not provide any information about A occurring, then neither does B^c)

Definition: Two Independent Events

Definition 1.6: Events A and B are independent if and only if

$$P[AB] = P[A]P[B]$$

. . . .

Equivalent definitions:

$$P[A|B] = P[A] \qquad\qquad P[B|A] = P[B]$$

Definition: Three Independent Events

Definition 1.7: A_1, A_2, and A_3 are independent if and only if

1. A_1 and A_2 are independent
2. A_2 and A_3 are independent
3. A_1 and A_3 are independent
4. $P[A_1 \cap A_2 \cap A_3] = P[A_1]P[A_2]P[A_3]$

Pairwise Independent Events

Definition: Two events A and B are independent if

$$P[A \cap B] = P[A] \cdot P[B]$$

- Recall: when $P[A], P[B] \neq 0$,
  $P[A \cap B] = P[A|B] \cdot P[B] = P[B|A] \cdot P[A]$
- So, equivalently, independence means:

  $$P[A|B] = P[A] \text{ and } P[B|A] = P[B]$$

  $\implies$ conditioning on an independent event does not change probability of the other

Remark:

- $A \cap B = \emptyset$ does **not** mean A and B are independent, e.g., consider A and A^c
- Independence is a property of events **and** of P

Pairwise Independent Events: Example

Example: Consider an experiment involving two rolls of a four-sided die, each outcome is equally likely with probability $1/16$

1. $A_i = \{\text{1st roll is i}\}$, $B_j = \{\text{2nd roll is j}\}$ independent? (Looks right. Yes!)
2. $A = \{\text{1st roll is 1}\}$, $B = \{\text{sum is 5}\}$ independent? (Does not look obvious, but calculate. Yes!)
3. $A = \{\text{max. of two rolls is 2}\}$, $B = \{\text{min. of two rolls is 2}\}$ independent? (Looks unlikely. No!)

Independence of a Collection of Events

- A collection $A_1, A_2, ..., A_n$ of events are mutually independent if for any finite subset S of $\{1, 2, ..., n\}$,

$$P[\cap_{i \in S} A_i] = \prod_{i \in S} P[A_i]$$

- Write out these conditions for three events: Three events are mutually independent if <u>all</u> these conditions hold:

$$\begin{aligned} P[A_1 \cap A_2 \cap A_3] &= P[A_1]P[A_2]P[A_3] \\ P[A_1 \cap A_2] &= P[A_1]P[A_2] \\ P[A_2 \cap A_3] &= P[A_2]P[A_3] \\ P[A_1 \cap A_3] &= P[A_1]P[A_3] \end{aligned}$$

- Warning: Pairwise independence $\not\Longrightarrow$ mutual independence

Independence of a Collection of Events: Example

Example: Two independent fair coin tosses

HH	**HT**
TH	**TT**

- $A = \{HT, HH\}$: First toss is H
- $B = \{TH, HH\}$: Second toss is H
- $C = \{TT, HH\}$: First toss $=$ second toss
- Verify they are pairwise independent, e.g., $P[C|A] = P[C]$
- Verify they are not mutually independent, $P[A \cap B \cap C] \neq P[A] \cdot P[B] \cdot P[C]$

Pairwise Independence and Mutual Independence

- Are there sets of random events which are pairwise independent but not mutually independent?

  Suppose a box contains 4 tickets labeled by:

  $\boxed{112}$ $\boxed{121}$ $\boxed{211}$ $\boxed{222}$

  Let us choose one ticket at random and consider random events

  $A_1 = \{1 \text{ occurs at the first place}\}$
  $A_2 = \{1 \text{ occurs at the second place}\}$
  $A_3 = \{1 \text{ occurs at the third place}\}$
  $P[A_1] = 1/2, P[A_2] = 1/2, P[A_3] = 1/2$
  $A_1A_2 = 112, A_1A_3 = 121, A_2A_3 = 211$
  $P[A_1A_2] = P[A_1A_3] = P[A_2A_3] = 1/4.$

  So, we conclude that the three events A_1, A_2, A_3 are pairwise independent. However, $A_1A_2A_3 = \emptyset$

  $P[A_1A_2A_3] = 0 \neq P[A_1]P[A_2]P[A_3] = (1/2)^3$

  $\therefore$ Pairwise independence of a given set of random events does <u>not</u> imply that these events are mutually independent.

Pairwise Independence and Mutual Independence (cont.)

- So, we have seen that pairwise independence $\not\Longrightarrow$ mutual independence
- What about $\Longleftarrow$ (i.e., does mutual independence imply pairwise independence?)
    - Yes! The definition of mutual independence implies it!

Pairwise Independence and Mutual Independence (cont.)

- Suppose A, B, C are random events satisfying just the relation $P[ABC] = P[A]P[B]P[C]$. Does it follow that A, B, C are pairwise independent?

  Toss two different standard dice, white and black.

  The sample space S of the outcomes consists of all ordered pairs $(i, j), i, j = 1, ..., 6$, $S = \{(1, 1), (1, 2), ..., (6, 6)\}$. Each point in S has probability $1/36$.

  $A_1 = \{\text{first die} = 1\text{,}2\text{, or } 3\}$, $A_2 = \{\text{first die} = 3\text{,}4\text{, or } 5\}$,
  $A_3 = \{\text{sum of faces is } 9\}$

  $A_1A_2 = \{31, 32, 33, 34, 35, 36\}$, $A_1A_3 = \{36\}$,
  $A_2A_3 = \{36, 45, 54\}$, $A_1A_2A_3 = \{36\}$

  $P[A_1] = 1/2, P[A_2] = 1/2, P[A_3] = 1/9$

  $P[A_1A_2A_3] = 1/36 = (1/2)(1/2)(1/9) = P[A_1]P[A_2]P[A_3]$

  $P[A_1A_2] = 1/6 \neq 1/4 = P[A_1]P[A_2]$

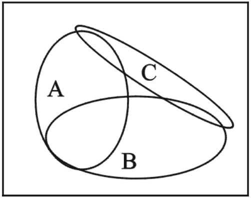
  $P[A_1A_3] = 1/36 \neq 1/18 = P[A_1]P[A_3]$

  $P[A_2A_3] = 1/12 \neq 1/18 = P[A_2]P[A_3]$

  $\therefore$ Just $P[ABC] = P[A]P[B]P[C]$ does not mean mutual independence. So, cannot imply pairwise independence, either.

Conditioning May Affect Independence

- Assume A and B are independent:

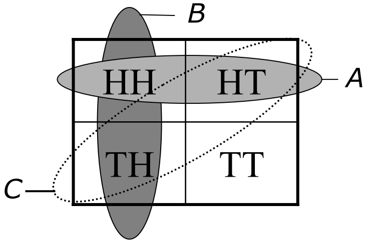


- If we are told that C occurred, are A and B independent?

Conditioning May Affect Independence

Example 1:

- Two independent fair ($p = 1/2$) coin tosses.
- Event A: First toss is H
- Event B: Second toss is H
- $P[A] = P[B] = 1/2$

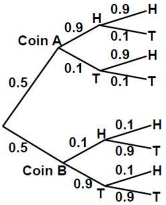


- Event C: The two outcomes are different.
- Conditioned on C, are A and B independent?

Conditioning May Affect Independence

Example 2: Choose at random among two unfair coins A and B

- $P[H|\text{coin } A] = 0.9$ and $P[H|\text{coin } B] = 0.1$
- Keep tossing the chosen coin



- Given coin A was selected, are future tosses independent? (Yes!)
- If we do not know which coin it is, are future tosses independent?
  - Compare $P[\text{toss } 11 = H]$ and $P[\text{toss } 11 = H \mid \text{first } 10 \text{ tosses are } H]$ ($\neq$!)

Conditionally Independent Events

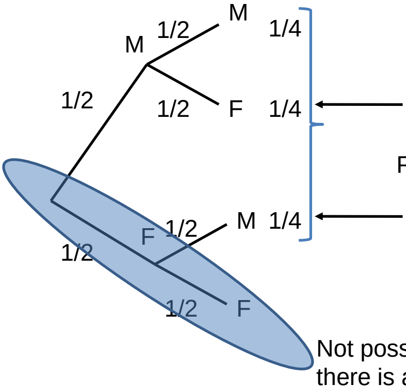
Definition: Two events A and B are conditionally independent given an event C with probability $P[C] \geq 0$ if

$$P[A \cap B | C] = P[A|C] \cdot P[B|C]$$

- We can show this is equivalent to $P[A|B \cap C] = P[A|C]$, i.e., if C is known, the additional information that B has occurred does not change the probability of A
- **Remark:** Independence does not imply conditional independence and vice versa!

The King's Sibling

- The king comes from a family of two children.
- What is the probability that his sibling is female?



$$P[\text{king has a sister}] = \frac{2/4}{3/4} = 2/3$$