CS70: Jean Walrand: Lecture 23.

Bayes' Rule, Independence, Mutual Independence

1. Conditional Probability: Review
2. Bayes' Rule: Another Look
3. Independence
4. Mutual Independence

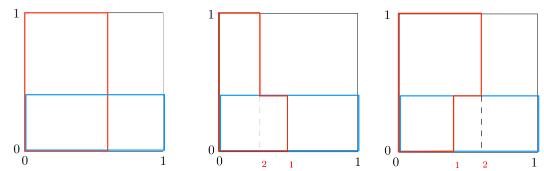
Conditional Probability: Review

Recall:

- $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$.
- Hence, $Pr[A \cap B] = Pr[B]Pr[A|B] = Pr[A]Pr[B|A]$.
- A and B are *positively correlated* if $Pr[A|B] > Pr[A]$,
  i.e., if $Pr[A \cap B] > Pr[A]Pr[B]$.
- A and B are *negatively correlated* if $Pr[A|B] < Pr[A]$,
  i.e., if $Pr[A \cap B] < Pr[A]Pr[B]$.
- Note: $B \subset A \Rightarrow A$ and B are positively correlated.
  ($Pr[A|B] = 1 > Pr[A]$)
- Note: $A \cap B = \emptyset \Rightarrow A$ and B are negatively correlated.
  ($Pr[A|B] = 0 < Pr[A]$)

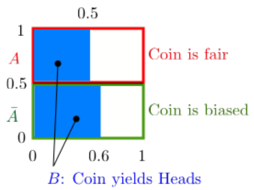
Conditional Probability: Pictures

Illustrations: Pick a point uniformly in the unit square



- Left: A and B are independent. $Pr[B] = b$; $Pr[B|A] = b$.
- Middle: A and B are positively correlated.
  $Pr[B|A] = b_1 > Pr[B|\bar{A}] = b_2$. Note: $Pr[B] \in (b_2, b_1)$.
- Right: A and B are negatively correlated.
  $Pr[B|A] = b_1 < Pr[B|\bar{A}] = b_2$. Note: $Pr[B] \in (b_1, b_2)$.

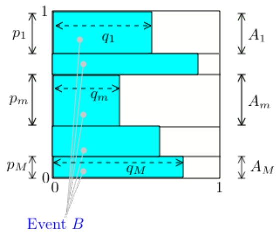
Bayes and Biased Coin



Pick a point uniformly at random in the unit square. Then

$Pr[A] = 0.5; Pr[\bar{A}] = 0.5$
$Pr[B|A] = 0.5; Pr[B|\bar{A}] = 0.6; Pr[A \cap B] = 0.5 \times 0.5$
$Pr[B] = 0.5 \times 0.5 + 0.5 \times 0.6 = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}]$
$$Pr[A|B] = \frac{0.5 \times 0.5}{0.5 \times 0.5 + 0.5 \times 0.6} = \frac{Pr[A]Pr[B|A]}{Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}]}$$
$\approx 0.46 =$ fraction of B that is inside A

Bayes: General Case

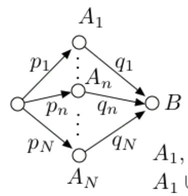


Pick a point uniformly at random in the unit square. Then

$Pr[A_m] = p_m, m = 1, \ldots, M$
$Pr[B|A_m] = q_m, m = 1, \ldots, M; Pr[A_m \cap B] = p_m q_m$
$Pr[B] = p_1 q_1 + \cdots p_M q_M$
$$Pr[A_m|B] = \frac{p_m q_m}{p_1 q_1 + \cdots p_M q_M} = \text{fraction of } B \text{ inside } A_m.$$

Bayes Rule

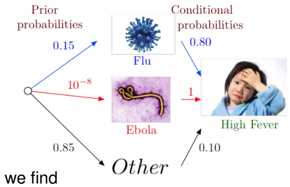
Another picture:



$p_n = Pr[A_n]$
$q_n = Pr[B|A_n]$
$A_1, \ldots, A_N$ disjoint
$A_1 \cup \cdots \cup A_N = \Omega$

$$Pr[A_n|B] = \frac{p_n q_n}{\sum_m p_m q_m}.$$

Why do you have a fever?



Using Bayes' rule, we find

$$Pr[\text{Flu}|\text{High Fever}] = \frac{0.15 \times 0.80}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.58$$

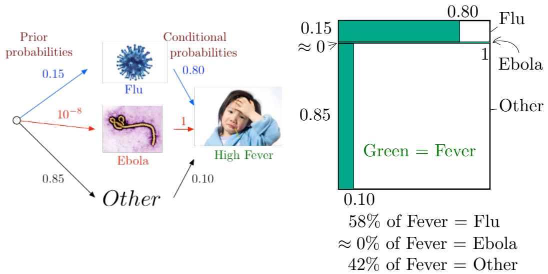
$$Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}$$

$$Pr[\text{Other}|\text{High Fever}] = \frac{0.85 \times 0.1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.42$$

The values $0.58, 5 \times 10^{-8}, 0.42$ are the posterior probabilities.

Why do you have a fever?

Our "Bayes' Square" picture:



Note that even though $Pr[\text{Fever}|\text{Ebola}] = 1$, one has

$$Pr[\text{Ebola}|\text{Fever}] \approx 0.$$

This example shows the importance of the prior probabilities.

Why do you have a fever?

We found

$$Pr[\text{Flu}|\text{High Fever}] \approx 0.58,$$
$$Pr[\text{Ebola}|\text{High Fever}] \approx 5 \times 10^{-8},$$
$$Pr[\text{Other}|\text{High Fever}] \approx 0.42$$

One says that 'Flu' is the Most Likely a Posteriori (MAP) cause of the high fever.
'Ebola' is the Maximum Likelihood Estimate (MLE) of the cause: it causes the fever with the largest probability.
Recall that

$$p_m = Pr[A_m], q_m = Pr[B|A_m], Pr[A_m|B] = \frac{p_m q_m}{p_1 q_1 + \cdots + p_M q_M}.$$

Thus,

- MAP = value of m that maximizes $p_m q_m$.
- MLE = value of m that maximizes q_m.

Independence

Definition: Two events A and B are **independent** if

$$Pr[A \cap B] = Pr[A]Pr[B].$$

Examples:

- When rolling two dice, $A =$ sum is 7 and $B =$ red die is 1 are independent;
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are not independent;
- When flipping coins, $A =$ coin 1 yields heads and $B =$ coin 2 yields tails are independent;
- When throwing 3 balls into 3 bins, $A =$ bin 1 is empty and $B =$ bin 2 is empty are not independent;

Independence and conditional probability

Fact: Two events A and B are **independent** if and only if

$$Pr[A|B] = Pr[A].$$

Indeed: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$, so that

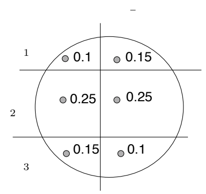
$$Pr[A|B] = Pr[A] \Leftrightarrow \frac{Pr[A \cap B]}{Pr[B]} = Pr[A] \Leftrightarrow Pr[A \cap B] = Pr[A]Pr[B].$$

Independence

Recall :

$$A \text{ and } B \text{ are independent}$$
$$\Leftrightarrow Pr[A \cap B] = Pr[A]Pr[B]$$
$$\Leftrightarrow Pr[A|B] = Pr[A].$$

Consider the example below:

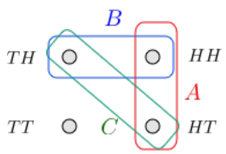


(A_2, B) are independent: $Pr[A_2|B] = 0.5 = Pr[A_2]$.
$(A_2, \bar{B})$ are independent: $Pr[A_2|\bar{B}] = 0.5 = Pr[A_2]$.
(A_1, B) are not independent: $Pr[A_1|B] = \frac{0.1}{0.5} = 0.2 \neq Pr[A_1] = 0.25$.

Pairwise Independence

Flip two fair coins. Let

- $A =$ 'first coin is H' $= \{HT, HH\}$;
- $B =$ 'second coin is H' $= \{TH, HH\}$;
- $C =$ 'the two coins are different' $= \{TH, HT\}$.



A, C are independent; B, C are independent;
$A \cap B, C$ are not independent. ($Pr[A \cap B \cap C] = 0 \neq Pr[A \cap B]Pr[C]$.)

If A did not say anything about C and B did not say anything about C, then $A \cap B$ would not say anything about C.

Example 2

Flip a fair coin 5 times. Let A_n = 'coin n is H', for $n = 1, \ldots, 5$.

Then,

$$A_m, A_n \text{ are independent for all } m \neq n.$$

Also,

$$A_1 \text{ and } A_3 \cap A_5 \text{ are independent.}$$

Indeed,

$$Pr[A_1 \cap (A_3 \cap A_5)] = \frac{1}{8} = Pr[A_1]Pr[A_3 \cap A_5]$$

. Similarly,

$$A_1 \cap A_2 \text{ and } A_3 \cap A_4 \cap A_5 \text{ are independent.}$$

This leads to a definition ....

Mutual Independence

Definition Mutual Independence

(a) The events $A_1, \ldots, A_5$ are mutually independent if

$$Pr[\cap_{k \in K} A_k] = \Pi_{k \in K} Pr[A_k], \text{ for all } K \subseteq \{1, \ldots, 5\}.$$

(b) More generally, the events $\{A_j, j \in J\}$ are mutually independent if

$$Pr[\cap_{k \in K} A_k] = \Pi_{k \in K} Pr[A_k], \text{ for all finite} K \subseteq J.$$

Thus, $Pr[A_1 \cap A_2] = Pr[A_1]Pr[A_2]$,
$Pr[A_1 \cap A_3 \cap A_4] = Pr[A_1]Pr[A_3]Pr[A_4], \ldots$.

Example: Flip a fair coin forever. Let A_n = 'coin n is H.' Then the events A_n are mutually independent.

Mutual Independence

Theorem

If the events $\{A_j, j \in J\}$ are mutually independent and if K_1 and K_2 are disjoint finite subsets of J, then any event V_1 defined by $\{A_j, j \in K_1\}$ is independent of any event V_2 defined by $\{A_j, j \in K_2\}$.

(b) More generally, if the K_n are pairwise disjoint finite subsets of J, then events V_n defined by $\{A_j, j \in K_n\}$ are mutually independent.

Proof:
See Lecture Note 25, Example 2.7. □

For instance, the fact that there are more heads than tails in the first five flips of a coin is independent of the fact there are fewer heads than tails in flips $6, \ldots, 13$.

Mutual Independence: Complements

Here is one step in the proof of the previous theorem.

Fact
Assume $A, B, C, \ldots, G, H$ are mutually independent.
Then, $A, B^c, C, \ldots, G^c, H$ are mutually independent.

Proof:

We show that
$Pr[A \cap B^c \cap C \cap \cdots \cap G^c \cap H] = Pr[A]Pr[B^c] \cdots Pr[G^c]Pr[H]$.

Assume that this is true when there are at most n complements.

Base case: $n = 0$ true by definition of mutual independence.

Induction step: Assume true for n. Check for $n+1$:

$A \cap B^c \cap C \cap \cdots \cap G^c \cap H =$
$A \cap B^c \cap C \cap \cdots \cap F \cap H \setminus A \cap B^c \cap C \cap \cdots \cap G \cap H$. Hence,

$$\begin{aligned} & Pr[A \cap B^c \cap C \cap \cdots \cap G^c \cap H] \\ &= Pr[A \cap B^c \cap C \cap \cdots \cap F \cap H] - Pr[A \cap B^c \cap C \cap \cdots \cap G \cap H] \\ &= Pr[A]Pr[B^c] \cdots Pr[F]Pr[H] - Pr[A]Pr[B^c] \cdots Pr[F]Pr[G]Pr[H] \\ &= Pr[A]Pr[B^c] \cdots Pr[F]Pr[H](1 - Pr[G]) \\ &= Pr[A]Pr[B^c] \cdots Pr[F]Pr[G^c]Pr[H]. \quad \square \end{aligned}$$

Summary.

Bayes' Rule, Independence, Mutual Independence

Main results:

- Bayes' Rule: $Pr[A_m|B] = p_m q_m / (p_1 q_1 + \cdots + p_M q_M)$.
- Mutual Independence: Events defined by disjoint collections of mutually independent events are mutually independent.