Probability Basics.

Probability Space.

1. **Sample Space:** Set of outcomes, Ω.

2. **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
   2.1 $0 \leq Pr[\omega] \leq 1$.
   2.2 $\sum_{\omega \in \Omega} Pr[\omega] = 1$.

Example: Two coins.

1. $\Omega = \{HH, HT, TH, TT\}$
   (Note: Not $\Omega = \{H, T\}$ with two picks!)
2. $Pr[HH] = \cdots = Pr[TT] = 1/4$

Consequences of Additivity

Theorem

(a) Inclusion/Exclusion: $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$;

(b) Union Bound: $Pr[A_1 \cup \cdots \cup A_n] \leq Pr[A_1] + \cdots + Pr[A_n]$;

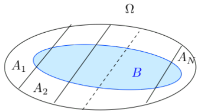
(c) Law of Total Probability:

If $A_1, \ldots A_N$ are a partition of Ω, i.e.,

pairwise disjoint and $\cup_{m=1}^{N} A_m = \Omega$, then

$$Pr[B] = Pr[B \cap A_1] + \cdots + Pr[B \cap A_N].$$

Proof Idea: Total probability.

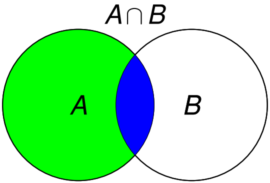


Add it up!

Conditional Probability.

Definition: The **conditional probability** of B given A is

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



In A!

In B?

Must be in $A \cap B$.

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

Note also:

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

Product Rule

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

Also: $Pr[A \cap B] = Pr[B|A]Pr[B]$

Theorem Product Rule
Let $A_1, A_2, \ldots, A_n$ be events. Then

$$Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}].$$

Simple Bayes Rule.

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

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

Bayes Rule: $Pr[A|B] = \frac{Pr[B|A]Pr[A]}{Pr[B]}$.

Is you coin loaded?

Your coin is fair w.p. $1/2$ or such that $Pr[H] = 0.6$, otherwise.

You flip your coin and it yields heads.

What is the probability that it is fair?

Analysis:

$$A = \text{ 'coin is fair'}, B = \text{ 'outcome is heads'}$$

We want to calculate $P[A|B]$.

We know $P[B|A] = 1/2, P[B|\bar{A}] = 0.6, Pr[A] = 1/2 = Pr[\bar{A}]$

Now,

$$\begin{aligned} Pr[B] &= Pr[A \cap B] + Pr[\bar{A} \cap B] = Pr[A]Pr[B|A] + Pr[\bar{A}]Pr[B|\bar{A}] \\ &= (1/2)(1/2) + (1/2)0.6 = 0.55. \end{aligned}$$

Thus,

$$Pr[A|B] = \frac{Pr[A]Pr[B|A]}{Pr[B]} = \frac{(1/2)(1/2)}{(1/2)(1/2) + (1/2)0.6} \approx 0.45.$$

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; $Pr[A \cap B] = \frac{1}{36}$, $Pr[A]Pr[B] = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right)$.
- When rolling two dice, $A =$ sum is 3 and $B =$ red die is 1 are not independent; $Pr[A \cap B] = \frac{1}{36}$, $Pr[A]Pr[B] = \left(\frac{2}{36}\right)\left(\frac{1}{6}\right)$.
- When flipping coins, $A =$ coin 1 yields heads and $B =$ coin 2 yields tails are independent; $Pr[A \cap B] = \frac{1}{4}$, $Pr[A]Pr[B] = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$.
- When throwing 3 balls into 3 bins, $A =$ bin 1 is empty and $B =$ bin 2 is empty are not independent;
  $Pr[A \cap B] = \frac{1}{27}$, $Pr[A]Pr[B] = \left(\frac{8}{27}\right)\left(\frac{8}{27}\right)$.

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

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

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 $Pr[A|B] > Pr[A]$,

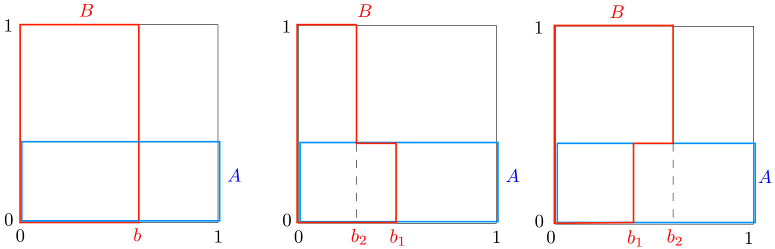
  i.e., if $Pr[A \cap B] > Pr[A]Pr[B]$.
- A and B are *negatively correlated* if $Pr[A|B] < Pr[A]$,

  i.e., if $Pr[A \cap B] < Pr[A]Pr[B]$.
- A and B are *independent* 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.
  $(Pr[A|B] = 0 < Pr[A])$

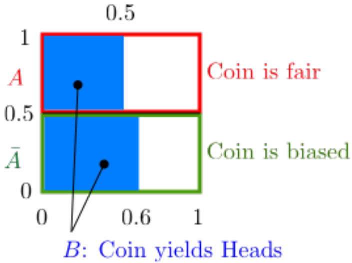
Conditional Probability: Pictures

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: $Pr[B] \in (b_1, b_2)$.

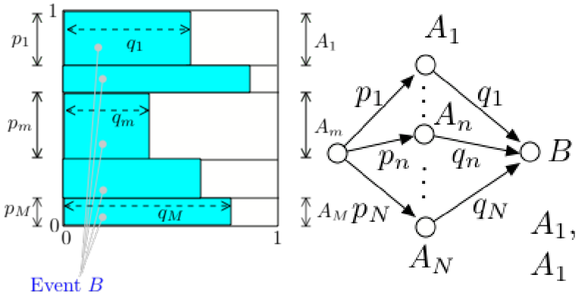
Bayes and Biased Coin



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

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

Bayes: General Case



$$p_n = Pr[A_n]$$
$$q_n = Pr[B|A_n]$$

$A_1, \dots, A_N$ disjoint

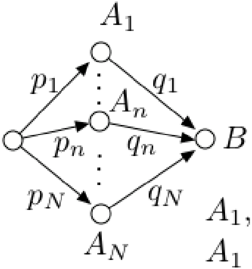
$A_1 \cup \cdots \cup A_N = \Omega$

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

$$Pr[A_n] = p_n, n = 1, \dots, N$$
$$Pr[B|A_n] = q_n, n = 1, \dots, N; Pr[A_n \cap B] = p_n q_n$$
$$Pr[B] = p_1 q_1 + \cdots p_N q_N$$
$$Pr[A_n|B] = \frac{p_n q_n}{p_1 q_1 + \cdots p_N q_N} = \text{ fraction of } B \text{ inside } A_n.$$

Bayes Rule

A general picture: We imagine that there are N possible causes $A_1, \dots, A_N$.



$$p_n = Pr[A_n]$$

$$q_n = Pr[B|A_n]$$

$$A_1, \dots, A_N \text{ disjoint}$$

$$A_1 \cup \cdots \cup A_N = \Omega$$

100 situations: $100 p_n q_n$ where A_n and B occur, for $n = 1, \dots, N$.
In $100 \sum_m p_m q_m$ occurences of B, $100 p_n q_n$ occurrences of A_n.

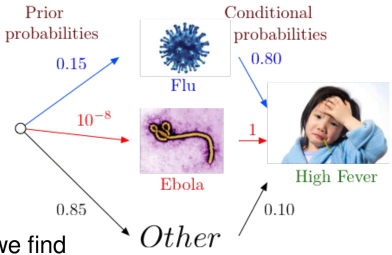
Hence,

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

But, $p_n = Pr[A_n], q_n = Pr[B|A_n], \sum_m p_m q - m = Pr[B]$, hence,

$$Pr[A_n|B] = \frac{Pr[B|A_n]Pr[A_n]}{Pr[B]}.$$

Why do you have a fever?



Using Bayes' rule, we find

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

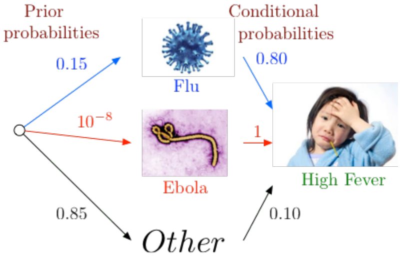
$$Pr[\text{Ebola}|\text{High Fever}] = \frac{10^{-8} \times 1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 5 \times 10^{-8}$$

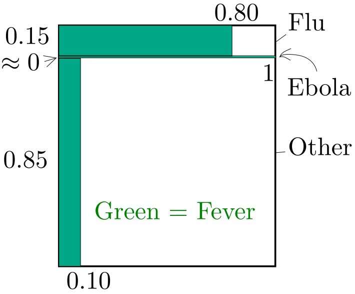
$$Pr[\text{Other}|\text{High Fever}] = \frac{0.85 \times 0.1}{0.15 \times 0.80 + 10^{-8} \times 1 + 0.85 \times 0.1} \approx 0.42$$

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

Why do you have a fever?

Our “Bayes’ Square” picture:





58% of Fever = Flu
$\approx 0\%$ of Fever = Ebola
42% of Fever = Other

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

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

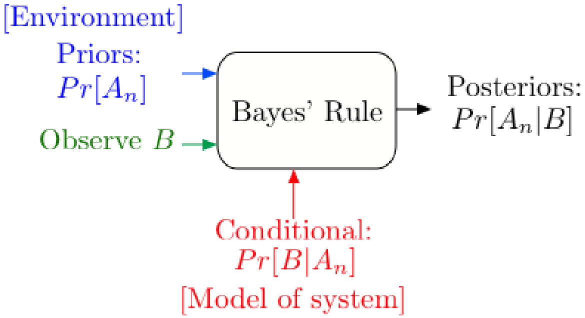
'Flu' is Most Likely a Posteriori (MAP) cause of high fever.
'Ebola' is Maximum Likelihood Estimate (MLE) of cause:
causes fever with largest probability.
Recall that

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

Thus,

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

Bayes' Rule Operations



Bayes' Rule: canonical example of how information changes our opinions.

Thomas Bayes

Thomas Bayes

Portrait used of Bayes in a 1936 book,[1] but it is doubtful whether the portrait is actually of him.[2] No earlier portrait or claimed portrait survives.

Born	c. 1701 London, England
Died	7 April 1761 (aged 59) Tunbridge Wells, Kent, England
Residence	Tunbridge Wells, Kent, England
Nationality	English
Known for	Bayes' theorem

Source: Wikipedia.

Thomas Bayes

A

FIG. 3. *Joshua Bayes (1671–1746).*

$$\xrightarrow{Pr[B|A]}$$

Max over B

B

Thomas Bayes?

A Bayesian picture of Thomas Bayes.

Testing for disease.

Random Experiment: Pick a random male.
Outcomes: $(test, disease)$
A - prostate cancer.
B - positive PSA test.

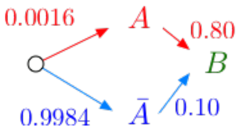
- $Pr[A] = 0.0016$, (.16 % of the male population is affected.)
- $Pr[B|A] = 0.80$ (80% chance of positive test with disease.)
- $Pr[B|\overline{A}] = 0.10$ (10% chance of positive test without disease.)

From http://www.cpcn.org/01_psa_tests.htm and http://seer.cancer.gov/statfacts/html/prost.html (10/12/2011.)

Positive PSA test (B). Do I have disease?

$$Pr[A|B]???$$

Bayes Rule.



Using Bayes' rule, we find

$$P[A|B] = \frac{0.0016 \times 0.80}{0.0016 \times 0.80 + 0.9984 \times 0.10} = .013.$$

A 1.3% chance of prostate cancer with a positive PSA test.

Surgery anyone?

Impotence...

Incontinence..

Death.

Quick Review

Events, Conditional Probability, Independence, Bayes' Rule

Key Ideas:

- Conditional Probability:

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

- Independence: $Pr[A \cap B] = Pr[A]Pr[B]$.
- Bayes' Rule:

  $$Pr[A_n|B] = \frac{Pr[A_n]Pr[B|A_n]}{\sum_m Pr[A_m]Pr[B|A_m]}.$$

  $Pr[A_n|B] =$ posterior probability; $Pr[A_n] =$ prior probability .

- All these are possible:

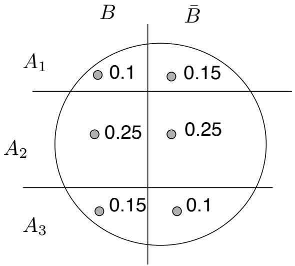
  $$Pr[A|B] < Pr[A]; Pr[A|B] > Pr[A]; Pr[A|B] = Pr[A].$$

Independence

Recall :

$$
\begin{aligned}
&A \text{ and } B \text{ are independent}\\
&\quad \Leftrightarrow Pr[A \cap B] = Pr[A]Pr[B]\\
&\quad \Leftrightarrow Pr[A|B] = Pr[A].
\end{aligned}
$$

Consider the example below:

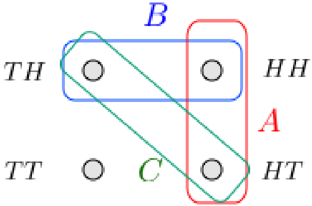


(A_2, B) are independent: $Pr[A_2|B] = 0.5 = Pr[A_2]$.
$(A_2, \bar{B})$ are independent: $Pr[A_2|\bar{B}] = 0.5 = Pr[A_2]$.
(A_1, B) are not independent: $Pr[A_1|B] = \frac{0.1}{0.5} = 0.2 \neq Pr[A_1] = 0.25$.

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. ($Pr[A \cap B \cap C] = 0 \neq Pr[A \cap B]Pr[C]$.)

False: 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

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

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

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

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

Mutual Independence

Theorem

(a) 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

$$\cap_{k \in K_1} A_k \text{ and } \cap_{k \in K_2} A_k \text{ are independent.}$$

(b) More generally, if the K_n are pairwise disjoint finite subsets of J, then the events

$$\cap_{k \in K_n} A_k \text{ are mutually independent.}$$

(c) Also, the same is true if we replace some of the A_k by $\bar{A}_k$.

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 $Pr[A \cap B] < Pr[A]Pr[B]$.
- A and B are *independent* 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.
  $(Pr[A|B] = 0 < Pr[A])$

Quick Review.

$\boxed{\text{Bayes' Rule, Mutual Independence, Collisions and Collecting}}$

Main results:

- Bayes' Rule: $Pr[A_m|B] = p_m q_m/(p_1 q_1 + \cdots + p_M q_M)$.
- Product Rule:
  $Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}]$.