Independence

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

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

**Examples:**
- When rolling two dice, $A = \text{sum is 7}$ and $B = \text{red die is 1}$ are independent; $\Pr[A \cap B] = \frac{1}{6}, \Pr[A]\Pr[B] = \frac{1}{36} \cdot \frac{1}{36} = \frac{1}{1296}$.
- When rolling two dice, $A = \text{sum is 3}$ and $B = \text{red die is 1}$ are not independent; $\Pr[A \cap B] = \frac{1}{6}, \Pr[A]\Pr[B] = \frac{1}{36} \cdot \frac{1}{36} = \frac{1}{1296}$.
- When flipping coins, $A = \text{coin 1 yields heads}$ and $B = \text{coin 2 yields tails}$ are independent; $\Pr[A \cap B] = \frac{1}{2}, \Pr[A]\Pr[B] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.
- When throwing 3 balls into 3 bins, $A = \text{bin 1 is empty}$ and $B = \text{bin 2 is empty}$ are not independent; $\Pr[A \cap B] = \frac{1}{8}, \Pr[A]\Pr[B] = \frac{1}{8} \cdot \frac{1}{8} = \frac{1}{64}$.

Causality vs. Correlation

Events $A$ and $B$ are **positively correlated** if

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

(E.g., smoking and lung cancer.)

A and $B$ being positively correlated does not mean that $A$ causes $B$ or that $B$ causes $A$.

Other examples:
- Tesla owners are more likely to be rich. That does not mean that poor people should buy a Tesla to get rich.
- People who go to the opera are more likely to have a good career. That does not mean that going to the opera will improve your career.
- Rabbits eat more carrots and do not wear glasses. Are carrots good for eyesight?

Proving Causality

Proving causality is generally difficult. One has to eliminate external causes of correlation and be able to test the cause/effect relationship (e.g., randomized clinical trials).

Some difficulties:
- $A$ and $B$ may be positively correlated because they have a common cause. (E.g., being a rabbit.)
- If $B$ precedes $A$, then $B$ is more likely to be the cause. (E.g., smoking.) However, they could have a common cause that induces $B$ before $A$. (E.g., smart, CS70, Tesla.)

More about such questions later. For fun, check "N. Taleb: Fooled by randomness."

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

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

Total probability

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_N$.

Then,

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

Indeed, $B$ is the union of the disjoint sets $A_n \cap B$ for $n = 1, \ldots, N$. Thus,

$$\Pr[B] = \Pr[A_1]\Pr[B|A_1] + \cdots + \Pr[A_N]\Pr[B|A_N].$$
**Total probability**

Assume that $\Omega$ is the union of the disjoint sets $A_1, \ldots, A_n$.

\[ P(A) = P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n). \]

**Bayes Rule**

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

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

Imagine 100 situations, among which $m := 100 \times (1/2)(1/2)$ are such that $A$ and $B$ occur. And among the $100 \times (1/2)(1/2)$ situations where $B$ occurred, there are $m := 100 \times (1/2)(1/2) \times (1/2)(1/2) = 25$ situations where $A$ occurred.

Hence,

\[ P(A|B) = \frac{P(A)P(B|A)}{P(B)} = \frac{(1/2)(1/2)(1/2)(1/2)(1/2)}{25} = 0.45. \]

**Is you coin loaded?**

Your coin is fair w.p. $1/2$ or such that $P(H) = 0.6$, 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 $P(B|A) = 1/2$, $P(B|\bar{A}) = 0.6$, $P(A) = 1/2 = P(\bar{A})$

Now,

\[ P(B) = P(A \cap B) + P(\bar{A} \cap B) = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) = (1/2)(1/2) + (1/2)(0.6) = 0.55. \]

Thus,

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

**Conditional Probability: Pictures**

Illustrations: Pick a point uniformly in the unit square

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

**Bayes and Biased Coin**

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

\[ P(A) = 0.5, P(\bar{A}) = 0.5 \]
\[ P(B|A) = 0.5, P(B|\bar{A}) = 0.6, P(A \cap B) = 0.5 \times 0.5 \]
\[ P(B) = 0.5 \times 0.5 + 0.5 \times 0.6 = P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) \]
\[ P(A|B) = \frac{0.5 \times 0.5 \times 0.5 \times 0.6}{0.5 \times 0.5 \times 0.5 \times 0.5 + 0.5 \times 0.5 \times 0.5 \times 0.6} = 0.46 \text{ fraction of } B \text{ that is inside } A \]
Bayes: General Case

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

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

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.

Bayes’ Rule Operations

\[ Pr[A_n] = p_n, n = 1, \ldots, N \]
\[ Pr[B|A_n] = q_n, n = 1, \ldots, N; \]
\[ Pr[A_n|B] = \frac{p_n q_n}{p_1 q_1 + \cdots + p_M q_M} \]

Thus,

- MAP = value of \( m \) that maximizes \( p_m q_m \).
- MLE = value of \( m \) that maximizes \( q_m \).

Bayes’ Rule is the canonical example of how information changes our opinions.

Why do you have a fever?

Our “Bayes’ Square” picture:

- 58% of Fever = Flu
- 42% of Fever = Other
- 0% of Fever = Ebola

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.

Thomas Bayes

Quick Review

Events, Conditional Probability, Independence, Bayes’ Rule

Key Ideas:
- Conditional Probability:
  \[ P[A|B] = \frac{P[A \cap B]}{P[B]} \]
- Bayes’ Rule:
  \[ P[A_i|B] = \frac{P[A_i|B]P[B|A_i]}{\sum_i P[A_i|B]P[B|A_i]} \]
  \( P[A_i|B] \) = posterior probability; \( P[A_i] \) = prior probability.
- All these are possible:
- Independence:
  \( A \) and \( B \) are independent
  \[ \iff P[A|B] = P[A|P[B] \]
  \[ \iff P[A|B] = P[A] \]
- Consider the example below:

\( (A_2, B) \) are independent: \( P[A_2|B] = 0.5 = P[A_2] \)
\( (A_2, B) \) are independent: \( P[A_2|B] = 0.5 = P[A_2] \)
\( (A_1, B) \) are not independent: \( P[A_1|B] = 0.2 \neq P[A_1] = 0.25 \).

Testing for disease.

Random Experiment: Pick a random male.
Outcomes: (test, disease)
- \( A \) - prostate cancer.
- \( B \) - positive PSA test.
- \( P[A] = 0.0016 \) (1.6% of the male population is affected.)
- \( P[B|A] = 0.80 \) (80% chance of positive test with disease.)
- \( P[B|\overline{A}] = 0.10 \) (10% chance of positive test without disease.)


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.

Bayes Rule.

\[ A \rightleftharpoons \begin{array}{c} 0.0016 \ A \ 0.80 \ B \ \ \ \\ \end{array} \]

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.

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" = \( \{TH, HT\}. \)

\[ A, C \text{ are independent}; \ B, C \text{ are independent}; \]
\( A \cap B, C \text{ are not independent}. \ (P[A \cap B \cap C] \neq 0 \neq P[A \cap B|P[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 $A_n = \text{`coin } n \text{ is H'},$ for $n = 1, \ldots, 5.$
Then, $A_m, A_n$ are independent for all $m \neq n.$
Also, $A_1$ and $A_3 \cap A_5$ 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,
$A_1 \cap A_2$ and $A_3 \cap A_4 \cap A_5$ are independent.
This leads to a definition ...

Mutual Independence

**Definition Mutual Independence**
(a) The events $A_1, \ldots, A_5$ are mutually independent if
$$\Pr[\cap_{k \in K} A_k] = \prod_{k \in K} \Pr[A_k],$$
for all $K \subseteq \{1, \ldots, 5\}.$
(b) More generally, the events $\{A_j, j \in J\}$ are mutually independent if
$$\Pr[\cap_{k \in K} A_k] = \prod_{k \in K} \Pr[A_k],$$
for all finite $K \subseteq J.$
Example: Flip a fair coin forever. Let $A_n = \text{`coin } n \text{ is H'}.$ 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
$$\cap_{k \in K_n} A_k$$
are mutually independent.
(c) Also, the same is true if we replace some of the $A_k$ by $\bar{A}_k.$
**Proof:**
See Notes 25, 2.7.