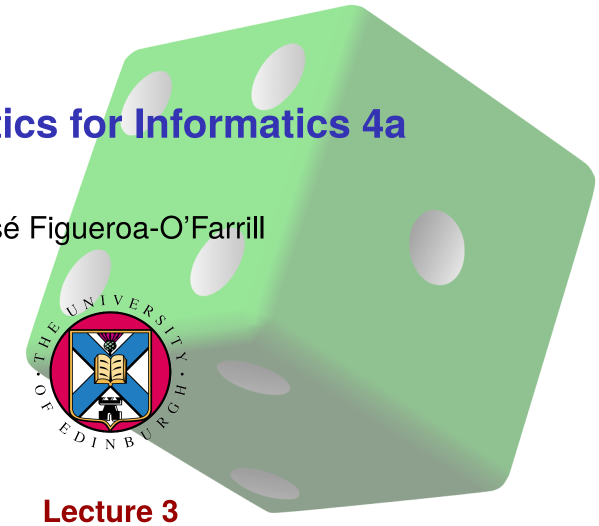
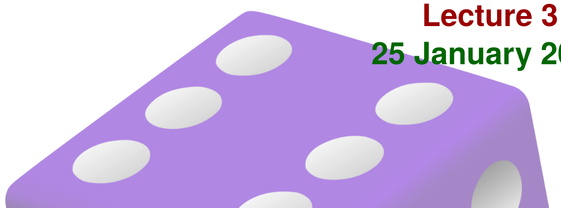



# Mathematics for Informatics 4a

José Figueroa-O'Farrill

**Lecture 3**

**25 January 2012**



## The story of the film so far...

With every experiment we associate a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where

- $\Omega$ is the set of all possible outcomes of the experiment;
- $\mathcal{F}$ is a $\sigma$-field of subsets of $\Omega$: containing $\Omega$ and closed under complementation and countable unions; and
- $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a function normalised to $\mathbb{P}(\Omega) = 1$ and countably additive over disjoint unions.

We also introduced uniform probability spaces with $\Omega$ a finite set and $\mathbb{P}(A) = |A|/|\Omega|$ for every event $A$.

**The thought of the day**

Probability is a measure of our ignorance and hence, when our knowledge about a system changes, the probability of events should also change to reflect the new knowledge.



**Example (Rolling two dice with increasing score)**

*We roll two fair dice in turn. Let $A$ be the event "the second die shows a higher score than the first". What is $\mathbb{P}(A)$?*

There are $6^2$ possible outcomes of which $\binom{6}{2} = 15$ lie in $A$.

Hence $\mathbb{P}(A) = \frac{15}{36} = \frac{5}{12}$.

*Now suppose that the first die turns out to be ⚄. What is $\mathbb{P}(A)$?*

There is now only one positive outcome: namely, (⚄, ⚅).

But the sample space has changed as well.

Once we know that the first die is ⚄, the sample space consists of six outcomes:

{(⚄, ⚀), (⚄, ⚁), (⚄, ⚂), (⚄, ⚃), (⚄, ⚄), (⚄, ⚅)}

and hence $\mathbb{P}(A) = \frac{1}{6}$.



**Example (Alice and Bob have children)**

*Alice and Bob have two children. What is the probability that they are both girls?*

Assuming a uniform distribution (i.e., boy or girl is equally likely), every outcome in our sample space

$$\Omega = \{(♀, ♀), (♀, ♂), (♂, ♀), (♂, ♂)\}$$

has probability $\frac{1}{4}$. The desired outcome $A = (♀, ♀)$ has $\mathbb{P}(A) = \frac{1}{4}$.

*Now suppose that we know that one of the children is a girl. What is the probability now?*

The sample space is now $\{(♀, ♀), (♀, ♂), (♂, ♀)\}$, whereas there is still only one positive outcome, whence $\mathbb{P}(A) = \frac{1}{3}$.

*What about if we know that the oldest child is a girl?*

The sample space is now $\{(♀, ♀), (♀, ♂)\}$, whence $\mathbb{P}(A) = \frac{1}{2}$.

## Conditional probability

**Definition**

Let $A, B$ be events. The **conditional probability** $\mathbb{P}(A|B)$ of the event of “$A$ occurs given that $B$ occured” is

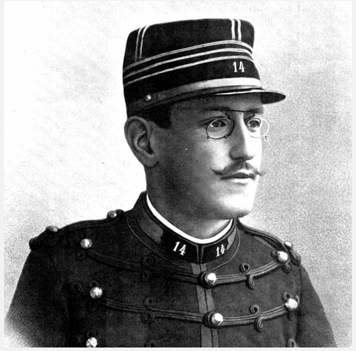
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

assuming that $\mathbb{P}(B) > 0$.

**Warning!**

$\mathbb{P}(A|B) \neq \mathbb{P}(B|A)$, unless $\mathbb{P}(A) = \mathbb{P}(B)$. The incorrect equality formed the basis of the prosecution’s case in the infamous *affaire Dreyfus* in France at the turn of the 20th century. To find out more, visit my colleague Andrew Ranicki’s page on this subject: `http://www.maths.ed.ac.uk/~aar/dreyfus.htm`.





This Dreyfus...

... not *this* Dreyfus!



## Conditional probability from uniform probability

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a uniform probability space (e.g., $\Omega$ finite and all outcomes equally probable) and $A, B$ are two events, we have

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} \quad \text{and} \quad \mathbb{P}(B) = \frac{|B|}{|\Omega|} .$$

If $B$ occurs, the possible outcomes are those in $B$ and they remain equally likely with probability $1/|B|$.
The event $A$ occurs if and only if $A \cap B$ occurs, whence the probability that $A$ occurs given that $B$ occurs is

$$\frac{|A \cap B|}{|B|} = \frac{|A \cap B|}{|\Omega|} \Big/ \frac{|B|}{|\Omega|} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A|B) .$$



## Conditional probability as relative frequency

We repeat the experiment $N$ times, with $B$ occuring $N(B)$ times. Since we know that $B$ occurs we are in one of those $N(B)$ trials. The event $A$ occurs in $N(A \cap B)$ of them, so the probability of $A$ occuring given that $B$ does is given by the limit $N \to \infty$ of

$$\frac{N(A \cap B)}{N(B)} = \frac{N(A \cap B)}{N} \Big/ \frac{N(B)}{N} \to \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A|B) .$$

We will now revisit the previous examples in the language of conditional probability.

### Example (Rolling two dice with increasing score)

Let $A$ be the event "the second die shows a greater score than the first" and let $B$ be the event "the first die shows a ⚄".
Then

$$B = \{(⚄, ⚀), (⚄, ⚁), (⚄, ⚂), (⚄, ⚃), (⚄, ⚄), (⚄, ⚅)\}$$
$$A \cap B = \{(⚄, ⚅)\}$$

whence

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{|A \cap B|}{|B|} = \tfrac{1}{6}.$$



### Example (Alice and Bob have children)

Let

$$A = \{(♀, ♀)\} \qquad \text{"both are girls"}$$
$$B = \{(♀, ♀), (♀, ♂), (♂, ♀)\} \qquad \text{"at least one is a girl"}$$
$$C = \{(♀, ♀), (♀, ♂)\} \qquad \text{"the oldest is a girl" .}$$

Then

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{|A \cap B|}{|B|} = \tfrac{1}{3}$$

and

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{|A \cap C|}{|C|} = \tfrac{1}{2}$$



### Example

*A box contains a double-headed coin, a double-tailed coin and a conventional coin. A coin is picked at random, tossed and the result is a head. What is the probability that it is the double-headed coin?*

Let $D$ be the event that we did pick the double-headed coin and let $H$ be the event that the coin we picked and tossed, came up heads. We want to calculate $\mathbb{P}(D|H)$.
There are 3 coins and hence 6 possible outcomes, of which 3 are heads. Therefore $\mathbb{P}(H) = \frac{3}{6} = \frac{1}{2}$.
Of those 3 outcomes, two come from picking the double-headed coin, whence $\mathbb{P}(D \cap H) = \frac{2}{6} = \frac{1}{3}$.
Therefore

$$\mathbb{P}(D|H) = \frac{\mathbb{P}(D \cap H)}{\mathbb{P}(H)} = \tfrac{1}{3}\Big/\tfrac{1}{2} = \tfrac{2}{3} .$$



## Conditional probability is a probability

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $B \in \mathcal{F}$ be an event with $\mathbb{P}(B) > 0$. Let $\mathbb{P}'(A) := \mathbb{P}(A|B)$.
Then $\mathbb{P}'$ is again a probability measure:

- $\mathbb{P}' : \mathcal{F} \to [0, 1]$
- $\mathbb{P}'(\Omega) = 1$
- $\mathbb{P}'$ is countably additive over disjoint unions

This means that all results proved for general probability measures apply to the conditional probability as well.
In fact, we can define a new probability space $(B, \mathcal{F}', \mathbb{P}')$, where

$$\mathcal{F}' = \{A \cap B | A \in \mathcal{F}\} .$$

Recall that $\mathcal{F}'$ is again a $\sigma$-field (cf. Tutorial Sheet 1), only this time of subsets of $B$.

## Multiplication rule

From the ▸ definition of conditional probability there follows the

**Multiplication rule**

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$$

which holds even if $\mathbb{P}(B) = 0$.

**Example**

*I have 5 red socks and 3 black socks in a drawer. I pick two socks at random. What is the probability that I get a black pair?*
Let $P =$ "the pair is black" and $B =$ "the first sock is black". Then clearly $P \subset B$, so $P = P \cap B$. Also $\mathbb{P}(B) = \frac{3}{8}$ and $\mathbb{P}(P|B) = \frac{2}{7}$, whence

$$\mathbb{P}(P) = \mathbb{P}(P \cap B) = \mathbb{P}(P|B)\mathbb{P}(B) = \tfrac{2}{7} \times \tfrac{3}{8} = \tfrac{3}{28}.$$



## Extended multiplication rule

**Theorem**

*Let $A_1, A_2, \dots, A_n$ be events with $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) > 0$. Then*

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_n|A_{n-1} \cap \cdots \cap A_1)\times \\ \times \mathbb{P}(A_{n-1}|A_{n-2} \cap \cdots \cap A_1) \cdots \mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$

**Proof.**

Just use that for $k = 2, \dots, n$,

$$\mathbb{P}(A_k|A_{k-1} \cap \cdots \cap A_1) = \frac{\mathbb{P}(A_k \cap \cdots \cap A_1)}{\mathbb{P}(A_{k-1} \cap \cdots \cap A_1)}$$

and multiply the terms in the RHS to obtain the LHS. □



**Example**

*A box contains 15 numbered balls: six balls have the number 1, four balls the number 2 and five balls the number 3. Suppose we draw three balls without replacement. What is the probability that we draw the balls in increasing numerical order?*
Let $A$ be event "first ball has number 1", $B$ the event "second ball has number 2" and $C$ the event "third ball has number 3".
We want to compute $\mathbb{P}(A \cap B \cap C)$.
By the extended multiplication rule,

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(C|A \cap B)\mathbb{P}(B|A)\mathbb{P}(A)$$

where $\mathbb{P}(A) = \frac{6}{15} = \frac{2}{5}$, $\mathbb{P}(B|A) = \frac{4}{14} = \frac{2}{7}$ and $\mathbb{P}(C|A \cap B) = \frac{5}{13}$, whence

$$\mathbb{P}(A \cap B \cap C) = \tfrac{2}{5} \times \tfrac{2}{7} \times \tfrac{5}{13} = \tfrac{4}{91} \simeq 4.4\%$$



**Example**

*A bag contains 26 cards, each one with a letter in $\{a, b, c, \dots, z\}$. You take 7 cards at random without replacement. What is the probability that you can spell "dreyfus" with them?*
Let $A_i$ be the event that the $i$th card is one with a letter in $\{d, r, e, y, f, u, s\}$. We are after $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_7)$. We use the extended multiplication rule:

$$\begin{aligned}\mathbb{P}(A_1 \cap \cdots \cap A_7) &= \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\cdots\mathbb{P}(A_7|A_6 \cap \cdots \cap A_1) \\ &= \tfrac{7}{26} \times \tfrac{6}{25} \times \tfrac{5}{24} \times \cdots \times \tfrac{1}{20} \\ &= \tfrac{7\times6\times\cdots\times1}{26\times25\times\cdots\times20} \times \tfrac{19!}{19!} \\ &= \binom{26}{7}^{-1} = \tfrac{1}{657800}.\end{aligned}$$

## Independence

Heuristically, two events $A$, $B$ are said to be independent if the chance of one occurring is not altered by the other's occurrence; that is, $\mathbb{P}(A|B) = \mathbb{P}(A)$. The multiplication rule the implies that

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B)$$

and, by symmetry, $\mathbb{P}(B|A) = \mathbb{P}(B)$.

**Definition**

Two events $A$, $B$ are **independent** if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

**Example**

"Not winning the lottery last week" and "winning the lottery this week" are independent events.



Suppose that we have three events $A$, $B$, $C$. We say that they are **independent** if any two are independent:

$$\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B) \qquad \mathbb{P}(A\cap C) = \mathbb{P}(A)\mathbb{P}(C) \qquad \mathbb{P}(B\cap C) = \mathbb{P}(B)\mathbb{P}(C)$$

**and in addition**

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) .$$

**Warning!**

Neither of the two sets of conditions imply the other.

The reason we impose all conditions is that independence means also that no two events occurring can influence the chance of a third event occurring.



**Example**

Consider again the example with the numbered balls, but now with replacement. The events $A$, $B$ and $C$ are now independent:

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{6}{15} \times \frac{4}{15} \times \frac{5}{15} = \frac{8}{225} \simeq 3.6\%$$

**Example**

*What is the probability of getting at least one ⚃ in two rolls of a fair die?* Let $A_i$ be the event of getting ⚃ in the $i$th roll. We are after $\mathbb{P}(A_1 \cup A_2)$. The two events are independent, whence by inclusion-exclusion

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)\mathbb{P}(A_1)\mathbb{P}(A_2)$$
$$= \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$



**Example**

*A fair coin is tossed* $10$ *times. What is the probability of getting all heads? of getting* $HTHTHTHTHT$*? of getting* $TTTTTTTTTH$*?*

Getting $H$ or $T$ in one toss of the coin is independent from getting $H$ or $T$ on a different toss. Since the coin if fair, $\mathbb{P}(H) = \mathbb{P}(T) = \frac{1}{2}$, whence the probability of getting any one desired outcome after $10$ tosses is $(\frac{1}{2})^{10}$.

*What about if the coin is not fair?*

Let $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = q = 1 - p$. Then by independence,

$$\mathbb{P}(HHHHHHHHHH) = \mathbb{P}(H)^{10} = p^{10}$$
$$\mathbb{P}(HTHTHTHTHT) = \mathbb{P}(H)\mathbb{P}(T)\mathbb{P}(H)\ldots\mathbb{P}(T) = p^5(1-p)^5$$
$$\mathbb{P}(TTTTTTTTTH) = \mathbb{P}(T)^9\mathbb{P}(H) = (1-p)^9 p$$

## Independence vs. pairwise independence

**Definition**

A family $A_1, A_2, \ldots, A_n$ of events is **pairwise independent** if

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j) \qquad \forall \quad 1 \leqslant i < j \leqslant n$$

and it is **independent** if, for any $1 \leqslant j_1 < j_2 < \cdots < j_k \leqslant n$,

$$\mathbb{P}(A_{j_1} \cap A_{j_2} \cap \cdots \cap A_{j_k}) = \mathbb{P}(A_{j_1})\mathbb{P}(A_{j_2}) \cdots \mathbb{P}(A_{j_k})$$

Whereas independence implies pairwise independence, the converse **is not true**.



**Example**

Two fair coins are tossed. Consider the events

- $A =$ "the first toss is a head"
- $B =$ "the second toss is a head"
- $C = A \triangle B =$ "exactly one toss is a head"

Then $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$ and

$$\mathbb{P}(A \cap B) = \tfrac{1}{4} = \mathbb{P}(A)\mathbb{P}(B)$$
$$\mathbb{P}(A \cap C) = \tfrac{1}{4} = \mathbb{P}(A)\mathbb{P}(C)$$
$$\mathbb{P}(B \cap C) = \tfrac{1}{4} = \mathbb{P}(B)\mathbb{P}(C) \ ,$$

whence the events are pairwise independent. However

$$\mathbb{P}(A \cap B \cap C) = 0 \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) \ ,$$

and the events are **not** independent.



## Summary

- The conditional probability of $A$ occuring given that $B$ occurs is
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$
- Conditional probability **is** a probability
- Multiplication rules:
$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$
$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_3|A_1 \cap A_2)\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$
- Events $A$, $B$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
- More generally, a countable family $\{A_i\}$, $i \in I$, of events are independent if
$$\mathbb{P}\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} \mathbb{P}(A_j)$$
for every finite subset $J \subseteq I$