Chapter 2

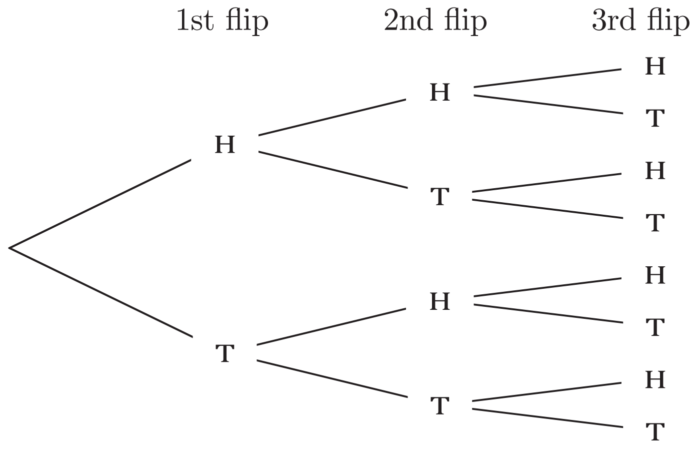
Probability

2.1 Combinatorial counting

Combinatorial counting is the study of different ways of arranging objects in a finite set.
Finite sets.
Defintion : 1. A non-empty set Ω is finite if we can index its elements as :
$\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. This n is the number of elements of Ω, it is called the cardinal of Ω and denoted by $Card\Omega$, $\#\Omega$ or $|\Omega|$.
2. The set ϕ (empty set) is finite with cardinal 0

2.1.1 The counting principle

Consider an experiment E which can be devided into r experiments $E_1, \ldots, E_r$. If E_1 has n_1 possible results and E_2 has n_2 possible results, ... , E_r has n_r results, then E has $n_1.n_2...n_r$ possible results.
Example : A child plays "heads or tails" by fliping a coin three times consicutively and noting the result of each flip. How many possibilities does the child encounter?



We shall now give the first elements of Combinatorics and apply them to simple situations.
$n-$Factorial : Denoted by $n!$ and defined by = product of the first n integers

$$\begin{cases} n! = n(n-1)(n-2)...3.2.1 \\ 0! = 1 \quad (\textit{by convention}) \end{cases}$$

2.1.2 Arrangement

Arrangement without repetition.
Among n distinct objects, we choose p distinct objects $(p \leq n)$ by classifying them in a particular order.
Number of Arrangement without repetition :

$$A_n^p = n\,(n-1)\ldots(n-p+1) = \frac{n!}{(n-p)!} \tag{2.1}$$

We must first select the 1^{st} object among n. There are n possibilities. Now, once the $1st$ is chosen, there are only $n-1$ possible choices for the 2^{nd} since the 2^{nd} must differ from the 1^{st}, etc. until we choose the p^{th} object among $(n-p+1)$ remaining possibilities.
Example : Consider 7 distinct objects $a_1, a_2, a_3, a_4, a_5, a_6, a_7$.
A triple like (a_1, a_6, a_3) is an arrangement of these 7 objects.
(a_1, a_3, a_6) is also an arrangement, different from the first.
What is the number of these arrangements?
If we try to form such a triple, we have 7 possible choices for the first object, 6 for the second, and 5 for the third. This gives

$$A_7^3 = 7.6.5 = \frac{7!}{4!}.$$

Arrangement with repetition.
Among n distinct objects, we choose p objects different or not (we can choose the same object several times) by classifying them in a particular order.
Number of Arrangement with repetition :

$$\overline{A_n^p} = n^P$$

Example : Number of passwords containing 5 digits is 10^5, and the number of passwords containing 5 different digits is A_{10}^5
Now, the number of passwords containing 5 different digits and (a) containing 1 and 7 is $A_5^2 \times A_8^3$, (b) containing the pattern 71 is $4 \times A_8^3$.

2.1.3 Permutation

Permutation without repetition.
We classify n distinct objects in a particular order. This is a special case of the previous one (2.1) $(p = n)$.
Number of Permutation without repetition

$$P_n = n!$$

Permutation with repetition.
We classify in a particular order n objects whose n_1 are identical of type 1, n_2 are identical of type 2, ..., n_p identical of type p $(n_1 + n_2 + \cdots + n_p = n)$
Number of Permutation without repetition

$$P(n_1, n_2, \ldots, n_p) = \frac{n!}{n_1! n_2! \ldots n_p!}.$$

2.1.4 Combinaison

Combinaison without repetition.
Among n objects we choose p objects without particular order. We have

$$C_n^p = \frac{n(n-1)\ldots(n-p+1)}{p!} = \frac{n!}{p!(n-p)!}$$

Remark : The combinaison is a subset of cardinal p, and C_n^p is the number of these subsets.
Combinaison with repetition.
We allow the repetition in subset.

$$\overline{C_n^p} = \frac{(n+p-1)!}{p!(n-1)!} = C_{n+p-1}^p$$

Properties :
1. Pascal's triangle : $C_n^0 = C_n^n = 1$, and $\forall n, p,\ 0 \leq p < n$, we have $C_{n-1}^{p-1} + C_{n-1}^p = C_n^p$.
2. Newton's binomial formula : $\forall a, b \in \mathbb{R}$ and $n \in \mathbb{N}$: $(a+b)^n = \sum_{p=0}^n C_n^p a^{n-p} b^p$.

2.2 Probability on a finite set

Historically, the notion of probability emerged from simple examples taken from games of hazard (the word hazard comes from the Arabic az-zahr : the die).
An Event : We will introduce this notion by associating it with an example: the game of die.

Definitions	Examples
A random experiment is an experiment which we cannot predict its result.	The experiment is : roll a die. The result of the experiment is the number indicated on the upper face of the die.
We can then associate to it the set of all possible results or outcomes called the sample space, and denoted by Ω	$\Omega = \{1, 2, 3, 4, 5, 6\}$.
A random event is a subset of Ω.	The event "obtaining an even number" is the subset $A = \{2, 4, 6\}$ of Ω.
We say that the event A is realized if the result of the experiment belong to A.	If the upper face of the die shows 5, A is not realized. If it shows 4, A is realized.
If an event contains only a single element ω, we say that it is an elementary event.	$B = \{1\}$ is one of the 6 elementary events of Ω.

Probabilistic models.
Given an experiment and its sample space Ω, with Ω is finite and $\Omega = \{\omega_1, \ldots, \omega_n\}$.
We define a probability on Ω
– When we associate to each elementary event ω_i a number $p_i = P(\omega_i)$, called the probability of ω_i, such that

$$\begin{cases} \forall i, \quad 0 \leq p_i \leq 1 \text{ and} \\ p_1 + p_2 + \cdots + p_n = 1 \end{cases},$$

– For any event $A \subset \Omega$, we define $P(A)$ as the sum of the probabilities of the elementary events that constitute A.

$$P(A) = \sum_{i/\omega_i \in A} P(\omega_i).$$

Example : The following table gives the probability distribution of a loaded die

ω_i	1	2	3	4	5	6
$P(\omega_i)$	0,20	0,15	0,15	0,10	0,05	0,35

Find the probability that the face 2 or 3 or 6 will show up when you roll the die.
Here $\Omega = \{1, 2, 3, 4, 5, 6\}$ and we want to compute the probability $P(E)$ of the event $E = \{2, 3, 6\}$. So

$$P(E) = P(\{2, 3, 6\}) = P(2) + P(3) + P(6) = 0,15 + 0,15 + 0,35 = 0,65.$$

Example : Consider a radar measuring the speed of a car. We know that this device can make an error of 0.01 km/h, 0.5 km/h, 1 km/h, 2 km/h, or 5 km/h. The probability that these errors occur is given by :

error	0,01	0,5	1	2	5
probability	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

What is the probability of event A: “the error made is less than 1 km/h” ?
$P(A) = P(0.01) + P(0.5) + P(1) = \frac{3}{8} + \frac{1}{4} + \frac{1}{8} = \frac{3}{4}$.
Properties of the probability

Parties of Ω	Vocabulary of events	Property
A	$\forall A \subset \Omega$	$0 \leq P(A) \leq 1$
ϕ	Impossible event	$P(\phi) = 0$
Ω	Certain event	$P(\Omega) = 1$
$A \cap B = \phi$	A and B are incompatible	$P(A \cap B) = P(A) + P(B)$
$A \subset B$	A implies B	$P(A) \leq P(B)$
$\overline{A}$	$\overline{A}$ is the complementary event of A	$P(\overline{A}) = 1 - P(A)$
$A \cup B$	A or B	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Uniform probability.
The following expressions "fair coin",“balanced or perfect die”, “ball drawn from the urn at random”, “indistinguishable balls” ... indicate that, for the experiments realized, the associated model is equiprobability or equally likely.
Definition :
We say that there is equiprobability or P is unifom when all elementary events have the same probability.
Calculations in the case of equiprobability
In a situation of equiprobability, if Ω has n elements then necessary we have $P(\omega_i) = \frac{1}{n}$, and if E is an event composed of m elementary events we have:

$$P(E) = \frac{cardE}{card\Omega} = \frac{\text{number of favorable cases}}{\text{number of possible cases}}.$$

Example : A letter is chosen at random from the letters of the English alphabet. Find the probability that
1. the letter is either I or U,
2. the letter is in the word $ALWAYS$, and
3. the letter is not in the word $NEVER$.
Here, the sample space is $\Omega = \{A, B, C, D; E, \ldots, X, Y, Z\}$. So $|\Omega| = 26$.
1. Let E be the event that the selected letter is I or U.
Then $E = \{I, U\}$.
So, $|E| = 2$
$P(E) = \frac{|E|}{|\Omega|} = \frac{2}{26}$.
2. Let F be the event that letter is in the word ALWAYS.
Then $F = \{A, L, W, Y, S\}$.

So, $|F| = 5$.
$P(F) = \frac{|F|}{|\Omega|} = \frac{5}{26}$.
3. Let G be the event that the letter is not in the word $NEVER$.
Then $G =$ all letters, except N, E, V, R.
So, $|G| = 26 - 4 = 22$.
$P(G) = \frac{|G|}{|\Omega|} = \frac{22}{26}$.

2.3 Conditional probability

Sometimes when new information becomes available, the probability of an event may have to be reevaluated in light of this new information. Suppose we have a sample space Ω and an event A. Now suppose we have new information that an event B has occurred.We will have to reevaluate the conditional probability of A, given the knowledge that B has occurred.
Definition. The conditional probability of an event A, given an event B with $P(B) > 0$, is defined by

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

So, we get the following formula (Multiplication rule)

$$P(A \cap B) = P(A|B)\, P(B).$$

Properties
1. All known properties of probability remain valid for conditional probability.
2. In the case where the possible outcomes are equally likely, we have

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

Example : We toss a fair coin three successive times. We wish to find the conditional probability $P(A|B)$ when A and B are the events

$$A = \{\text{more heads than tails come up}\} \quad B = \{\text{1st toss is a head}\}.$$

The sample space consists of eight sequences,

$$\Omega = \{HHH,\ HHT,\ HTH,\ HTT,\ THH,\ THT,\ TTH,\ TTT\},$$

which we assume to be equally likely. The event B consists of the four elements HHH, HHT, HTH, HTT, so its probability is

$$P(B) = \frac{4}{8}.$$

The event $A \cap B$ consists of the three elements HHH, HHT, HTH, so its probability is

$$P(A \cap B) = \frac{3}{8}.$$

Thus, the conditional probability $P(A|B)$ is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/8}{4/8} = \frac{3}{4}.$$

Because all possible outcomes are equally likely here, we can also compute $P(A|B)$ using a shortcut

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{3}{4}.$$

2.4 Total probability Theorem and Bayes' rule

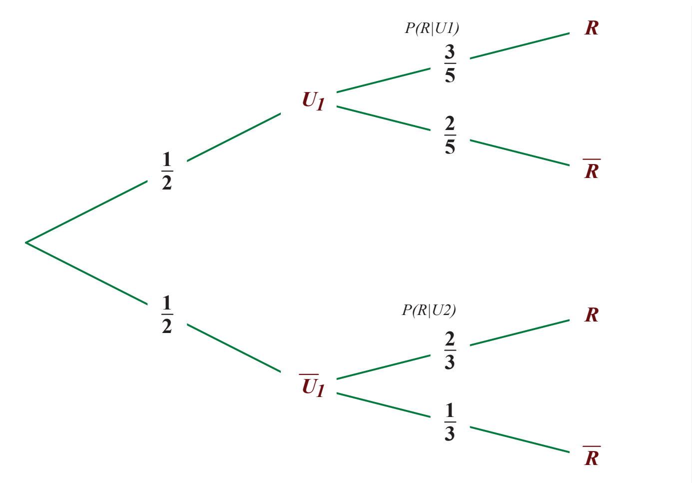
Total Probability Theorem
Let the event A such that $P(A) > 0$. Then, for any event B, we have

$$\begin{aligned} P(B) &= P(A \cap B) + P\left(\overline{A} \cap B\right) \\ &= P(A)P(B|A) + P\left(\overline{A}\right) P\left(B|\overline{A}\right). \end{aligned}$$

Bayes' Rule For any event A and B with $P(A) > 0$, $P(B) > 0$, we have

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

Example : Two indistinguishable urns U_1 and U_2 contain respectively :
– urn U_1 : 3 red balls, 2 green balls;
– urn U_2 : 2 red balls, 1 green ball.
An urn is chosen at random and a ball is drawn from it. What is the probability that it is red?
Let – $U1$ (respectively $U2$): "the ball drawn from urn U_1" (respectively U_2);
– R (respectively V): "the ball drawn is red" (respectively green).



The total probability theorem gives:

$$P(R) = P(R \cap U_1) + P(R \cap U_2)$$
$$P(R \cap U_1) = P(R|U_1)\,P(U_1) + P(R|U_2)\,P(U_2)$$

We deduce that

$$P(R) = \frac{3}{5} \times \frac{1}{2} + \frac{2}{3} \times \frac{1}{2} = \frac{19}{30}.$$

2.5 Independent Events

If the conditional probability $P(A|B) = P(A)$, then we say that A and B are independent.
Note that by the definition this is equivalent to, $P(A \cap B) = P(A)\,P(B)$.
Definition. Two events A and B are independent if

$$P(A \cap B) = P(A)\,P(B).$$

Remark : For 3 or more events $A_1, A_2, \ldots, A_n$, we say they are independent if the "multiplication rule" applies for any number of them. For example, 3 events A, B, C are independent if all of the following holds :

$$P(A \cap B) = P(A)\,P(B)$$
$$P(A \cap C) = P(A)\,P(C)$$
$$P(B \cap C) = P(B)\,P(C)$$
$$P(A \cap B \cap C) = P(A)\,P(B)\,P(C).$$