# 1 Introduction

Probability Theory is a mathematical theory whose aim is to quantify uncertainity. We want to make sense of statements such as

- The probability that tomorrow will rain is 0.75.

- The probability that the the next time that I will toss a coin, the result will be heads is 1/2.

The intuitive meaning is clear: in the first example tomorrow we would bring with us an umbrella, while in the second example we are saying that it is equally probable that the next time that I will toss a coin the result will be heads or tails.
We will come back in this introduction to the meaning of probability, and now we focus on the structure of the above examples. We say that E=" Tomorrow will rain" and F=" The next time that I will toss a coin, the result will be heads" are events. Having defined $E$ and $F$, the above two probability statements can be rewritten as

- The probability of $E$ is 0.75

- The probability of $F$ is 0.5

Using the standard mathematical notation, the two above statements can be rewritten as $\mathbb{P}(E) = 0.75$ and $\mathbb{P}(F) = 0.5$. This is the general structure of probability statements

- The probability of the event $E$ is $p$, where $p$ is a number between 0 and 1, or, in mathematical notation, $\mathbb{P}(E) = p$.

Concretely, the probability of an event $E$ is a number denoted by $\mathbb{P}(E)$ between 0 and 1 attached to each event. It is intuitive that it quantifies the uncertainity we have on events, in the sense that if $\mathbb{P}(E) = 0.6 < \mathbb{P}(F) = 0.8$, we regard the event $F$ more likely than the event $E$. We note that the

probability is not unique and it depends on the individual who is deciding it and on its state of information. As a simple example, the probability that I would assign to the event E="Real Madrid will win the Champions League this year" would be 1/16 if I only knew that there are 16 teams left and between them there is Real Madrid and had no information of any kind about football. Since I sometimes watch football, the probability that I assign to the event $E$ is greater than 1/16, say 1/3. It would decrease if I knew that Benzema is not in shape.
A possible definition of the probability of an event $E$, that is, of the number $\mathbb{P}(E)$ is as the degree of belief an individual( the one assigning the probability) has on the event $E$.
The first notion that we need to precise is the one of event

## 1.1 Definition of an event

For us an event something defined by an unambiguous proposition which can be either true or false (but you are uncertain whether they are true or false). By unambiguous we informally mean that a possible bet or insurance based upon it can be decided without question. Examples of events are: " Nadal will win the next Australian Open ", " Democrats will win the next USA elections ", a given shop this year will sell more products than last year... We note that even if events are stated in an unambiguous way, they are usually unknown to you or you are uncertain about them ( and a probability will be just a quantitative version of this uncertainty). In order to make a more concrete example to which we will come back later, consider a given shop. You can consider the following events "Tomorrow 1 person will enter to the shop", "Tomorrow 1 woman will enter to the shop", " Tomorrow the number of people that will enter will be less or equal than 30 and more or equal than 5"...

## 1.2 Operation and relations on events

From two events $E$ and $F$ you can build the events "$E$ and $F$", "$E$ or $F$" and not $F$. In order to make an example, consider again the example of the shop. Define E=" Tomorrow only 1 person will enter in the shop ", F=" Tomorrow it will rain", G="More than 10 people will enter in the shop tomorrow".

- "E and F"= " Tomorrow only one person will enter in the shop and it will be raining ".

- "E or G"= "Tomorrow either more than 10 persons or only one person will enter in the shop".

- "not F "= " Tomorrow it won't be raining".

Given two events E or F we say that $E$ implies $F$ if, given that $E$ occurs, then $F$ occurs. Note that it can very well happen that neither $E$ implies $F$ nor $F$ implies $E$. As an example, let $E$, $F$ and $G$ be defined as above. Then $E$ implies "$notG$", $E$ does not imply $F$ and $F$ does not imply $G$.

**Exercise 1:**
Convince yourself about the previous statements

## 1.3 Elementary Events and Sample Space

When we want to do a probabilistic analysis, we don't want to take into account every event. To be more concrete, if we are interested in the result of a coin toss we are interested in the event E="The result of the coin toss is tails" but not on the event "The result of the coin toss is tails and tomorrow it will rain in Rome".
Coming back to the example of the shop, you could be interested in predicting the probability of the number of people that will enter in the shop tomorrow. In this case you can consider events of the form "The total number of people that will enter in the shop tomorrow is more or equal than 10 and less or equal than 100", " Tomorrow 20 people will enter in the shop", ... But you could also want to make a finer analysis by distinguishing the gender of the people entering in the shop tomorrow. In this case you want to to consider events of the form "The number of women entering in the shop tomorrow will be less than 20 and the man entering in the shop tomorrow will be more than 30", and so on.

**Definition 1.1** *A choice of events (on which we wish to make a probabilistic analysis) is called (in this notes) a system.*

Once that you fixed a system, that is, once you have chosen the events you want to consider, you can distinguish a special kind of events, the *elementary events*. Those are the events that you cannot subdivide more. By subdivide an event $E$ into $F$ and $G$ we mean that $E = F$ or $G$.

**Definition 1.2** *Fix the events you want to consider. The elementary events are those events which cannot be subdivided into other events you want to consider. More precisely, the event $E$ is elementary if $E$ cannot be written as $E =$ "A or B", for $A$ and $B$ events that you want to consider.*

For the sake of concreteness, let's come back to the example of the shop. Imagine that you only want to make a probabilistic analysis of the number of people entering in the shop tomorrow, disregarding the gender. We have already seen the events that we want to consider in this case. The event E="The total number of people that will enter in the shop tomorrow is more or equal than 10 and less or equal than 100 " can be subdivided, for instance, into E="F or G", where F="The total number of people that will enter in the shop tomorrow is more or equal than 10 and less or equal than 50", and G="The total number of people that will enter in the shop tomorrow is strictly more than 50 and less or equal than 100", so that it is not elementary. On the contrary, the event H=" Tomorrow precisely 30 people will enter in the shop " cannot be subdivided into other events that we want to take into account. We don't want to consider the event "Tomorrow exactly 30 people will enter in the shop and it will snow". H is an elementary event. If instead we take into account the gender of the people entering, $H$ could be decomposed in smaller events such as L=" Tomorrow exactly 20 man and 10 woman will enter in the shop". In this new system, $L$ will be an elementary event.

**Example 1.3 (Two dice rolled)** *Two dice have been rolled, and we don't know their result. First assume that you can only look at the sum of the results. Then events are of the form*

- *"The sum is an odd number", "The sum is greater than 8",...*

*and the elementary events are*

- *"The sum is 2"*
- *" The sum is 3"*

  ⋮
- *"The sum is 12",*

*For example, we don't want to distinguish inside the event " The sum is 6 and Nadal will win the next tennis tournament".*

*Assume now that we can look at the result of each single die. Now the previosly elementary event " The sum is 3" can be decomposed into the elementary events " The first die gave 1 and the second 2 " and the event "The first die gave 2 and the second 1". In this case the elementary events are*

- *" The first die gave 1 and the second 1"*
- *" The first die gave 1 and the second 2"*

  ⋮
- *" The first die gave 1 and the second 6"*
- *" The first die gave 2 and the second 1 "*

  ⋮
- *" The first die gave 6 and the second 6"*

Elementary events are also called outcomes of an experiment. In the shop example, you can think that you are counting the number of people entering through some kind of detector placed above the entrance. The number give to you by the detector is the elementary event. With this interpretation, a system is also called an experiment.

We end this section with an important definition

**Definition 1.4 (Sample Space)** *Assume you want to do a probabilistic analysis of a system, so that you have chosen the events of which you want to calculate the probability and consequently you have defined the elementary events. The set of elementary events is denoted by* $\Omega$ *and is called the sample space.*

If you like to think at elementary events as the outcomes of the experiments and to probability systems as experiments, then the sample space $\Omega$ can be thought as the set of outcomes of the experiment

As we will see, in the text of the exercises it is the sample space which is usually given to you. In the previous example we have recognized the elementary events, so that we just need to place them together to obtain the sample space:

**Example 1.5 (Two dice have been rolled)** *If we are looking only at the sum of the two dice, the sample space is*

$$\Omega = \{\text{"The sum is 2"}, \text{"The sum is 3"}, ..., \text{The sum is 12"}\},$$

*while if we are looking at the result of each single dice*

$$\Omega = \{ \text{"The first die gave 1 and the second 1"}, \text{"The first die gave 1 and the second 2"}, \\ ....\text{"The first die gave 6 and the second 6"}\}$$

**Example 1.6 (A chess match)** *Assume that you are observing a chess match. A full description of the game is given by describing the moves of each pleyer untill the king is killed. Usually this description is given using what is called algebraic notation. An example of a match is given by (It is a Fisher vs. Kasparov match)*
*1.d4Nf6 2.c4e6 3.Nc3Bb4 4.Nf3c5 5.e3Nc6 6.Bd3Bxc3+ 7.bxc3d6 8.e4e5 9.d5Ne7 10.Nh4h6 11.f4Ng6 12.Nxg6fxg6 13.fxe5dxe5 14.Be3b6 15.O-OO-O 16.a4a5 17.Rb1Bd7 18.Rb2Rb8 19.Rbf2Qe7 20.Bc2g5 21.Bd2Qe8 22.Be1Qg6 23.Qd3Nh5 24.Rxf8+Rxf8 25.Rxf8+Kxf8 26.Bd1Nf4 27.Qc2Bxa40–1*
*The above is an elementary event of the system "A chess match between Player1 and Player 2". A game easier than chess will be the object of an exercise at the end of this lecture. Note that describing the event "The white will win" in terms of elementary events is a task too difficult to be made.*

## 1.4 Exercises

**Ex. 1** — From the 2019-2020 course " Mathematics and Statistics" given by professors Khovanskaya, Bubilin, Shurov, Filimonov and Sonin at HSE. A six faced die is rolled. Find the elementary events that compose the event

1. "The result is 6"
2. "The result is a number less or equal 2"
3. "The result is an even number"
4. "The result is a number strictly greater than 4"
5. "The result is seven"

**Answer (Ex. 1)** —

1. It is already an elementary event
2. "The result is a number less or equal than 2"= "The result is 1" or "The result is 2".
3. "The result is a number strictly greater than 4" = "The result is 5" or "The result is 6 ".
4. The event is the impossible event, also denoted by $\emptyset$.

**Ex. 2** — A coin is tossed twice. We are interested in the face which the coin shows when it lands: heads or tails. The set of elementary events is $\{HH, HT, TH, TT\}$, where the event HH corresponds to "the first toss gave heads, the second heads", the event HT corresponds... Write the elementary events that compose the events

1. "We obtained two heads".
2. "The first toss gave heads".
3. "We obtained 1 tails "
4. "At least one toss gave tails"

**Answer (Ex. 2)** —

1. $HH$.
2. $TH, TT$.
3. $TH, HT$.
4. $HT, TH, TT$.

**Ex. 3** — A coin is tossed 4 times. We are interested in the face that it shows when it lands: heads or tails. How many elementary events are there? which elementary events are contained in the events

1. The first result was heads
2. The second result was tails
3. The first result was heads and the second tails
4. All the 4 tosses gave the same result

**Answer (Ex. 3)** — The elementary events are HHHH, HHHT,HHTH,....,TTTT. There are $2^4 = 16$ of them.

1. HHHH, HHHT,....,HTTT
2. HTHH,HTHT,HTTH,...,TTTT
3. HTHH, HTHT,HTTH,HTTT
4. HHHH, TTTT

**Ex. 4** — ([**Ross**] Chapter 2 Exercise 1) A box contains 3 marbles: 1 red, 1 green, 1 blue. Consider the experiment that consists in taking 1 marble from the box and then replacing it in the box and then drawing a second marble. Describe the sample space.

**Answer (Ex. 4)** — The sample space can be thought as the set of the outcomes of the experiment. In this case the possible outcomes are "The first marble is red and the second marble is red", "The first marble is red and the second is green",... A more efficient way to write an outcome is to write an array with two entries, i.e, (r,g), where the first entry represents the colour of the firs marble, and the second entry represents the colour of the second marble. That is, we write the elementary event "The first marble is red and the second is blue" by (r,b). The sample space becomes

$$\Omega = \{(r,r),(r,g),(r,b),(g,r),(g,g),(g,b),(b,r),(b,g),(b,b)\} \quad \boxed{\texttt{ex1}}$$

**Ex. 5** — Repeat Exercise 1 when the second marble is drawn without replacing the first marble.

**Answer (Ex. 5)** — The outcomes (r,r), (g,g) and (b,b) are no longer possible. The sample space is

$$\Omega = \{(r,g),(r,b),(g,r),(g,b),(b,r),(b,g)\} \quad \boxed{\texttt{ex2}}$$

**Ex. 6** — You toss a coin 3 times. Describe the elementary events.

**Answer (Ex. 6)** — $\Omega = \{(H,H,H),(H,H,T),....,(T,T,T)\}$, where the event (H,T,H) is "The result of the first toss was head, the result of the second...".

## 1.5 A graphical way to think about the sample space

A way to visualize the sample space $\Omega = \{\omega_1, ..., \omega_n\}$ is by drawing a portion of area in the plane as $\Omega$, and subdivide it in $n$ disjoint regions such as in the picture.

Suppose that now we want to make a finer analysis in our system, and that now each $\omega_i$ is further divided into the new elementary events $\omega_i'$. By performing this division on the previous image we obtain the new sample space division of the sample space.

One can think at the point $(x,y)$ of the plane where $\Omega$ is pictured as being limiting points that you can obtain from elementary events after making further and further refinements.

## 1.6 Examples

The aim of the next examples is to introduce set once for all a convenient notation for the examples that we will be treating during the course.

**Example 1.7 (A dice has been rolled)** *The elementary events are " The result is 1 ",..." The result is 6". We use the shortcut 1=" The result is 1" ,...., 6=" The result is 6". The sample spase becomes* $\Omega = \{1,2,3,4,5,6\}$. *Note that you don't have to consider 1,2,3,4,5,6 as numbers, but as events. Concretely 1+1=2 makes no sense since "The result is 1"+ "The result is 2 " makes no sense.*

**Example 1.8 (Two dice rolled)** *The sample space is* $\Omega = \{(1,1),(1,2),...,(1,6),(2,1),...(6,6)\}$ $\{(i,j)\,,\, i,j = 1,...,6\}$, *where the* $(i,j)$ *is a shortcut for the elementary event "the first die gave i and the second j". The event "The sum of the results is 5" corresponds to the subset of* $\Omega$ $E = \{(1,4),(2,3),(3,2),(4,1)\} = \{(i,j) \in \Omega\,,\, i+j = 8\}$.

**Example 1.9 (A coin toss)** *In a coin toss there are only two elementary events: 0=" The result is tails" and 1 = " The result is heads".* $\Omega_1 = \{0, 1\}$. *We will often use the word succes instead of the word heads.*

**Example 1.10 (two, three and $n$ coin tosses)** *Assume that we are looking at two coin tosses. Then* $\Omega_2 = \{00, 01, 10, 11\} = \{x_1x_2\,,\, x_i \in \{0, 1\}i = 1, 2\}$*, where, for example, 01="The first toss gave tails and the second tails". For three tosses,* $\Omega_3 = \{000, 001, 010, 011, 100, 101, 110, 111\} = \{x_1x_2x_3\,,\, x_i \in \{0, 1\}i = 1, 2, 3\}$*. The event E= "There has been exactly one succes" is* $E = \{001, 010, 100\} = \{x_1x_2x_3\,,\, x_1 + x_2 + x_3 = 1\}$.
*In general, the sample space of n coin tosses is* $\Omega_n = \{x_1x_2...x_n\,,\, x_i \in \{0, 1\}i = 1, ..., n\}$*, and the event* $E^n_k$*= " There have been k successes" is* $E^n_k = \{x_1...x_n \in \Omega_n\,,\, x_1 + ... + x_n = k\}$.

**Example 1.11 (finite but undetermined amount of coin tosses)** *Suppose that we are observing some coin tosses but we are uncertain also on how many of them there will be. The sample space is* $\Omega = \bigcup_{i=1}^{\infty} \Omega_i$ *where the in* $\Omega_i$ *have been defined previously.*

**Example 1.12 (Infinite coin tosses)** *A non-trivial generalization of n coin tosses is the one of an infinite amount of coin tosses. In this case the sample space is* $\tilde{\Omega} = \{x_1...x_n...\,,\, x_i \in \{0, 1\}\}$*. In this case you know that the amount of times that you flip the coin is infinite.*

**Example 1.13 (Sampling)** *Suppose that a sample of 100 people is taken in order to estimate how many people smoke. If you are only interested in the number of smokers, disregarding age, gender, etc., the sample space is* $\Omega = \{0, 1..., 100\}$*, where, for instance,* 1 *means the event "There is exactly 1 smoker". The event "The majority of people sampled are smokers" is composed by the simple events 51,52,...100.*

## 1.7 Events as subsets of the sample space

Up to now what we said was an operative way to define your sample space in terms of the events of which you are interested in doing a probabilistic analysis. However, the mathematical theory needs only the sample space as a primitive notion. The other notions are based on $\Omega$, included the one of event. Indeed the previous examples show that an event can be represented as the set of elementary events that compose it, that is, an event is a subset of the sample space. When we roll a dice, the event "The result is odd" can be written as $\{1, 3, 5\}$.

**Definition 1.14** *Let $\Omega$ be the sample space. An $E$ subset of $\Omega$ is said to be an event. We also consider the impossible event $\emptyset$ as an event.*

For example, the sample space of a three faced die is $\Omega = \{1, 2, 3\}$, and the events are $\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$. The event $\{1, 3\}$ can be written as a proposition as "The result is an odd number".

Given a set, the set of his subsets is called the power set and is denoted. Using this notion, we could say that the set ov events is the power set of the sample space. A nice exercise is to prove that if $\Omega$ has $n$ elements, the power set has $2^n$ elements.

To every proposition we associate a subset of the sample space, as we did with the event "The resut is an odd number", but the analogy does not end here. The operations "or", "and" " not", can be traduced in terms of $\cup$, $\cap$ and $^c$. Consider the sample space $\Omega_3 = \{000, 001, ..., 111\}$, and consider the events E="The number of successes? is less than one" and $F$= " The number of successes is more than 1". Then $E = \{000, 001, 010, 100\}$ and $F = \{001, 010, 100, 011, 101, 110, 111\}$, and the event G=" E and F"=" The number of successes is exactly 1" is simply given by $G = E \cap F = \{000, 001, 010, 100\} \cap \{001, 010, 100, 011, 101, 110, 111\} = \{001, 010, 100\}$. Note also that K=" not E"=" The number of successes is strictly more than 1", and that, indeed, $K = \{011, 101, 110, 111\} = \{000, 001, 010, 100\}^c = \Omega_3 \setminus \{000, 001, 010, 100\}$.

**Ex. 7** — ([**Feller**], Chapter 1.8) Let $\Omega$ be a sample space and $A, B, C$ be three arbitrary events. Find the expressions for the events that of $A, B, C$:

1. Only $A$ occurs.
2. Both $A$ and $B$, but not $C$ occur.
3. All three events occur.
4. At least one occurs.
5. At least two occur.
6. One and no more occurs.
7. Two and no more occur
8. None occurs.
9. Not more than two occur.

**Answer (Ex. 7)** —

1. $A \setminus (B \cup C) = A \cap B^c \cap C^c$.
2. $(A \cap B) \setminus C = A \cap B \cap C^c$.
3. $A \cap B \cap C$.
4. $A \cup B \cup C$.
5. ex:set1 At least two occur if occur one bertween $A \cap B$, $A \cap C$ or $B \cap C$. Thus $(A \cap B) \cup (A \cap C) \cup (B \cap C)$.
6. It has to happen one between $A$ $B$ and $C$, that is, it has to happen $A \cup B \cup C$, but it does not have to happen two of them, that is, it cannot happen the event described in 5:
7. $(A \cap B) \cup (A \cap C) \cup (B \cap C)$. Thus the event is $A \cup B \cup C \setminus ((A \cap B) \cup (A \cap C) \cup (B \cap C))$.
8. $((A \cap B) \cup (A \cap C) \cup (B \cap C)) \setminus A \cap B \cap C$ which is the same as $((A \cap B) \cup (A \cap C) \cup (B \cap C)) \cap (A^c \cup B^c \cup C^c)$ .
9. $(A \cup B \cup C)^c = A^c \cap B^c \cap C^c$.
10. This event happens when the event all three events occur does not happen. Thus the event is $(A \cap B \cap C)^c = A^c \cup B^c \cup C^c$.

# 2 Exercises

**Ex. 8** — ([**Ross**] Chapter 2 Exercise 2) In an experiment, die is rolled continually until a 6 appears, at which point the experiment stops. What is the sample space of this experiment? Let $E_n$ denote the event " The experiment is completed before n rolls". Write it as a subset of the sample space. Describe $\left(\bigcup_{i=1}^{\infty} E_i\right)^c$.

**Answer (Ex. 8)** — The sample space is the space of outputs of the experiment. We encode the event "The first roll gave 4, the second 2, the third 6 and then stops" in the string 426. That is, we encode the output of the experiment in strings of digits 1,2...,5,6, where the digit correspond to the result of the die, and the position of the digit corresponds to when has the digit appeared. The string must stop at after the first 6, since the game stops when the die gives 6.
We have to keep in mind that we can observe events in which we never stop rolling the die, which correspond to infinite strings of 1,2...5, for example

1212121.... or 121233...34453236....

$\Omega = \{6, 16, 26, 36, 46, 56, 116, ...\} \cup \{x_1 x_2 .... x_n ... : x_i \in \{1, 2, 3, 4, 5\}\} = E \cup G.$

The set $E$ is the event that the game stops in a finite number of rolls. The set $G$ is the event that the game goes on without ending. For any reasonable probability, the event G has probability 0, and it is usually neglected. The events $E_i$ can be written as subsets of the state space in the following way: $E_1 = \{6\}$, $E_2 = \{6, 16, 26..., 56\}$, $E_3 = E_2 \cup \{116, 126, ..., 556\}$,... (In a more compact form $E_i = \{x_1...x_n 6 : x_i \in \{1, ..., 5\}\, n \leq i\}$).
The event

$$\bigcup_{i=1}^{\infty} E_n = E_0 \cup E_1 \cup E_2, .... = "E_0 \; or E_1 \; or .... E_n \; \; or ..."$$

is, by definition of "or", the event that $E_n$ happens for at least one $n$. That is, is the event "The game ends before $n$ rolls at least one $n$ "="The game ends after $n$ rolls for some $n$", which coincides with $E$. The set event $E^c$ corresponds to $G$.

**Ex. 9** — ([**Ross**] Chapter 2 Exercise 3) Two dice are thrown. Let $E$ be the event that the sum of the dice is odd, and $F$ the event that at least one of the two dice lands in 1 and let $G$ be the event that the sum is 5. Describe the events $E \cap F$, $E \cup F$, $F \cap G$, $E \cap F^c = E \setminus F$ and $E \cap F \cap G$

**Answer (Ex. 9)** —

$$E = \{(1, 2), (1, 4), (1, 6), (2, 1), (2, 3), (2, 5), (3, 2)...(6, 5)\}$$

$$F = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1)...(6, 1)\}$$

$$G = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$$

$$E \cup F = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), (2, 3), (2, 5), (3, 1), (3, 2), (3, 4)..(6, 5)\}$$

$$E \cap F = \{(1, 2), (1, 4), (1, 6), (2, 1), (4, 1), (6, 1)\}$$

$$F \cap G = \{(1, 4), (4, 1)\}$$

$$E \cap F \cap G = F \cap G = \{(1, 4), (4, 1)\}$$

$$E \cap F^c = \{((2, 3), (2, 5), (3, 2), (3, 4), (3, 6), (4, 3), ...(6, 5)\}$$

**Ex. 10** — ([**Ross**] Chapter 2 Exercise 4) Anton, Barbara and Carlo take turns flipping a coin. The first one to get head wins. The sample space to this experiment can be defined as $\Omega = \{1, 01, 001, 0001, ...\} \cup \{00000...\}$.

1. Interpret the sample space.
2. In terms of the sample space, write the following events: A="Anton wins", B="Barbara wins" and $(A \cap B)^c$.

**Answer (Ex. 10)** —

1. 1="First toss is heads"(So that Anton wins) 01="The first toss is tails, the second heads "(So that Barbara wins)... The event 00000...="Head never arises, the game never stops".
2. As it is easy to seen, Anton wins whenever the digit 1 appears in the 1st, 4th, 7th.. position. This means that

$$A = \{1, 0001, 0000001, ...\}.$$

In the same way, the event that Barbara wins corresponds is

$$B = \{01, 00001, 00000001, ...\}$$

where the digit appears in the after $3 \times k + 1$ zeros, for some $k \in \mathbb{N}$. $(A \cup B)^c = \Omega \setminus (A \cup B) = \{001, 000001, ...\} \cup \{0000...\}$. The event $(A \cup B)^c$ is the event "Carlo wins or the game doesn't end"

# 3 Probability

Once we choose the events of which we are uncertain about and interest us, we want to quantify this uncertainty by introducing the probability. This is done by attaching to each event $E$ a number $\mathbb{P}(E) \in [0, 1]$, denoted as the probability of the event $E$.

Note that when considering two (or more) events, say $E$ and $F$, the numbers $\mathbb{P}(E)$ and $\mathbb{P}(F)$ are not arbitrary and have to satisfy some relations that ensure that the probability behaves as we expect to behave. Think about the events E=" 5 Star Movement will win next Italy elections "and F=" 5 Star Movement will win next Italy elections and on January 12 2023 it will rain". Then the numbers $\mathbb{P}(E)$ and $\mathbb{P}(F)$ are numbers in $[0, 1]$, but they also need to satisfy $\mathbb{P}(E) \geq \mathbb{P}(F)$, since if $F$ happens then $E$ happens. The relations that $\mathbb{P}(E)$ have to satisy when calculated on different events (

so that $\mathbb{P}$ can be properly called a probability) are the *Kolmogorov Axioms.* We restrict our attention to finite sample spaces $\Omega = \{\omega_1, ..., \omega_n\}$. In this case we don't need the full generality of the Kolmogorov axioms, and we the probability of an event $E \subset \Omega$ can be restated in terms of elementary events. Imagine that we roll a *balanced* die, so that the state space is $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\mathbb{P}(1) = ... = \mathbb{P}(6) = 1/6$. What is the probability of $\{1, 3, 5\}$= "The result is an odd number? The probability of $\{1, 3, 5\}$ is the sum of the probabilities of the elementary events that compose it:

$$\mathbb{P}(\{1, 3, 5\}) = \mathbb{P}(1) + \mathbb{P}(3) + \mathbb{P}(5) = 1/6 + 1/6 + 1/6 = 1/2.$$

**Example 3.1 (Unbalanced die)** *You roll an unbalanced die whose probabilities are given by $\mathbb{P}(1) = \mathbb{P}(4) = 1/8$, $\mathbb{P}(2) = \mathbb{P}(3) = 1/3$ $\mathbb{P}(6) = 0$ and $\mathbb{P}$ you buy exactly one, so that the sample space is $\Omega = \{1, 2, 3, ...10\}$, where 1= "You have bought the copon number 1",.... Assume that the probability of having bought the i-th coupon are gven by*

$$\mathbb{P}$$

*The sample space is $\Omega = \{1, 2, 3..., 1000\}$ (we are assuming that no more that 1000 people will enter in that day). Assume that the probabilities of the elementary events are given by $\mathbb{P}(i)$*

## 3.1 Definition on a finite Sample Space

Let $\Omega = \{\omega_1, ..., \omega_n\}$ be a finite sample space. A probability on the elementary events is a set of $n$ numbers $\mathbb{P}(\omega_1)$, $\mathbb{P}(\omega_2)$,...,$\mathbb{P}(\omega_n)$ satisfying

$$\mathbb{P}(\omega_i) \in [0, 1] \text{ for every } i = 1, ..., n \qquad \boxed{\texttt{e:pos}}$$

and satisfying the followinf *fundamental assumption*

$$\mathbb{P}(\omega_1) + ... + \mathbb{P}(\omega_n) = 1. \qquad \boxed{\texttt{e:fun}}$$

The meaning that you can give to the number $\mathbb{P}(\omega_i)$ is thedegree of belief that you have on the event $\omega_i$, but once you have those numbers, probability theory tells you what you can do with those numbers disregarding their interpretation.

**Example 3.2 (balanced dice)** *Assume that we throw a balanced dice. Then the sample space is* $\Omega = \{1, 2, 3, 4, 5, 6\}$*, and the hypotheses of the dice being balanced means that there is not favourable outcome:* $\mathbb{P}(1) = 1/6$, $\mathbb{P}(2) = 1/6$, $\mathbb{P}(3) = 1/6$, $\mathbb{P}(4) = 1/6$, $\mathbb{P}(5) = 1/6$, $\mathbb{P}(6) = 1/6$.

**Example 3.3 (unbalanced dice)** *Assume you throw a dice, so that the sample space is* $\Omega = \{1, 2, 3, 4, 5, 6\}$*. An possible probability for an unbalanced dice would be* $\mathbb{P}(1) = 1/6$, $\mathbb{P}(1) = 1/6$, $\mathbb{P}(3) = 1/12$, $\mathbb{P}(4) = 3/12$, $\mathbb{P}(5) = 1/6$, $\mathbb{P}(6) = 1/6$,

**Definition 3.4** *Let* $\Omega = \{\omega_1, ..., \omega_n\}$ *be the sample space, and assume that the probabilities of the elementary events* $\mathbb{P}(\omega_1)$*,...,*$\mathbb{P}(\omega_n)$ *have been assigned. The prbability of an event* $E$ *is the sum of the probabilities of the elementary events that compose it:*

$$\mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\omega). \tag{1}$$

d:prob

*In a less obscure way, if* $E = \{\omega_1, ..., \omega_k\}$*, then* $\mathbb{P}(E) = \mathbb{P}(\omega_1) + .... + \mathbb{P}(\omega_k)$.

**Example 3.5 (The dice)** *Let's calculate the probability of the event "The result is an odd number " with the two different probabilities chosen above. The event is* $E = \{1, 3, 5\}$*, and, in the balanced case, its probability is*

$$\mathbb{P}(E) = \mathbb{P}(1) + \mathbb{P}(3) + \mathbb{P}(5) = 1/6 + 1/6 + 1/6 = 1/2.$$

*For the unbalanced case the probability is*

$$\mathbb{P}(E) = \mathbb{P}(1) + \mathbb{P}(3) + \mathbb{P}(5) = 1/6 + 1/12 + 1/6 = 5/12.$$

Observe that the fundamental assumption is equivalent ot $\mathbb{P}(\Omega) = 1$, that is, that your experiment will have an outcome for sure. Of course, the probability of the impossible event $\emptyset$ is zero:$\mathbb{P}(\emptyset) = 0$.

## 3.2 Uniform probablity

When you have no reason to think that one of the outcomes, or elements of the sample space is more likely than the other, then each $\mathbb{P}(\omega_i)$ should be equal, say to $p \in [0,1]$ for each $i = 1, ..., n$. In this case the fundamental assumption tells us that

$$\mathbb{P}(\omega_1) + ... + \mathbb{P}(\omega_n) = p + .... + p = np = 1,$$

that is, $p = 1/n$. Therefore each elementary event has the same probability and this probability is $1/n$, where we recall that $n$ is the number of elements of $\Omega$.

Let's consider the event $E = \{\omega_1, ..., \omega_k\}$ (or more generally $E = \{\omega_{i_1}, ..., \omega_{i_k}\}$). The definition (1) tells us that $\mathbb{P}(E) = 1/n + .... + 1/n = k/n$. This can be written in the suggestive form

$$P(E) = \frac{\text{number of elements of E}}{\text{number of elements of } \Omega}.$$

In this case calculate the probability reduces to counting. In many cases this can be almost an impossible task. You can interpret computing the probability as counting the elements $\omega_i$ with a weight $\mathbb{P}(\omega_i)$, where the usual way of counting (each element weights the same) is obtained when considering the uniform probability

**Example 3.6 (Extractions with replacement )**

**Example 3.7 (Extraction without replaacement)**