

# Probability

Marco Biroli

September 7, 2020



# Contents

<b>1</b>	<b>Founding Blocks</b>	<b>5</b>
1.1	Definitions . . . . .	5
<b>2</b>	<b>Coin tossing games.</b>	<b>9</b>
2.1	The model . . . . .	9
2.2	Graphical Representation. . . . .	9
2.3	Interpretation of the model. . . . .	9
2.3.1	Coin tossing game . . . . .	9
2.3.2	Random Walk . . . . .	9
2.4	Distribution or law of $S_n$ . . . . .	9
2.6	Equalization or return to 0 . . . . .	10
2.7	The lamplighter walk. . . . .	10
2.8	Reflection principle. . . . .	10
2.9	The ballot theorem. . . . .	11
2.10	End of the computation. . . . .	11
2.11	Fundamental Lemma . . . . .	11
2.12	Last tie . . . . .	11



# Chapter 1

## Founding Blocks

### 1.1 Definitions

**Definition 1.1.1** (Universe). We consider a random experiment, then the set of all possible outcomes of the experiment is denoted by  $\Omega$  and is called the universe.

**Definition 1.1.2** (Event). An event is usually denoted by  $E$ . An event is a set of results for which we can compute the probability.

**Definition 1.1.3** (Collection). The collection of all events is denoted by  $\mathcal{F}$ . Hence  $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ .

**Definition 1.1.4** (Disjoint Events). Two events  $A, B \in \mathcal{F}$  are disjoint or incompatible if they cannot occur simultaneously. In other words if  $A \cap B = \emptyset$ .

**Remark.** We require that the collection  $\mathcal{F}$  of the events is an algebra of sets.

**Definition 1.1.5** (Algebra of Sets). An element  $\mathcal{F}$  is called an algebra of sets if  $\mathcal{F} \neq \emptyset$  and:

1.  $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
2.  $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}$

**Remark.** For the scope of this course we further require that  $\mathcal{F}$  is stable under countable unions. In other words the second condition above is replaced by:

$$(A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$$

**Definition 1.1.6** ( $\sigma$ -algebra). A  $\sigma$ -algebra is an algebra of sets where the second condition is replaced by the stronger condition requiring stability under countable union.

**Definition 1.1.7** (Probability). The probability  $P(E)$  of  $E$  is the theoretical value for the proportion of experiments in which  $E$  occurs. Thus the probability is a function from  $\mathcal{F}$  to  $[0, 1]$ . Such that:

1.  $P(\Omega) = 1$ .
2.  $A, B \in \mathcal{F}, A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ .

In other words,  $P$  is an additive set function from  $\mathcal{F}$  to  $[0, 1]$ .

**Remark.** This definition however is not very well suited to infinite event sets. Then modern probability theory adds a condition to the above.

**Definition 1.1.8** (Modern Probability). A modern probability  $P(E)$  of  $E$  is a probability with the stronger condition:

$$\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{F}^{\mathbb{N}}, (\forall n, m \in \mathbb{N}, n \neq m \Rightarrow A_n \cap A_m = \emptyset) \Rightarrow P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n)$$

**Definition 1.1.9** (Probability Space). A probability space is a triple  $(\Omega, \mathcal{F}, P)$ . Where  $\Omega$  is the universe of all possible results,  $\mathcal{F}$  is a  $\sigma$ -field on  $\Omega$ , and  $P$  is a modern probability function on  $\mathcal{F}$ .

**Remark.** The mathematical framework which defines probability theory actually comes from another mathematical framework called measure theory. This is why the elements of the  $\sigma$ -field are sometimes called the measurable sets and the probability function is sometimes called a probability measure.

**Definition 1.1.10** (Finite Space). We consider the case where  $\Omega$  is a finite set, we write  $\Omega = \{x_1, \dots, x_n\}$ . The natural  $\sigma$ -field on  $\Omega$  is  $\mathcal{P}(\Omega)$ . It is the only  $\sigma$ -field which contains the singletons. Then let  $P$  be a probability on  $\Omega$  and let us set  $\forall i \in \llbracket 1, n \rrbracket, p_i = P(\{x_i\})$ . Then the numbers  $p_i$  satisfy:

$$(\forall i \in \llbracket 1, n \rrbracket, 0 \leq p_i \leq 1) \wedge \sum_{i=1}^n p_i = 1$$

Then for any  $A \subset \Omega$  we have by additivity that:

$$P(A) = \sum_{x \in A} P(\{x\}) = \sum_{i: x_i \in A} p_i$$

Hence  $P$  is completely determined by the numbers  $p_i$ .

**Remark.** Notice that conversely if we are given the numbers  $p_i$  summing to 1 we can define a probability  $P$  on  $\Omega$  by stating  $P(\{x_i\}) = p_i$  and  $P$  will indeed be a probability measure.

**Definition 1.1.11** (Countable Spaces). We suppose that  $\Omega$  is countable and we set  $\Omega = \{x_n, n \in \mathbb{N}\}$ . The natural  $\sigma$ -field on  $\Omega$  is again the power set of  $\Omega$ . Then the definitions are an immediate generalization of the ones for a finite space.

**Definition 1.1.12** (Continuous Spaces). If we take the simplest example of  $\Omega = \mathbb{R}$  then the intuitive  $\sigma$ -field being the power set turns out to be too complicated to be useful. Hence we take for  $\mathcal{F}$  the Borel tribe of  $\mathbb{R}$ ,  $\mathcal{B}(\mathbb{R})$ . The Borel  $\sigma$ -field corresponds to taking a countable union of all possible closed intervals of  $\mathbb{R}$ .

**Definition 1.1.13** (Random Variable). Let  $(\Omega, \mathcal{F}, P)$  be a probability space. A random variable  $X$  on  $(\Omega, \mathcal{F}, P)$  is map from  $\Omega$  to  $\mathbb{R}$ . Which satisfies:

$$\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

**Remark.** This definition is equivalent to: for all interval  $I$  of  $\mathbb{R}$  we have that  $X^{-1}(I) \in \mathcal{F}$ .

**Notation.** The event  $X^{-1}(I)$  is denoted by  $\{X \in I\}$  or even simply  $X \in I$ . Secondly random variables are denoted by capital letters typically  $X, Y, U, V$  and their possible values are denoted by the corresponding lowercase letters.

**Definition 1.1.14** (Law of a random variable). Let  $(\Omega, \mathcal{F}, P)$  be a probability space and let  $X$  be a random variable defined on  $\Omega \rightarrow \mathbb{R}$ . The law of  $X$  is the probability measure on  $\mathbb{R}$  defined by:

$$\forall B \in \mathcal{B}(\mathbb{R}), P_X(B) = P(X \in B)$$

*Proof.* Let us check that  $P_X$  is indeed a probability measure. We have that:

$$P_X(\mathbb{R}) = P(X \in \mathbb{R}) = 1.$$

Furthermore let  $(B_n)_{n \in \mathbb{N}} \in \mathcal{B}(\mathbb{R})^{\mathbb{N}}$  be a disjoint sequence of Borel sets. Then:

$$P_X\left(\bigcup_{n \in \mathbb{N}} B_n\right) = P\left(X \in \bigcup_{n \in \mathbb{N}} B_n\right) = P\left(\bigcup_{n \in \mathbb{N}} \{X \in B_n\}\right) = \sum_{n \in \mathbb{N}} P(X \in B_n) = \sum_{n \in \mathbb{N}} P_X(B_n)$$

□

**Notation.** The law  $P_X$  of  $X$  is sometimes called the distribution of  $X$ . We furthermore say that two variables  $X, Y$  have the same law if  $P_X = P_Y$ . The object of primary interest for a random variable is its law.

**Definition 1.1.15** (Law). Let  $f$  be a non-negative function  $\mathbb{R} \rightarrow \mathbb{R}^+$  which is integrable and  $\int_{\mathbb{R}} f(x)dx = 1$ . We define next:

$$\forall A \in \mathcal{B}(\mathbb{R}) \quad P(A) = \int_A f(x)dx$$

This formula defines a probability measure on  $\mathbb{R}$ , called the probability measure with density function  $f$ .

**Definition 1.1.16** (Expectation). We say that the random variable  $X$  has an expectation or that it is integrable if:

$$\int_{\mathbb{R}} |x| dP_X(x) < +\infty$$

Then the expectation is defined as:

$$E(X) = \int_{\mathbb{R}} x dP_X(x) = \int_{\Omega} X dP = \int_{\omega \in \Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}} \text{Id}_{\mathbb{R}} dP_X$$

From this formula we see that the expectation is completely dependent on the law of the random variable.





# Chapter 2

## Coin tossing games.

### 2.1 The model

We take a fair coin and consider the experiment which consists in throwing  $n$  times the coin. If -1 denotes tail and +1 denotes head then the result of the experiment is  $r \in \{-1, 1\}^n = \Omega$ . Since  $\Omega$  is finite we immediately have  $\mathcal{F} = \mathcal{P}(\Omega)$ . Since we assumed the coin to be fair we have by symmetry that all the results are equiprobable. Hence:

$$\forall \omega \in \Omega, \quad P(\omega) = \frac{1}{|\Omega|} = \frac{1}{2^n}$$

Now let  $X_k$  be the random variable corresponding to the result of the  $k$ -th throw. Formally  $X_k$  is the map:

$$\begin{aligned} X_k : \Omega &\longrightarrow X_k(\Omega) \\ (\omega_1, \dots, \omega_n) &\longmapsto \omega_k \end{aligned}$$

### 2.2 Graphical Representation.

To the sequence  $X_1, \dots, X_n$  we associate the partial sums  $S_0 = 0, S_1 = X_1, \dots, S_n = X_1 + \dots + X_n$ . The sequence  $S_0, \dots, S_n$  contains exactly the same information as  $X_1, \dots, X_n$ . We can therefore represent the result of the experiment by a polygonal line, which is the line which joins successively the points:  $(0, S_0), \dots, (n, S_n)$ . Such a polygonal line, associated to a sequence of signs, will be called a path.

### 2.3 Interpretation of the model.

#### 2.3.1 Coin tossing game

Potter and Voldemort play the following game. Potter throws a coin and Voldemort tries to guess the result. If Voldemort is wrong he pays 1 euro to Potter however if he is right Potter gives 1 euro to Voldemort. Then  $S_n$  represents the algebraic gain of Potter after  $n$  turns.

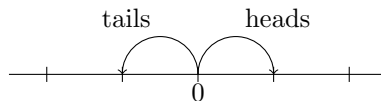
#### 2.3.2 Random Walk

We consider a drunkard which performs a random walk on  $\mathbb{Z}$  with the following mechanism. At  $t = 0$  we center the origin on his position, then at every step he tosses a fair coin and he goes left or right according to whether he obtains tails or heads. Then  $S_n$  corresponds to the final position after  $n$  steps.

### 2.4 Distribution or law of $S_n$

**Proposition 2.5.** *The law of  $S_n$  is the probability distribution on  $\{-n, \dots, n\}$  given by:*

$$\forall k \in \{-n, \dots, n\} \quad P(S_n = k) = \binom{n}{\frac{n+k}{2}} \frac{1}{2^n}$$



*Proof.* A simple proof can be done geometrically. We know that all results are equiprobable hence it suffices to count the number of possible choice that leads to  $S_n = k$ . Let us consider such a path and let us denote by  $\alpha$  the number of ascending steps and  $\beta$  the number of descending steps, then we have that:

$$\begin{cases} \alpha - \beta = k \\ \alpha + \beta = n \end{cases} \Rightarrow \alpha = \frac{n+k}{2}$$

Then the number of possible paths is simply given by the number of possible choices for the ascending steps which are immediately given by:  $\binom{n}{\alpha}$ .  $\square$

## 2.6 Equalization or return to 0

We say that there is an equalization or return to 0 at time  $n$  if  $S_n = 0$ . Since  $n$  and  $S_n$  have the same parity this occurs only at even times. Then from what we got previously we immediately get that:

$$P(S_{2n} = 0) = \frac{1}{2^{2n}} \binom{2n}{n} = \frac{1}{2^{2n}} \frac{(2n)!}{(n!)^2}$$

Then Stirling's formula tells us that:  $n! = \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \left(1 + \frac{1}{12n} + \frac{1}{288n^2} + o\left(\frac{1}{n^2}\right)\right)$  and applying it above we obtain:

$$P(S_{2n} = 0) \stackrel{n \rightarrow +\infty}{\sim} \frac{1}{\sqrt{\pi n}}$$

This gives an excellent approximation even for small values of  $n$ .

## 2.7 The lamplighter walk.

Imagine you have an infinite street with lanterns every meter and one lamplighter whose job is to light all the lanterns. He also starts at the origin and lights the lantern at the origin. Then he throws a coin and goes left or right according to if it is tails or heads and lights the successive lantern. The position at time  $n$  is given by  $S_n$  and the process  $(S_n)_{n \in \mathbb{N}}$  is the symmetric random walk on  $\mathbb{Z}$ . Now we are interested in the following question: What is the probability that the lamplighter comes back to the original lantern? Mathematically this is described by:

$$\begin{aligned} P(\exists n \in \mathbb{N}^*, S_{2n} = 0) &= P\left(\bigcup_{n \geq 1} \{S_{2n} = 0\}\right) \\ &= P(\{S_2 = 0\} \cup (\{S_4 = 0\} \setminus \{S_2 = 0\}) \cup \dots \cup (S_{2n=0} \setminus (\{S_2 = 0\} \cup \dots \cup \{S_{2n-2} = 0\}))) \\ &= P\left(\bigcup_{n \geq 1} (\{S_{2n} = 0\} \setminus (\{S_2 = 0\} \cup \dots \cup \{S_{2n-2} = 0\}))\right) \end{aligned}$$

Now since the events are disjoint we can use the  $\sigma$ -additivity property. Hence we get:

$$P(\exists n \in \mathbb{N}^*, S_{2n} = 0) = \sum_{n \geq 1} P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} = 0)$$

Now necessarily all the  $S_{2i}$  for  $i \in \llbracket 1, n-1 \rrbracket$  must have the same sign. By symmetry the probability that they are all positive is equal to the probability that they are all negative. Hence:

$$\begin{aligned} P(\exists n \in \mathbb{N}^*, S_{2n} = 0) &= 2P(S_1 > 0, \dots, S_{2n-1} > 0, S_{2n} = 0) \\ &= \frac{2}{2^{2n}} |\{\text{paths from } (1, 1) \text{ to } (2n-1, 1) \text{ which do not touch the axis}\}| \end{aligned}$$

Now we decide to count the number of paths that do touch the axis instead of the one that do not touch the axis. Since we know the total number of paths this is an equivalent problem.

## 2.8 Reflection principle.

Let  $A = (a, \alpha)$  and  $B = (b, \beta)$  be two points with  $0 < a < b$  and  $\alpha, \beta > 0$ . My problem now is to count the number of paths that go from  $A$  to  $B$  which touch or cross the axis. Notice that this problem is perfectly

equivalent to counting the number of paths that go from  $A' = (a, -\alpha)$  to  $B$ . To formalize this we need to make sure that this is a one-to-one map. Consider a path  $s = (s_a, s_{a+1}, \dots, s_b)$  be a path from  $A$  to  $B$  which touches the axis and  $t$  be the first time it touches in other words  $t = \min\{i \geq a : s_i = 0\}$ . Let  $T = (t, 0)$  and to the path  $s$  we associate the path  $\phi(s)$  obtained from  $s$  by taking the reflexion of the portion  $AT$  with respect to the  $x$ -axis. Now we claim that  $\phi$  is a one-to-one correspondence from the paths from  $A$  to  $B$  which touch the axis unto the paths from  $A'$  to  $B$ . In fact  $\phi$  is an involution ( $\phi^2 = \text{Id}$ ) and hence it is injective and obviously surjective therefore bijective.

## 2.9 The ballot theorem.

Let  $x, n > 0$ . The number of paths from  $(0, 0)$  to  $(n, x)$  which do not touch the axis after time 0 is equal to the number of paths from  $(1, 1)$  to  $(n, x)$  minus the number of paths from  $(1, -1)$  to  $(n, x)$ . This immediately yields:

$$\frac{x}{n} \binom{n}{\frac{n+x}{2}}$$

An example of application is the following. In an election the candidate  $P$  scores  $p$  votes and  $Q$  scores  $q$  votes with  $p > q$ . The probability that the winning candidate is always ahead during the reading of the votes is given by:

$$\frac{p-q}{p+q}$$

## 2.10 End of the computation.

Now going back to our lamplighter computation applying the reflexion principle we have that:

$$P(\exists n \in \mathbb{N}^*, S_{2n} = 0) = \frac{2}{2^{2n}} \frac{1}{2n-1} \binom{2n-1}{n} = \frac{1}{2n-1} P(S_{2n} = 0)$$

Hence the probability that the lamplighter returns to 0 is given by:

$$P(\text{return to } 0) = \sum_{n \geq 1} \frac{1}{2n-1} P(S_{2n} = 0) = \sum_{n \geq 1} P(S_{2n-2} = 0) - P(S_{2n} = 0) = 1$$

## 2.11 Fundamental Lemma

From the previous computation we have gotten that:

$$P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} = 0) = P(S_{2n-2} = 0) - P(S_{2n} = 0)$$

Yet we also have that:

$$P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} = 0) = P(S_2 \neq 0, \dots, S_{2n-2} \neq 0) - P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} \neq 0)$$

Hence for any  $n$  larger than 1 we have that:

$$P(S_{2n-2} = 0) - P(S_{2n} = 0) = P(S_2 \neq 0, \dots, S_{2n-2} \neq 0) - P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} \neq 0)$$

Moreover we have that:

$$P(S_2 \neq 0) = \frac{1}{2} = P(S_2 = 0)$$

Which gives the fundamental lemma:

$$P(S_2 \neq 0, \dots, S_{2n} \neq 0) = P(S_{2n} = 0)$$

## 2.12 Last tie

Consider a coin tossing game of length  $2n$  and the time  $T$  of the last tie before  $2n$ :  $T = \max\{k \leq 2n : S_k = 0\}$ . Now we want to find the distribution of  $T$ .

**Proposition 2.13.** *The law of  $T$  is called the arcsinus law and is given by:*

$$P(T = k) = P(S_{2k} = 0)P(S_{2n-2k} = 0)$$