# <u>Probabilité et Simulation</u>

### PolyTech INFO4 (Grenoble) – TD

Last updated: 2025-10-03

Contact: davide.legacci@univ-grenoble-alpes.fr

**References**

- Fundamentals: [1] B. Jourdain, *Probabilités et statistique pour l'ingénieur.* 2018.
- Further reading: [2] P. Billingsley, *Probability and Measure.* John Wiley & Sons, 2012.

**Websites**

- CM: https://github.com/jonatha-anselmi/INFO4-PS
- TD: https://github.com/davidelegacci/probasim24

## Contents

# 1. Theory recap - Probability on finite set

Reminder: a set is

- *finite* (*fini*), if it has a finite number of elements
- *countable* (*dénombrable*), if either it is finite, or it can be made in one to one correspondence with the set of natural numbers

## 1.1. Finite probability space

- *Random process*: one random outcome out of finitely many
- *Sample space* $\Omega =$ *finite* set of possible *outcomes* $\omega$
- *Probability* on $\Omega =$ set of weights $\mathbb{P}(\omega) \in \mathbb{R}$ on each $\omega \in \Omega$ such that
  - $\mathbb{P}(\omega) > 0 \forall \omega \in \Omega$
  - $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$

---

- *Event* $A \subseteq \Omega =$ any subset of the sample space

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

- *Complementary event* $A^c = \Omega / A$ (pronounced "*not* $A$")
- "$A$ *and* $B$" $= A \cap B$
- "$A$ *or* $B$" $= A \cup B$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \tag{2}$$

- Probability of complementary event

$$\Omega = A^c \sqcup A \Rightarrow 1 = \mathbb{P}(A^c) + \mathbb{P}(A) \tag{3}$$

---

- *Indicator function* of event $A \subseteq \Omega$

$$\mathbb{1}_A : \Omega \to \mathbb{R}, \quad \mathbb{1}_A(\omega) = \begin{cases} 1 \text{ if } \omega \in A \\ 0 \text{ if } \omega \notin A \end{cases} \tag{4}$$

- Denoted by $|S|$ the *cardinality* of a set $S$
- Denoted by $S^n$ the *cartesian product* of $S$ with itself $n$ times

$$S^n = S \times S \times ... \times S = \{(s_1, ..., s_n) : s_i \in S \text{ for all } i = 1, ..., n\} \tag{5}$$

- *Cardinality of cartesian product*

$$|S^n| = |S|^n \tag{6}$$

## 1.2. Uniform probability

- *Every outcome* $\omega \in \Omega$ *has the same weight*

$$\mathbb{P}(\omega) = \frac{1}{|\Omega|} \tag{7}$$

- *Uniform probability of the event* $A \subseteq \Omega$

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|} \tag{8}$$

## 1.3. Counting

- Number of *permutations* of $k$ elements:
  - Number of ways to *order* $k$ elements
  - **Only order matters**

$$P_k = k! \tag{9}$$

Permutation of 5 elements

| | | | | |
|---|---|---|---|---|
| 1 | 5 | 4 | 3 | 2 |

- Number of *dispositions* of $k$ elements out of $n$ $(k \leq n)$:
  - Number of ways to *choose and order* $k$ elements out of $n$
  - **Order and elements** matter
  - Number of injections : $\{1, ..., k\} \to \{1, ..., n\}$

$$D_{n,k} = \underbrace{n(n-1)...}_{k \text{ times}} = n(n-1)...(n-k+1) = \frac{n!}{(n-k)!} \tag{10}$$

Disposition of 3 elements out of 5

| | | | | |
|---|---|---|---|---|
| | 2 | 1 | 3 | |

- Number of *combinations* of $k$ elements out of $n$ $(k \leq n)$:
  - Number of ways to *choose* $k$ elements out of $n$
  - **Only elements** matter
  - Number of subsets of cardinality $k$ of a set of cardinality $n$
  - Number of dispositions modulo number of permutations

$$C_{n,k} = \frac{D_{n,k}}{P_k} = \frac{n!}{k!(n-k)!} = \binom{n}{k} = \text{choose}(n, k) \tag{11}$$

Combination of 3 elements out of 5

| | | | | |
|---|---|---|---|---|
| | | X | X | X |

- *Binomial theorem*

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j \tag{12}$$

## Exercises

**Exercise 1.1** *(Handshakes and kisses)*

There are $f$ girls and $g$ boys in a room. Boys exchange handshakes, girs exchange kisses, boys and girls exchange kisses. How many kisses in total?

The number of kisses exchanegs among girls is the number of subsets of cardinality 2 of a set of cardinality $f$, that is $\binom{f}{2} = \frac{f(f-1)}{2}$. Or, think that each girl gives $f-1$ kisses, and one needs a factor of one half to avoid double counting.

For the number of kisses exchanged between boys and girls: the first girl gives $g$ kisses, the second girl gives $g$ kisses, and so on, so we have $fg$ in total.

$$\text{number of kisses} = \frac{f(f-1)}{2} + fg \tag{13}$$

**Exercise 1.2** *(Throwing a dice)* Throw a fair dice with $f$ faces $n$ times. What's the prob to never get the same result twice?

**General strategy**

- Identify sample space $\Omega$ (write in set-theoretic notation!) and its cardinality $|\Omega|$
- Identify event $A \subseteq \Omega$ (write in set-theoretic notation!) and its cardinality $|A|$
- Uniform probability? If so, use $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$

Let $\mathcal{N} = \{1, ..., n\}$ and $\mathcal{F} = \{1, ..., f\}$. The sample space is

$$\Omega = \{\omega = (\omega_1, ..., \omega_n) : \omega_i \in \mathcal{F} \text{ for all } i \in \mathcal{N}\} = \mathcal{F}^n \tag{14}$$

with cardinality

$$|\Omega| = |\mathcal{F}^n| = |\mathcal{F}|^n = f^n \tag{15}$$

Endow the sample space with the uniform probability (since every outcome of the experiment is equiprobable).

The event we're looking at is

$$A = \left\{\omega \in \Omega : \omega_i \neq \omega_j \text{ for all } i \neq j \in \mathcal{N}\right\} \tag{16}$$

Clearly if $n > f$ then $\mathbb{P}(A) = 0$. Let $n \leq f$. The (uniform) probability of the event $A$ is $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$, with

$$\begin{aligned} |A| &= \#\text{ of ways to choose and order } n \text{ elements out of } f \\ &= \underbrace{f(f-1)...}_{n} = f(f-1)...(f-n+1) = \frac{f!}{(f-n)!} \end{aligned} \tag{17}$$

$$\mathbb{P}(A) = \frac{f!}{f^n(f-n)!} \tag{18}$$

**Exercise 1.3** *(Birthday paradox)* What is the probability that at least 2 people out of $n$ have the same birthday? (Assume: uniform birth probability and year with $y$ number of days).

**Quick solution**

$$
\begin{aligned}
\mathbb{P}(A) &= 1 - \mathbb{P}\Bigg(\underbrace{\text{no two people have the same birthday}}_{\text{Ex. 2}}\Bigg) \\
&= 1 - \frac{y!}{y^n (y-n)!}
\end{aligned} \tag{19}
$$

**Formal solution** Let $\mathcal{N} = \{1, ..., n\}$ and $\mathcal{Y} = \{1, ..., y\}$ with $n \leq y$. The sample space is

$$
\begin{aligned}
\Omega &= \text{distributions of possible birthdays of } n \text{ people} \\
&= \{\omega = (\omega_1, ..., \omega_n) : \omega_i \in \mathcal{Y} \text{ for all } i \in \mathcal{N}\} = \mathcal{Y}^n
\end{aligned} \tag{20}
$$

where $\omega_i$ is the birthday of the $i$-th person. The cardinality of the sample space is

$$
|\Omega| = |\mathcal{Y}^n| = |\mathcal{Y}|^n = y^n \tag{21}
$$

The event we're looking at is

$$
A = \{\omega \in \Omega : \exists i \neq j \in \mathcal{N} : \omega_i = \omega_j\} \tag{22}
$$

Note that this is the complementary event to the event defined in Equation 16 of Exercise 2. Thus we can compute its probability as

$$
\mathbb{P}(A) = 1 - \mathbb{P}(A^c) \tag{23}
$$

in agreement with Equation 19.

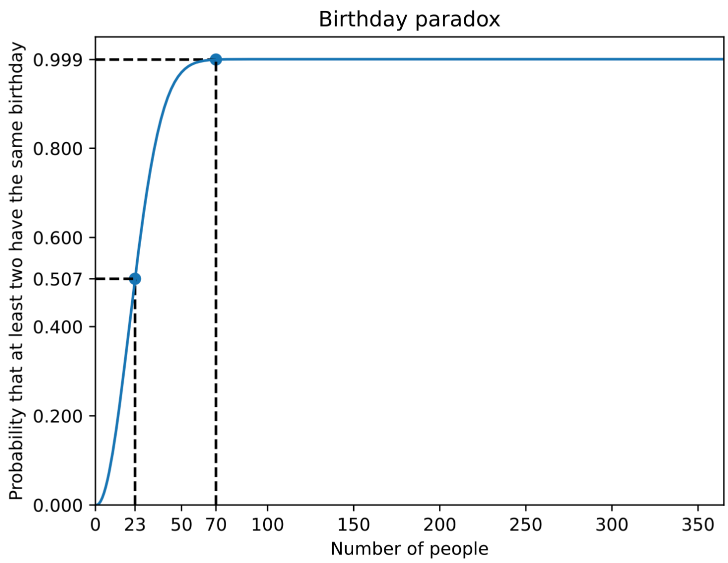


Figure 4: Birthday paradox probability. Code available.

> **Exercise 1.4** *(Same birthday as the prof)* What is the probability that at least 1 student out of $n$ has the same birthday of the prof? (Assume: uniform birth probability and year with $y$ number of days).

**Quick solution**

$$
\begin{aligned}
\mathbb{P}(A) &= 1 - \mathbb{P}\Big(\underbrace{\text{nobody has the prescribed birth date}}\Big) \\
&= 1 - \left(\frac{y-1}{y}\right)^n
\end{aligned} \tag{24}
$$

**Formal solution 1** As above $\mathcal{N} = \{1, ..., n\}$ and $\mathcal{Y} = \{1, ..., y\}$ with $n \leq y$. The sample space is $\Omega = \mathcal{Y}^{n+1}$, that is the set of possible birthdays of $n+1$ people, the $(n+1)$-th being the prof. Its cardinality is $|\Omega| = y^{n+1}$. The event we're looking at is

$$A = \{\omega \in \Omega : \exists i \in \mathcal{N} : \omega_i = \omega_{n+1}\} \tag{25}$$

with complementary event

$$A^c = \{\omega \in \Omega : \omega_i \neq \omega_{n+1} \forall i \in \mathcal{N}\} \tag{26}$$

As usual $\mathbb{P}(A) = 1 - \mathbb{P}(A^c) = 1 - \frac{|A^c|}{|\Omega|}$, with

$$|A^c| = \underbrace{y}_{\text{prof}} \cdot \underbrace{(y-1)^n}_{\text{students}} \tag{27}$$

So, $\mathbb{P}(A) = 1 - \frac{y(y-1)^n}{y^{n+1}} = 1 - \left(\frac{y-1}{y}\right)^n$, in agreement with Equation 24.

Note that a factor $\frac{y}{y}$, corresponding to the prof's birthday, simplifies in the last step. Alternatively, you can fix the birthday of the prof and exclude it from the analysis from the beginning. In this case $|\Omega| = y^n$ and $|A^c| = (y-1)^n$, leading to the same result.

**Formal solution 2** Using the probability of the complementary event is often the smartest way to proceed, but for the sake of completeness let's see how to get the same result directly. Let $A_j$ be the event *"exactly $j$ students out of $n$ have the same birthday as the prof"*. The event we look at then is

$$A = \sqcup_{j \in \mathcal{N}} A_j \tag{28}$$

with probability (cf Equation 2)

$$\mathbb{P}(A) = \sum_{j \in \mathcal{N}} \mathbb{P}(A_j) = \frac{\sum_{j \in \mathcal{N}} |A_j|}{|\Omega|} \tag{29}$$

The cardinality of $A_j$ is

$$\begin{aligned} |A_j| &= \underbrace{1...1}_{j \text{ times}} \cdot \underbrace{(y-1)...(y-1)}_{n-j \text{ times}} \cdot \underbrace{y}_{\text{prof}} \cdot \underbrace{\binom{n}{j}}_{\text{number of ways to choose } j \text{ elements out of } n} \\ &= y(y-1)^{n-j}\binom{n}{j} \end{aligned} \tag{30}$$

By an application of the binomial theorem (Equation 12) and a short manipulation,

$$\sum_{j=1}^{n} |A_j| = y(y^n - (y-1)^n) \tag{31}$$

which leads back to Equation 24.

# 2. Theory recap - Conditional probability and independence

## 2.1. Conditional probability

- Let $\mathbb{P}$ be a probability on $\Omega$, and consider the events $A, B \subseteq \Omega$.
- *Conditional probability*: probability of $A$ *given* $B$

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ if } \mathbb{P}(B) \neq 0 \tag{32}$$

- not reallty defined if $\mathbb{P}(B) = 0$, cf [2] pag. 427.
- often used as

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B) \tag{33}$$

- Conditional probability and complementary event (proof: simple exercise.)

$$\mathbb{P}(A \mid B) + \mathbb{P}(A^c \mid B) = 1 \tag{34}$$

- *Total Probability Theorem*

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c) \tag{35}$$

- *Bayes theorem*

$$\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A) \Rightarrow \mathbb{P}(A \mid B)\mathbb{P}(B) = \mathbb{P}(B \mid A)\mathbb{P}(A) \tag{36}$$

See this notebook for an example of Bayes theorem in action.

## 2.2. Independent events

Let $\Omega$ be equipped with a probability $\mathbb{P}$.

- two events $A, B \subseteq \Omega$ are said *independent* if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \tag{37}$$

- equivalently, by definition of conditional probability, $A$ and $B$ are independent if

$$\mathbb{P}(A \mid B) = \mathbb{P}(A) \tag{38}$$

- $n$ events $A_1, ..., A_n$ are said *independent* if

$$\mathbb{P}(\cap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i) \text{ for all } I \subseteq \{1, ..., n\} \tag{39}$$

- pairwise independence does not imply independence of all events!

# Exercises

**Exercise 2.1** *(Pile ou Face)* Jet de 2 pieces, $\Omega = \{\mathrm{PP}, \mathrm{PF}, \mathrm{FP}, \mathrm{FF}\}$. Cet espace est muni de la probabilite uniforme. Soient les evenements:

- $A = \text{1ere piece donne P}$
- $B = \text{2eme piece donne F}$
- $C = \text{le deux pieces donnent le meme resultat}$

Questions:

- $A$ et $B$ sont indépendantes?
- $A, B$ et $C$ sont indépendants?

$$\begin{aligned}
A &= \{PP, PF\} & \mathbb{P}(A) &= 1/2 & \\
B &= \{PF, FF\} & \mathbb{P}(B) &= 1/2 & \\
C &= \{PP, FF\} & \mathbb{P}(C) &= 1/2 & \\
A \cap B &= \{PF\} & \mathbb{P}(A \cap B) &= 1/4 & &= \mathbb{P}(A)\mathbb{P}(B) \\
A \cap C &= \{PP\} & \mathbb{P}(A \cap C) &= 1/4 & &= \mathbb{P}(A)\mathbb{P}(C) \\
B \cap C &= \{FF\} & \mathbb{P}(B \cap C) &= 1/4 & &= \mathbb{P}(B)\mathbb{P}(C) \\
A \cap B \cap C &= \emptyset & \mathbb{P}(A \cap B \cap C) &= 0 & &\neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).
\end{aligned}$$

Ainsi les événements $A, B$ et $C$ sont 2 à 2 indépendants mais pas indépendants.

Figure 5: Pairwise independence does not imply independence of all events!

**Exercise 2.2** *(Pieces mecaniques defectueuses)* Parmi 10 pièces mécaniques, 4 sont déefectueuses. On prend successivement deux pièces au hasard dans le lot (sans remise). Quelle est la probabilité pour que les deux pièces soient correctes?

**Solution 1** Let $A_i$ be the event *the i-th drawn piece is good*, with $i \in \{1, 2\}$. We need the probability of the event $A_2 \cap A_1$. By definition of conditional probavility,

$$\mathbb{P}(A_2 \cap A_1) = \underbrace{\mathbb{P}(A_2 \mid A_1)}_{\frac{5}{9}} \underbrace{\mathbb{P}(A_1)}_{\frac{6}{10}} = \frac{1}{3}. \tag{40}$$

**Solution 2** The sample space is the set of subsets of cardinality 2 of a set of cardinality 10, so $|\Omega| = \binom{10}{2}$. The event we consider is the set of subsets of cardinality 2 of a set of cardinality 6, so $|A| = \binom{6}{2}$. Then

$$\mathbb{P}(A) = \frac{\binom{6}{2}}{\binom{10}{2}} = \frac{6 \cdot 5}{10 \cdot 9} = \frac{1}{3}. \tag{41}$$

**Exercise 2.3** *(Betting on cards)* We have three cards:

- a *red* card with both faces red;
- a *white* card with both faces white;
- a *mixed* card with a red face and a white face.

One of the three cards is drawn at random and one of the faces of this card (also chosen at random) is exposed. This face is red. You are asked to bet on the color of the hidden face. Do you choose red or white?

**Intuitive solution** Since we observe a red face, the white card cannot be on the table. There are three possibilities left: 1. we're observing a face of the red card (in which case the hidden face is red); 2. we are observing the other face of the red card (in which case the hidden face is red); 3. we are observing the red face of the mixed card (in which case the hidden face is white). So the hidden face is red 2 out of 3 times.

**Formal solution** Denote by $R$ a red face and by $W$ a white face, and denote by a subscript $o$ the observed face, and by a subscript $h$ the hidden face. The possible outcomes then are

$$\Omega = \{R_h \cap R_o, R_h \cap W_o, W_h \cap R_o, W_h \cap W_o\}. \tag{42}$$

For example, $W_h \cap R_o$ is the event "*the hidden face is white and the observed face is red*". We are given the information that the observed face is red, so we need to find the probability that the hidden face has a certain color *given* the fact that the observed face is red. In other words, we need to find the probabilities

$$\mathbb{P}(W_h \mid R_o) \text{ and } \mathbb{P}(R_h \mid R_o). \tag{43}$$

Clearly it suffices to find one of them, since by Equation 34 they sum to 1. By definition of conditional probability (Equation 32)

$$\mathbb{P}(R_h \mid R_o) = \frac{\mathbb{P}(R_h \cap R_o)}{\mathbb{P}(R_o)}. \tag{44}$$

The numerato of Equation 44 corresponds to the probability of drawing the red card (event denoted by $r$), which is $\frac{1}{3}$:

$$
\begin{aligned}
\mathbb{P}(R_h \cap R_o) &= \mathbb{P}(R_o \cap r) \\
&= \underbrace{P(R_o \mid r)}_{1} \underbrace{P(r)}_{\frac{1}{3}} = \frac{1}{3}.
\end{aligned} \tag{45}
$$

The denominator of Equation 44 is the probability of drawing the red card, plus the probability of drawing the mixed card (event denoted by $m$) and observing the red face, that is $\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{2}$. This follows formally form the Total Probability Theorem (Equation 35) and the definition of conditional probability (Equation 32):

$$
\begin{aligned}
\mathbb{P}(R_o) &= \mathbb{P}(R_o \cap R_h) + \mathbb{P}(R_o \cap W_h) \\
&= \frac{1}{3} + \qquad \mathbb{P}(R_o \cap m) \\
&= \frac{1}{3} + \qquad \underbrace{P(R_o \mid m)}_{\frac{1}{2}} \underbrace{P(m)}_{\frac{1}{3}} = \frac{1}{2}
\end{aligned} \tag{46}
$$

In conclusion, given that the observed face is red, the hidden face is red with probability $2/3$:

$$
\mathbb{P}(R_h \mid R_o) = \frac{1/3}{1/2} = \frac{2}{3}. \tag{47}
$$

**Exercise 2.4** *(Russian roulette)* You are playing two-person Russian roulette with a revolver featuring a rotating cylinder with six bullet slots. Each time the gun is triggered, the cylinder rotates by one slot. Two bullets are inserted one next to the other into the cylinder, which is then randomly positioned. Your opponent is the first to place the revolver against her temple. She presses the trigger and… she stays alive. With great display of honor, she offers you to rotate the barrel again at random before firing in turn. What do you decide?

The bullets occupy the positions $x$ and $x + 1 \bmod 6$:

$$
\Omega = \{12, 23, 34, 45, 56, 61\} \tag{48}
$$

Say the revolver shots from position 1. The event "*the first player dies*" is

$$
\mathrm{die}_1 = \{12, 61\} \tag{49}
$$

so $\mathbb{P}(\mathrm{die}_1) = \frac{1}{3}$ and $\mathbb{P}(\mathrm{live}_1) = \frac{2}{3}$. We need to compute

$$
\mathbb{P}(\mathrm{die}_2 \mid \mathrm{live}_1) = \frac{\mathbb{P}(\mathrm{die}_2 \cap \mathrm{live}_1)}{\mathbb{P}(\mathrm{live}_1)} \tag{50}
$$

Since the cylynder rotates after being triggered we have $\mathrm{die}_2 = \{56, 61\}$ and $\mathrm{die}_2 \cap \mathrm{live}_1 = \{56\}$, so $\mathbb{P}(\mathrm{die}_2 \mid \mathrm{live}_1) = \frac{1}{6} / \frac{2}{3} = \frac{1}{4} < P(\mathrm{die}_1)$. So you don't shuffle the barrel.

# 3. Theory recap - Probability space, dicrete random variables, and distributions

## 3.1. Probability measure

The relevant references are [1] pag. 11 and [2], pag. 22 and 160.

**Definition 3.1.1** *(sigma-field)* Let $\Omega$ be any set. A *$\sigma$-field $\mathcal{A}$* on $\Omega$ is a collection of subsets of $\Omega$ that[1]

1. is closed under complement: if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$;
2. contains the whole set: $\Omega \in \mathcal{A}$;
3. is closed under countable union: if $A_1, A_2, \ldots$ is a countable family of sets of $\mathcal{A}$ then their union $\cup_{i=1}^{\infty} A_i$ is also in $\mathcal{A}$.

A subset of $\Omega$ that is in $\mathcal{A}$ is called *event.*

**Definition 3.1.2** *(Measure)* Given a set $\Omega$ and a $\sigma$-algebra $\mathcal{A}$ on $\Omega$, a *measure* $\mu$ is a function

$$\mu : \mathcal{A} \to \mathbb{R}_{\geq 0} \tag{51}$$

such that
1. $\mu(\emptyset) = 0$
2. *countable additivity* (also called *$\sigma$-additivity*) is fulfilled, namely the measure of a *disjoint* countable union of sets in $\mathcal{A}$ is the sum of their measures:

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i). \tag{52}$$

**Definition 3.1.3** *(Probability measure)* Given a set $\Omega$ and a $\sigma$-algebra $\mathcal{A}$ on $\Omega$, a *probaility measure* $\mathbb{P}$ is a measure (in the sense above) with the additional requirement that

$$\mathbb{P}(\Omega) = 1. \tag{53}$$

- Note that this implies that $\mathbb{P}(A) \leq 1$ for all events $A \in \mathcal{A}$.
- A triple $(\Omega, \mathcal{A}, \mathbb{P})$ where $\mathcal{A}$ is a $\sigma$-algebra on $\Omega$ and $\mathbb{P}$ is a probability measure is called *probability space.*

**Tip** Putting all together, a probability measure $\mathbb{P} : \mathcal{A} \to [0, 1]$ on a space $\Omega$ is a function from a "well-behaved" collection of subsets of $\Omega$ (the $\sigma$-field) to $[0, 1]$, such that $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$, and fulfilling countable additivity.

## 3.2. Discrete random variables

**Definition 3.2.1** *(Discrete random variable)* Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a *discrete random variable* $X$ is a function $X : \Omega \to F$ such that
1. $F$ is a *countable* set; we call it the space of *values* of the random variable, and we say that any $x \in F$ is a *value.*
2. *the level sets of $X$ are events*, that is

$$\{X = x\} := \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{A} \text{ for all } x \in F \tag{54}$$

- clearly, $\{X = x\} = \emptyset \in \mathcal{A}$ for all $x \in F \setminus \operatorname{Im}(X)$
- property 2. guarantees that $\mathbb{P}\{X = x\}$ is well-defined for all $x \in F$, which allows for the following definition:

**Definition 3.2.2** *(Distribution of a discrete random variable)* The *distribution* (or *law*) of a random variable $X$ is the function $\mu_X : F \to [0, 1]$ defined by

$$\mu_X(x) = \mathbb{P}\{X = x\} \text{ for all } x \in F. \tag{55}$$

In words: for any value, the law of a random variable returns the probability that the random variable takes the given value.

[1]In french, this set is called *tribu* on $\Omega$. The term *$\sigma$-algebra* is also used – and is more common in the context of pure analysis, c.f. [3] – whereas the term *$\sigma$-field* is more common in the context of probabioloty theory, c.f. [2].

- two discrete random variables $X$ and $Y$ taking values resp. in $F$ and $G$ are *independent* if

$$\mathbb{P}\{X = x, Y = y\} = \mathbb{P}\{X = x\}\mathbb{P}\{Y = y\} \text{ for all } x \in F, y \in G \tag{56}$$

- it is understood that $\{X = x, Y = y\}$ is a shorthand for the event

$$\{\omega \in \Omega : X(\omega) = x\} \cap \{\omega \in \Omega : Y(\omega) = y\} \in \mathcal{A}. \tag{57}$$

- the definition generalises to collections of DRVs, see Section 2.2.3 in [1].

> **Tip** A discrete random variable is a *function on* $\Omega$ with *countably many values*. Think of it as an experiment with countably many random outcomes. Its *law, or distribution*, gives the probability to observe each of the possible (countable) values of the random variable.

> **Tip**
> - $(\Omega, \mathcal{A}, \mathbb{P})$ with $\mathbb{P} : \mathcal{A} \to [0, 1]$ and $\mathbb{P}(\Omega) = 1$
> - $X : \Omega \to F$ countable values space, with $\{X = x\} \in \mathcal{A}$ is an event for all $x \in F$
> - $\mu_X : F \to [0, 1]$ such that $\mu_X(x) = \mathbb{P}\{X = x\}$ is the distribution of $X$

## 3.3. Standard discrete distributions

### 3.3.1. Bernoulli $\mathcal{B}(p)$

- The Bernoulli distribution models a random experiment which has two possible outcomes.
- More precisely, the Bernoulli distribution is the distribution of a discrete random variable $X$ that can assume only values in $F = \{0, 1\}$.
- The distribution is parametrized by $p \in [0, 1]$, the probability to observe 1. We then have

$$\begin{aligned} \mu : \{0, 1\} &\to [0, 1] \\ 1 &\longmapsto \quad p \\ 0 &\longmapsto \quad 1 - p \\ x &\longmapsto \quad p^x (1 - p)^{1-x} \end{aligned} \tag{58}$$

- When a random variable $X$ follows the Bernoulli distribution we write $X \sim \mathcal{B}(p)$.

### 3.3.2. Binomial $\mathcal{B}(n, p)$

- Distribution of the discrete random variable $X = X_1 + \ldots + X_n$, where the $X_i$-s are independent Bernoulli variables of parameter $p \in [0, 1]$.
- $F = \{0, \ldots, n\}$; $k \in F$ is value of the sum

$$\begin{aligned} \mu : F &\to [0, 1] \\ k &\longmapsto \binom{n}{k} p^k (1 - p)^{n-k} \end{aligned} \tag{59}$$

### 3.3.3. Poisson $\mathcal{P}(\lambda)$

- probability of observing a given number of independent events occurring at constant rate $\lambda > 0$
- $F = \mathbb{N}$; $n \in F$ is number of observed events

$$\begin{aligned} \mu : F &\to [0, 1] \\ n &\longmapsto e^{-\lambda} \frac{\lambda^n}{n!} \end{aligned} \tag{60}$$

### 3.3.4. Geometric $\mathcal{G}(p)$

- First successful event from a sequence of independent $p$-Bernoulli events.

- $F = \mathbb{N}^*$; $k \in F$ is first succesful event

$$\begin{aligned} \mu : F &\to [0,1] \\ k &\longmapsto p(1-p)^{k-1} \end{aligned} \tag{61}$$

### 3.4. Useful stuff

- Vandermonde's identity

$$\sum_{k_1=0}^{k} \binom{n_1}{k_1}\binom{n_2}{k-k_1} = \binom{n_1+n_2}{k} \tag{62}$$

## Exercises

**Exercise 3.1** *(Sum of independent binomial distributions)* Let $X_i \sim \mathcal{B}(n_i, p)$ with $i \in \{1,2\}$ be independent discrete random variables following the Bernoulli law. Find the law of $X_1 + X_2$.

*Hint: c.f. derivation of binomial distribution [1] pag. 16.*

The laws $\mu_i : F_i = \{0, ..., n_i\} \to [0,1]$ of the two variables are given by

$$\mu_i(k_i) = \mathbb{P}(X_i = k_i) = \binom{n_i}{k_i} p^{k_i}(1-p)^{n_i-k_i} \tag{63}$$

The law of $X_1 + X_2$ takes value in $F = \{0, ..., n_1 + n_2\}$ and for all $k \in F$ is given by

$$\begin{aligned} \mu(k) &= \mathbb{P}(X_1 + X_2 = k) \\ &= \mathbb{P}\left(\bigsqcup_{\substack{k_i \in F_i \\ k_1+k_2=k}} \{X_1 = k_1, X_2 = k_2\}\right) \\ &= \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \mathbb{P}(X_1 = k_1)\mathbb{P}(X_2 = k_2) && \text{by c. add and indep.} \\ &= \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \mu(k_1)\mu(k_2) && \text{by def of law} \\ &= \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \binom{n_1}{k_1}\binom{n_2}{k_2} p^{k_1+k_2}(1-p)^{n_1+n_2-k_1-k_2} \\ &= p^k(1-p)^{n_1+n_2-k} \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \binom{n_1}{k_1}\binom{n_2}{k_2} \end{aligned} \tag{64}$$

Let's focus on the sum. For each fixed $k_1 \in F_1$, $k_2$ is constrained to be $k - k_1$. Furthermore, in order for $k_2$ to be $\geq 0$, $k_1$ can be at most equal to $k$. So the constraints

$$\begin{aligned} k_1 &\in \{0, ..., n_1\} \\ k_2 &\in \{0, ..., n_2\} \\ k_1 + k_2 &= k \end{aligned} \tag{65}$$

can be replaced by the constraints

$$
\begin{aligned}
k_1 &\in \{0, ..., k\} \\
k_2 &= k - k_1
\end{aligned} \tag{66}
$$

namely

$$
\sum_{\substack{k_i \in F_i \\ k_1 + k_2 = k}} \binom{n_1}{k_1} \binom{n_2}{k_2} = \sum_{k_1=0}^{k} \binom{n_1}{k_1} \binom{n_2}{k - k_1} = \binom{n_1 + n_2}{k} \tag{67}
$$

where the second step follows by by Vandermonde's identity.

**Remark 3.4.1** Note that it is correct to have $k_1$ running from 0 to $k$:

- If $k \leq n_1$, $k_1$ can be at most $k$ so that $k_2 = k - k_1 \geq 0$, so we have directly Equation 67.
- If $k > n_1$, we have

$$
\begin{aligned}
\sum_{\substack{k_i \in F_i \\ k_1 + k_2 = k}} \binom{n_1}{k_1} \binom{n_2}{k_2} &= \sum_{k_1=0}^{n_1} \binom{n_1}{k_1} \binom{n_2}{k - k_1} \\
&= \sum_{k_1=0}^{n_1} \binom{n_1}{k_1} \binom{n_2}{k - k_1} + \sum_{k_1 = n_1 + 1}^{k} \binom{n_1}{k_1} \binom{n_2}{k - k_1}
\end{aligned} \tag{68}
$$

since each summand in the the second sum is zero[2], and we get again Equation 67.

So in conclusion

$$
\mu(k) = \binom{n_1 + n_2}{k} p^k (1 - p)^{n_1 + n_2 - k} \tag{69}
$$

namely $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$.

**Exercise 3.2** *(Sum of independent Poisson distributions)* Let $X_i \sim \mathcal{P}(\lambda_i)$ with $i \in \{1, 2\}$ be independent discrete random variables following the Poisson law. Find the law of $X_1 + X_2$.

*Hint: c.f. previous exercise and binomial theorem.*

Analogously to before, with $i \in \{1, 2\}$, we look for the law $\mu : \mathbb{N} \to [0, 1]$ given by

$$
\begin{aligned}
\mu(n) = \mathbb{P}(X_1 + X_2 = n) &= \sum_{\substack{n_i \in \mathbb{N} \\ n_1 + n_2 = n}} \mu_1(n_1) \mu_2(n_2) \\
&= \sum_{\substack{n_i \in \mathbb{N} \\ n_1 + n_2 = n}} e^{-\lambda_1} \frac{\lambda_1^{n_1}}{n_1!} e^{-\lambda_2} \frac{\lambda_2^{n_2}}{n_2!}
\end{aligned} \tag{70}
$$

As before we replace the consraint by $n_1 \in \{0, ..., n\}$ and $n_2 = n - n_1$, so

[2] Recall that $\binom{a}{b} = 0$ if $b > a$.

$$\begin{aligned}\mu(n) &= e^{-(\lambda_1+\lambda_2)} \sum_{n_1=0}^{n} \frac{\lambda_1^{n_1}}{n_1!} \frac{\lambda_2^{n-n_1}}{n-n_1!} \cdot \frac{n!}{n!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{n_1=0}^{n} \binom{n}{n_1} \lambda_1^{n_1} \lambda_2^{n-n_1} \\ &= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!}\end{aligned} \tag{71}$$

So $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

**Exercise 3.3** *(Min of independent geometric distributions)* Let $X_i \sim \mathcal{G}(p_i)$ with $i \in \{1, 2\}$ be independent DRVs following the geometric law. Find the law of $\min\{X_1 + X_2\}$.

*Hint: set up the problem in terms of inequalities. Recall that the geometric series $\sum_{i=0}^{\infty} r^i$ with $r \in [0, 1)$ converges to $\frac{1}{1-r}$.*

Let $Z = \min\{X_1 + X_2\}$. We look for the law $\mu : \mathbb{N}^* \to [0, 1]$ such that

$$\begin{aligned}\mu(k) &= \mathbb{P}(Z = k) \\ &= \mathbb{P}(Z \geq k) - \mathbb{P}(Z \geq k+1) \\ &= \mathbb{P}(X_1 \geq k, X_2 \geq k) - \mathbb{P}(X_1 \geq k+1, X_2 \geq k+1) \\ &= \mathbb{P}(X_1 \geq k)\mathbb{P}(X_2 \geq k) - \mathbb{P}(X_1 \geq k+1)\mathbb{P}(X_2 \geq k+1)\end{aligned} \tag{72}$$

Let's drop the subscript for a moment. For a DRV $X \sim \mathcal{G}(p)$ and for $k \in \mathbb{N}^*$ we need

$$\begin{aligned}\mathbb{P}(X \geq k) &= \mathbb{P}\left(\bigsqcup_{i \geq k} (X = i)\right) \\ &= \sum_{i \geq k} \mathbb{P}(X = i) \\ &= \sum_{i \geq k} p(1-p)^{i-1} \\ &= p(1-p)^{k-1} + p(1-p)^k + p(1-p)^{k+1} + \dots \\ &= p(1-p)^{k-1}\left(1 + (1-p) + (1-p)^2 + \dots\right) \\ &= p(1-p)^{k-1} \sum_{j=0}^{\infty} (1-p)^j \\ &= (1-p)^{k-1}\end{aligned} \tag{73}$$

Plugging in Equation 72 we get

$$\begin{aligned}\mu(k) &= \mathbb{P}(X_1 \geq k)\mathbb{P}(X_2 \geq k) - \mathbb{P}(X_1 \geq k+1)\mathbb{P}(X_2 \geq k+1) \\ &= (1-p_1)^{k-1}(1-p_2)^{k-1} - (1-p_1)^k(1-p_2)^k \\ &= (1-p_1)^{k-1}(1-p_2)^{k-1}[1 - (1-p_1)(1-p_2)]\end{aligned} \tag{74}$$

Let $\alpha = (1-p_1)(1-p_2)$ and $\beta = 1 - \alpha = p_1 + p_2 - p_1 p_2$, so $\mu(k) = \beta(1-\beta)^{k-1}$, i.e. $Z \sim \mathcal{G}(\beta)$.

**Exercise 3.4** *(*Lost messages)* If there is extra time, do Exercise 4.1, else do in next session.