Probabilité et Simulation

PolyTech INFO4 (Grenoble) – 2024-2025 – Practical Sessions

Last updated: 2024-10-02. Info: davide.legacci@univ-grenoble-alpes.fr

1. Theory recap 11.9.24

- *Jet set* Ω = finite set of possible outcomes ω
- *Probability* on Ω = set of weights $P(\omega) \in \mathbb{R}$ on each $\omega \in \Omega$ such that
  - $P(\omega) > 0 \forall \omega \in \Omega$
  - $\sum_{\omega \in \Omega} P(\omega) = 1$
- *Event* $A \subseteq \Omega$ = subset of the jet set
- *Complementary event* $A^c = \Omega / A$
- The cardinality of a set S is denoted by $|S|$
- *Uniform probability of the event* A

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

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

- *Binomial theorem*

$$(x+y)^n = \sum_{j=0}^{n} \binom{n}{j} x^{n-j} y^j \tag{3}$$

1.1. Counting

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

$$P_k = k! \tag{4}$$

- 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 $f : \{1, ..., k\} \to \{1, ..., n\}$

$$D_{n,k} = \underbrace{n(n-1)...}_{k \text{ times}} = n(n-1)...(n-k+1) = \frac{n!}{(n-k)!} \tag{5}$$

- 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

$$C_{n,k} = \frac{D_{n,k}}{P_k} = \frac{n!}{k!(n-k)!} = \binom{n}{k} = \text{choose}(n, k) \tag{6}$$

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.

$$\text{number of kisses} = \frac{f(f-1)}{2} + fg \tag{7}$$

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?

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

$$\Omega = \{\omega = (\omega_1, ..., \omega_n) : \omega_i \in \mathcal{F} \text{ for all } i \in \mathcal{N}\} = \mathcal{F}^n \tag{8}$$

with cardinality

$$|\Omega| = |\mathcal{F}^n| = |\mathcal{F}|^n = f^n \tag{9}$$

The event we're looking at is

$$A = \left\{\omega \in \Omega : \omega_i \neq \omega_j \text{ for all } i \neq j \in \mathcal{N}\right\} \tag{10}$$

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

$$\begin{aligned} |A| &= \#\text{ of ways to choose and order } n \text{ elements out of } f \\ &= \underbrace{f(f-1)...}_{n} = f(f-1)...(f-n+1) = \frac{f!}{(f-n)!} \end{aligned} \tag{11}$$

$$P(A) = \frac{f!}{f^n(f-n)!} \tag{12}$$

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

$$\begin{aligned} P(A) &= 1 - P\left(\underbrace{\text{no two people have the same birthday}}_{\text{Ex. 2}}\right) \\ &= 1 - \frac{y!}{y^n(y-n)!} \end{aligned} \tag{13}$$

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

$$
\begin{aligned}
\Omega &= \text{distributions of possible birthdays of } n \text{ people} \\
&= \{\omega = (\omega_1, \ldots, \omega_n) : \omega_i \in \mathcal{Y} \text{ for all } i \in \mathcal{N}\} = \mathcal{Y}^n
\end{aligned} \tag{14}
$$

where ω_i is the birthday of the i-th person. The cardinality of the jet set is

$$
|\Omega| = |\mathcal{Y}^n| = |\mathcal{Y}|^n = y^n \tag{15}
$$

The event we're looking at is

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

Note that this is the complementary event to the event defined in Equation 10 of Exercise 2. Thus we can compute its probability as

$$
P(A) = 1 - P(A^c) \tag{17}
$$

in agreement with Equation 13.

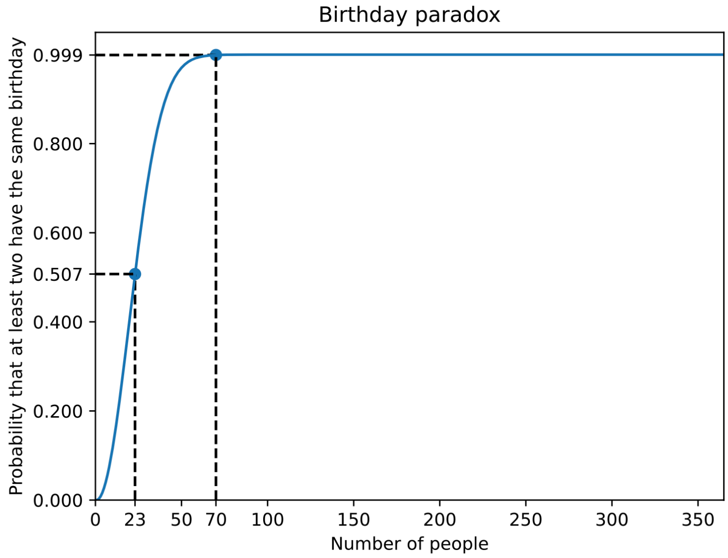

Figure 1: Birthday paradox probability. Code available.

> **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

$$
\begin{aligned}
P(A) &= 1 - P\Big(\underbrace{\text{nobody has the prescribed birth date}}\Big) \\
&= 1 - \left(\frac{y-1}{y}\right)^n
\end{aligned} \tag{18}
$$

Formal solution 1 As above $\mathcal{N} = \{1, \ldots, n\}$ and $\mathcal{Y} = \{1, \ldots, y\}$ with $n \leq y$. The jet set 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

$$
A = \{\omega \in \Omega : \exists i \in \mathcal{N} : \omega_i = \omega_{n+1}\} \tag{19}
$$

with complementary event

$$A^c = \{\omega \in \Omega : \omega_i \neq \omega_{n+1} \forall i \in \mathcal{N}\} \tag{20}$$

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

$$|A^c| = \underbrace{y}_{\text{prof}} \cdot \underbrace{(y-1)^n}_{\text{students}} \tag{21}$$

So, $P(A) = 1 - \frac{y(y-1)^n}{y^{n+1}} = 1 - \left(\frac{y-1}{y}\right)^n$, in agreement with Equation 18.

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

$$A = \sqcup_{j \in \mathcal{N}} A_j \tag{22}$$

with probability (cf Equation 2)

$$P(A) = \sum_{j \in \mathcal{N}} P(A_j) = \frac{\sum_{j \in \mathcal{N}} |A_j|}{|\Omega|} \tag{23}$$

The cardinality of A_j is

$$\begin{aligned} |A_j| &= \underbrace{1...1}_{j \text{ times}} \cdot \underbrace{(y-1)...(y-1)}_{n-j \text{ times}} \cdot \underbrace{y}_{\text{prof}} \cdot \underbrace{\binom{n}{j}}_{\text{number of ways to choose } j \text{ elements out of } n} \\ &= y(y-1)^{n-j}\binom{n}{j} \end{aligned} \tag{24}$$

By an application of the binomial theorem (Equation 3) and a short manipulation,

$$\sum_{j=1}^{n} |A_j| = y(y^n - (y-1)^n) \tag{25}$$

which leads back to Equation 18.

2. Theory recap 18.9.24

2.1. Conditional probability

- *Conditional probability*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \text{ if } P(B) \neq 0 \tag{26}$$

- not reallty defined if $P(B) = 0$, cf [1] pag. 427.
- often used as

$$P(A \cap B) = P(A \mid B)P(B) \tag{27}$$

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

$$P(A \mid B) + P(A^c \mid B) = 1 \tag{28}$$

- *Total probability theorem*

$$P(A) = P(A \mid B)P(B) + P(A \mid B^c)P(B^c) \tag{29}$$

- *Bayes theorem*

$$P(A \cap B) = P(B \cap A) \Rightarrow P(A \mid B)P(B) = P(B \mid A)P(A) \tag{30}$$

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

2.2. Independent events

Let Ω be equipped with a probability P.

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

$$P(A \cap B) = P(A)P(B) \tag{31}$$

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

$$P(\cap_{i \in I} A_i) = \prod_{i \in I} P(A_i) \text{ for all } I \subseteq \{1, ..., n\} \tag{32}$$

- 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:

- $A =$ 1ere piece donne P
- $B =$ 2eme piece donne F
- $C =$ 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 2: 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,

$$P(A_2 \cap A_1) = \underbrace{P(A_2 \mid A_1)}_{\frac{5}{9}} \underbrace{P(A_1)}_{\frac{6}{10}} = \frac{1}{3}. \tag{33}$$

Solution 2 The jet set 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

$$P(A) = \frac{\binom{6}{2}}{\binom{10}{2}} = \frac{6 \cdot 5}{10 \cdot 9} = \frac{1}{3}. \tag{34}$$

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 The cards are RR, RW, WW with W for white and R for red. Call RR the "red" card, WW the "white" card, and WR the "mixed" card. 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 The jet set contains the possible outcomes of a sequence of two events: 1. draw a card (out of three), and 2. observe a face (out of two). 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

$$\Omega = \{R_h \cap R_o, R_h \cap W_o, W_h \cap R_o, W_h \cap W_o\} \tag{35}$$

where the first entry indicates the hidden face, and the second entry indicates the observed face. For example, $W_h \cap R_o$ is the event "*the hidden face is white and the observed face is red*", and similarly for the others.

In this formulation, every element in the jet set is the intersection of two (dependent) events of the type 1. a face is hidden, and 2. a face is observed. Note that the event $W_h \cap R_o$ is equivalent to the event "*the mixed card is drawn, and the red face is observed.*" Under this second point of view, each outcome in Ω is the intersection of two (dependent) events of the type 1. a card is drawn, and 2. a face is observed. Denoting the event *"draw the red card"* by r, the event *"draw the white card"* by w, and the event *"draw the mixed card"* by m, the jet set is equivalently

$$\Omega = \{r \cap R_o, m \cap W_o, m \cap R_o, w \cap W_o\} \tag{36}$$

This formulation helps to understand that the probability on Ω **is not uniform**. The probabilities of the events in Ω are computed by Equation 27:

$$P(R_h \cap R_o) = P(r \cap R_o) = \frac{P(r \mid R_o)}{R_o} \tag{37}$$

However, we do not know the probabilities on the right hand side. As a simple trick, remember that $P(A \cap B) = P(B \cap A)$, so we can turn this around:

$$\begin{aligned} P(R_h \cap R_o) &= P(R_o \cap r) \\ &= \underbrace{P(R_o \mid r)}_{1} \underbrace{P(r)}_{\frac{1}{3}} = \frac{2}{6} \end{aligned} \tag{38}$$

$$
\begin{aligned}
P(R_h \cap W_o) &= P(W_o \cap m) \\
&= \underbrace{P(W_o \mid m)}_{\frac{1}{2}} \underbrace{P(m)}_{\frac{1}{3}} = \frac{1}{6}
\end{aligned} \tag{39}
$$

$$
\begin{aligned}
P(W_h \cap R_o) &= P(R_o \cap m) \\
&= \underbrace{P(R_o \mid m)}_{\frac{1}{2}} \underbrace{P(m)}_{\frac{1}{3}} = \frac{1}{6}
\end{aligned} \tag{40}
$$

$$
\begin{aligned}
P(W_h \cap W_o) &= P(W_o \cap w) \\
&= \underbrace{P(W_o \mid w)}_{1} \underbrace{P(w)}_{\frac{1}{3}} = \frac{2}{6}
\end{aligned} \tag{41}
$$

Now by Equation 26 and using these probabilities,

$$
\begin{aligned}
P(W_h \mid R_o) &= \frac{P(W_h \cap R_o)}{P(R_o)} \\
&= \frac{P(W_h \cap R_o)}{P(R_h \cap R_o) + P(W_h \cap R_o)} = \frac{1}{3}
\end{aligned} \tag{42}
$$

$$
\begin{aligned}
P(R_h \mid R_o) &= \frac{P(R_h \cap R_o)}{P(R_o)} \\
&= \frac{P(R_h \cap R_o)}{P(R_h \cap R_o) + P(W_h \cap R_o)} = \frac{2}{3} \\
&= 1 - P(W_h \mid R_o)
\end{aligned} \tag{43}
$$

where the last line follows from Equation 28 and gives directly the answer. So in conclusion, given the fact that we observe a red face, the hidden face is also red with probability $2/3$.

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 $x + 1 \bmod 6$:

$$
\Omega = \{12, 23, 34, 45, 56, 61\} \tag{44}
$$

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

$$
\mathrm{die}_1 = \{12, 61\} \tag{45}
$$

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

$$
P(\mathrm{die}_2 \mid \mathrm{live}_1) = \frac{P(\mathrm{die}_2 \cap \mathrm{live}_1)}{P(\mathrm{live}_1)} \tag{46}
$$

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

3. Theory recap 26.9.24

3.1. Probability measure

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

Definition 3.1.1 *(sigma-field)* Let Ω be any set. A *σ-field* $\mathcal{A}$ on Ω is a collection of subsets of Ω 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 Ω that is in $\mathcal{A}$ is called *event*.

Definition 3.1.2 *(Measure)* Given a set Ω and a σ-algebra $\mathcal{A}$ on Ω, a *measure* μ is a function

$$\mu : \mathcal{A} \to \mathbb{R}_{\geq 0} \tag{47}$$

such that

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

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i). \tag{48}$$

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

$$P(\Omega) = 1. \tag{49}$$

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

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

3.2. Discrete random variables

Definition 3.2.1 *(Discrete random variable)* Given a probability space $(\Omega, \mathcal{A}, P)$, a *discrete random variable* X is a function $X : \Omega \to F$ such that

1. F is a *countable* set;
2. *the level sets of X are events*, that is

$$\{X = x\} := \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{A} \text{ for all } x \in F \tag{50}$$

- clearly, $\{X = x\} = \emptyset \in \mathcal{A}$ for all $x \in F \setminus \operatorname{Im}(X)$

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

- the second property guarantees that $P\{X = x\}$ 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 $\mu : F \to [0, 1]$ defined by

$$\mu(x) = P\{X = x\} \text{ for all } x \in F. \tag{51}$$

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

$$P\{X = x, Y = y\} = P\{X = x\}P\{Y = y\} \text{ for all } x \in F, y \in G \tag{52}$$

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

$$\{\omega \in \Omega : X(\omega) = x\} \cap \{\omega \in \Omega : Y(\omega) = y\} \in \mathcal{A}. \tag{53}$$

- the definition generalises to collections of DRVs, see Section 2.2.3 in [2].

Take-away A discrete random variable is a *function on* Ω with *countable range*. Think of it as an experiment with a random outcome. Its *law, or distribution*, gives the probability to observe each of the possible (countable) *values* of the random variable.

Take-away
- $(\Omega, \mathcal{A}, P)$ with $P : \mathcal{A} \to [0, 1]$ and $P(\Omega) = 1$
- $X : \Omega \to F$ countable, with $\{X = x\} \in \mathcal{A}$ for all $x \in F$
- $\mu : F \to [0, 1]$ such that $\mu(x) = P\{X = x\}$

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 $p \in [0, 1]$, and we write $X \sim \mathcal{B}(p)$.

$$\begin{aligned} \mu : F \to\ & [0, 1] \\ 1 \longmapsto\ & p \\ 0 \longmapsto\ & 1 - p \\ x \longmapsto\ & p^x (1 - p)^{1-x} \end{aligned} \tag{54}$$

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

$$\begin{aligned} \mu : F \to\ & [0, 1] \\ k \longmapsto\ & \binom{n}{k} p^k (1 - p)^{n-k} \end{aligned} \tag{55}$$

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

$$
\begin{aligned}
\mu : F &\to [0,1] \\
n &\longmapsto e^{-\lambda}\frac{\lambda^n}{n!}
\end{aligned}
\tag{56}
$$

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

$$
\begin{aligned}
\mu : F &\to [0,1] \\
k &\longmapsto p(1-p)^{k-1}
\end{aligned}
\tag{57}
$$

3.4. Useful stuff

- Vandermonde's identity

$$
\sum_{k_1=0}^{k} \binom{n_1}{k_1}\binom{n_2}{k-k_1} = \binom{n_1+n_2}{k}
\tag{58}
$$

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 [2] pag. 16.

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

$$
\mu_i(k_i) = P(X_i = k_i) = \binom{n_i}{k_i} p^{k_i}(1-p)^{n_i-k_i}
\tag{59}
$$

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

$$
\begin{aligned}
\mu(k) &= P(X_1 + X_2 = k) \\
&= P\left(\bigsqcup_{\substack{k_i \in F_i \\ k_1+k_2=k}} \{X_1 = k_1, X_2 = k_2\}\right) \\
&= \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} P(X_1 = k_1)P(X_2 = k_2) && \text{by c. add and indep.} \\
&= \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \mu(k_1)\mu(k_2) && \text{by def of law} \\
&= \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \binom{n_1}{k_1}\binom{n_2}{k_2} p^{k_1+k_2}(1-p)^{n_1+n_2-k_1-k_2} \\
&= p^k(1-p)^{n_1+n_2-k} \sum_{\substack{k_i \in F_i \\ k_1+k_2=k}} \binom{n_1}{k_1}\binom{n_2}{k_2}
\end{aligned}
\tag{60}
$$

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 ≥ 0, k_1 can be at most equal to k. So the constraints

$$
\begin{aligned}
&k_1 \in \{0, ..., n_1\} \\
&k_2 \in \{0, ..., n_2\} \\
&k_1 + k_2 = k
\end{aligned}
\tag{61}
$$

can be replaced by the constraints

$$
\begin{aligned}
&k_1 \in \{0, ..., k\} \\
&k_2 = k - k_1
\end{aligned}
\tag{62}
$$

namely

$$
\sum_{\substack{k_i \in F_i \\ k_1 + k_2 = k}} \binom{n_1}{k_1}\binom{n_2}{k_2} = \sum_{k_1=0}^{k} \binom{n_1}{k_1}\binom{n_2}{k - k_1} = \binom{n_1 + n_2}{k} \tag{63}
$$

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 63.
- If $k > n_1$, we have

$$
\begin{aligned}
\sum_{\substack{k_i \in F_i \\ k_1 + k_2 = k}} \binom{n_1}{k_1}\binom{n_2}{k_2} &= \sum_{k_1=0}^{n_1} \binom{n_1}{k_1}\binom{n_2}{k - k_1} \\
&= \sum_{k_1=0}^{n_1} \binom{n_1}{k_1}\binom{n_2}{k - k_1} + \sum_{k_1=n_1+1}^{k} \binom{n_1}{k_1}\binom{n_2}{k - k_1}
\end{aligned}
\tag{64}
$$

since each summand in the the second sum is zero[2], and we get again Equation 63.

So in conclusion

$$
\mu(k) = \binom{n_1 + n_2}{k} p^k (1 - p)^{n_1 + n_2 - k} \tag{65}
$$

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

$$
\begin{aligned}
\mu(n) = P(X_1 + X_2 = n) &= \sum_{\substack{n_i \in \mathbb{N} \\ n_1 + n_2 = n}} \mu_1(n_1)\mu_2(n_2) \\
&= \sum_{\substack{n_i \in \mathbb{N} \\ n_1 + n_2 = n}} e^{-\lambda_1} \frac{\lambda_1^{n_1}}{n_1!} e^{-\lambda_2} \frac{\lambda_2^{n_2}}{n_2!}
\end{aligned}
\tag{66}
$$

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 $b > a$.

$$\begin{aligned}\mu(n) &= e^{-(\lambda_1+\lambda_2)} \sum_{n_1=0}^{n} \frac{\lambda_1^{n_1}}{n_1!} \frac{\lambda_2^{n-n_1}}{n-n_1!} \cdot \frac{n!}{n!} \\ &= \frac{e^{-(\lambda_1+\lambda_2)}}{n!} \sum_{n_1=0}^{n} \binom{n}{n_1} \lambda_1^{n_1} \lambda_2^{n-n_1} \\ &= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!}\end{aligned} \tag{67}$$

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.

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

$$\begin{aligned}\mu(k) &= P(Z = k) \\ &= P(Z \geq k) - P(Z \geq k+1) \\ &= P(X_1 \geq k, X_2 \geq k) - P(X_1 \geq k+1, X_2 \geq k+1) \\ &= P(X_1 \geq k)P(X_2 \geq k) - P(X_1 \geq k+1)P(X_2 \geq k+1)\end{aligned} \tag{68}$$

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

$$\begin{aligned}P(X \geq k) &= P\left(\bigsqcup_{x \geq k} (X = i)\right) \\ &= \sum_{i \geq k} P(X = i) \\ &= \sum_{i \geq k} p(1-p)^{i-1} \\ &= p(1-p)^{k-1} + p(1-p)^k + p(1-p)^{k+1} + \ldots \\ &= p(1-p)^{k-1}\left(1 + (1-p) + (1-p)^2 + \ldots\right) \\ &= p(1-p)^{k-1} \sum_{j=0}^{\infty} (1-p)^j \\ &= (1-p)^{k-1}\end{aligned} \tag{69}$$

Plugging in Equation 68 we get

$$\begin{aligned}\mu(k) &= P(X_1 \geq k)P(X_2 \geq k) - P(X_1 \geq k+1)P(X_2 \geq k+1) \\ &= (1-p_1)^{k-1}(1-p_2)^{k-1} - (1-p_1)^k(1-p_2)^k \\ &= (1-p_1)^{k-1}(1-p_2)^{k-1}[1 - (1-p_1)(1-p_2)]\end{aligned} \tag{70}$$

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. $Z \sim \mathcal{G}(\beta)$.

4. Theory recap 2.10.24

Recall from last week that

Take-away
- $(\Omega, \mathcal{A}, P)$ with $P : \mathcal{A} \to [0, 1]$ and $P(\Omega) = 1$
- $X : \Omega \to F$ countable, with $\{X = x\} \in \mathcal{A}$ for all $x \in F$
- $\mu : F \to [0, 1]$ such that $\mu(x) = P\{X = x\}$

4.1. Expected value of a discrete random variable

In this section when we say *"X is a RV"* we mean "*$X : \Omega \to F \subset \mathbb{R}$ is a discrete random variable with real values.*"

Definition 4.1.1 *(Expected value)* A RV is *integrable* if $\sum_{x \in F} |x| P(X = x) < +\infty$, and in this case its *expected value* $\mathbb{E}(X)$ is the real number

$$\mathbb{E}(X) := \sum_{x \in F} x P(X = x). \tag{71}$$

Proposition 4.1.2 *(Linearity of expectation)*

$$\mathbb{E}(X + aY) = \mathbb{E}(X) + a\mathbb{E}(Y) \tag{72}$$

Some **properties** of the expected value of a RV:

- $\mathbb{E}(\text{constant}) = \text{constant}$
- Sufficient condition, positivity, monotonicity: see [2] pag. 20.

Some **examples**:
- $X \sim \mathcal{B}(p) \Rightarrow \mathbb{E}(X) = p$
- $X \sim \mathcal{B}(n, p) \Rightarrow \mathbb{E}(X) = np$ (immediate by linearity from the above)
- $X \sim \mathcal{P}(\lambda) \Rightarrow \mathbb{E}(X) = \lambda$
- $X \sim \mathcal{G}(p) \Rightarrow \mathbb{E}(X) = \frac{1}{p}$

Theorem 4.1.3 *(Expectation of function)* Let $X : \Omega \to F \subset \mathbb{R}$ be RV, and consider some function $f : F \to \mathbb{R}$. Then

$$\mathbb{E}(f(X)) = \sum_{x \in F} f(x) P(X = x) \tag{73}$$

whenever defined (see [2] Th. 2.3.6 for details).

Proposition 4.1.4 *(Expectation and independence)* Let X, Y be RV and f, g two functions on their values such that all the expectations are well-defined (i.e. al the random variables are integrable). Then

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)) \tag{74}$$

4.2. Variance of a discrete random variable

Definition 4.2.1 *(Variance)* A RV X is called *square integrable* if X^2 is integrable, that is if $\sum_{x \in F} x^2 P(X = x) < +\infty$, and in this case

$$\mathrm{Var}(X) := \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \tag{75}$$

- If X is square integrable then the variance is well defined, cf [2] Remark 2.3.11

- The variance is a measure of the spreading, dispersion, of a random variable around its expected value

The two following properties of the variance are very useful for concrete calculations:

Lemma 4.2.2 *(Variance as difference of expectations)*

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \tag{76}$$

The variance is in general **not linear**:

Lemma 4.2.3 *(Variance after scaling and shifting)*

$$\mathrm{Var}(aX + b) = a^2 \ \mathrm{Var}(X) \text{ for all } a, b \in \mathbb{R} \tag{77}$$

Proof Exercise; both lemmas follow by linearity of the expectation. □

Proposition 4.2.4 *(Variance and independence)* Let $(X_i)_{i \in \{1,\dots,n\}}$ be a family of square integrable random variables. Then their sum is square integrable, and **if the X_i are independent** then

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) \tag{78}$$

Some **examples**:
- $X \sim \mathcal{B}(p) \Rightarrow \mathrm{Var}(X) = p(1-p)$, see Exercise 4.3 and proof.
- $X \sim \mathcal{B}(n, p) \Rightarrow \mathrm{Var}(X) = np(1-p)$ (immediate by Proposition 4.2.4)
- $X \sim \mathcal{P}(\lambda) \Rightarrow \mathrm{Var}(X) = \lambda$, see proof.

Exercises

Exercise 4.1 *(Lost messages)* On a telecommunication channel, it has been estimated that in T time units there arrives a number of messages that can be estimated by a DRV $\sim \mathcal{P}(\lambda T)$. Each message has a loss probability equal to p, independent of the other messages. Find the probability that the number of lost message in T units of time is equal to l.

Without loss of generality rescale $\lambda \leftarrow \lambda T$. We need to find the discrete random variable L whose range $\{0, 1, 2, \dots\} \ni l$ contains the possible numbers l of lost messages in one time unit. The probability $P(L = l)$ to lose l message is then by definition by the law of L.

Let X_i be the DRV for the event "*the i-th message is lost*". Since each message is lost with probability p, $X_i \sim \mathcal{B}(p)$ for all $i \in \{1, 2, \dots\}$.

Let $L_a = \sum_{i=1}^{a} X_i$ be the DRV whose range $\mathrm{Im}(L_a) = \{0, 1, \dots, a\} \ni l$ contains the numbers l of possible lost messages out of a arrived. Since L_a is the sum of a independent p-Bernoulli DRVs, L_a follows the binomial distribution:

$$L_a \sim \mathcal{B}(a, p). \tag{79}$$

Finally, let A be the DRV estimating the number of arrived messages $a \in \{0, 1, \dots, \}$ in one time unit; we are given that $A \sim \mathcal{P}(\lambda)$.

The law of L is given by

$$P(L = l) = P\left[\bigsqcup_{a=l}^{\infty} \{L_a = l \cap A = a\}\right], \tag{80}$$

that is, we look at the disjoint union of all the events in which, given a arrived messages, l are lost. By countable additivity and independence,

$$P(L = l) = \sum_{a=l}^{\infty} P(L_a = l) P(A = a) \tag{81}$$

Now L_a follows a binomial distribution and A follows a Poisson distribution, so

$$\begin{aligned} P(L = l) &= \sum_{a=l}^{\infty} \binom{a}{l} p^l (1-p)^{a-l} e^{-\lambda} \frac{\lambda^a}{a!} \cdot \frac{\lambda^l}{\lambda^l} \\ &= \lambda^l p^l \frac{1}{l!} e^{-\lambda} \sum_{a=l}^{\infty} \frac{1}{(a-l)!} (1-p)^{a-l} \lambda^{a-l} \\ &= \lambda^l p^l \frac{1}{l!} e^{-\lambda} \sum_{j=0}^{\infty} \frac{(\lambda - \lambda p)^j}{j!} \\ &= \lambda^l p^l \frac{1}{l!} e^{-\lambda} e^{\lambda - \lambda p} \\ &= \frac{(\lambda p)^l}{l!} e^{-\lambda p}. \end{aligned} \tag{82}$$

So $V \sim \mathcal{P}(\lambda p)$.

Exercise 4.2 *(Poisson expectation)* Let $N \sim \mathcal{P}(\lambda)$. Find $\mathbb{E}\left(X := \frac{1}{N+1}\right)$.

By Theorem 4.1.3 we have

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} \frac{1}{n+1} e^{-\lambda} \frac{\lambda^n}{n!} \tag{83}$$

Multiply and divide by λ and shift the running index to get $\mathbb{E}(X) = \frac{1-e^{-\lambda}}{\lambda}$.

Exercise 4.3 *(Archery)* An archer shots n arrows at a target. The shots are independent, and each shot hits the target with probability p. Let X be the random variable "*number of times the target is hit*".

1. What is the law of X?
2. What is the expectation of X?
3. What is the value of p that maximises the variance of X?

The archer bets on his result. He gets g euros when he hits the target, and loses l euros when he misses it. Let Y be the random variable that represent the net gain of the archer at the end of the n shots.

4. What is the expectation of Y?
5. What is the relation between g and l that guarantees the archer an expected gain of zero?

1. X is the sum of n independent p-Bernoullli variables, hence $X \sim \mathcal{B}(n, p)$ (binomial distribution).
2. We have to compute the expectation of a binomial random variable $X = X_1 + \ldots + X_n$, where each X_i is a Bernoulli variable. Since expectations are linear we can compute the expectation of the Bernoulli variables, and sum them:

$$\mathbb{E}(X_i) = 1 \cdot p + 0 \cdot (1-p) = p \tag{84}$$

$$\mathbb{E}(X_1 + ... + X_n) = \mathbb{E}(X_1) + ... + \mathbb{E}(X_n) = np \tag{85}$$

For example, if $p = 0.5$ and $n = 10$, this means that the archer expects to hit the target 5 times.

3. Let's compute the variance of a Bernoulli and a binomial variable by Equation 76:

$$\mathbb{E}(X_i^2) = 1^2 \cdot p + 0^2 \cdot (1 - p) = p \tag{86}$$

$$\mathrm{Var}(X_i) = \mathbb{E}(X_i^2) - \mathbb{E}(X_i)^2 = p(1 - p) \tag{87}$$

By independence, $\mathrm{Var}(X_1 + ... + X_n) = \mathrm{Var}(X_1) + ... + \mathrm{Var}(X_n) = np(1 - p)$. To find the value of p that maximises the variance differentiate: $n(1 - 2p) = 0 \Rightarrow p = 0.5$.

4. Let $Y_i =$ gain of i-th shot. Then $\mathbb{E}(Y_i) = gp - l(1 - p)$, and

$$\mathbb{E}(Y = Y_1 + ... + Y_n) = n[gp - l(1 - p)] \tag{88}$$

For example if $n = 10$, $g = 1$, $l = 2$, we have $\mathbb{E}(Y) = 30p - 20$; and if furthermore $p = 0.5$ then $\mathbb{E}(Y) = -5$.

5. To find the value of relation between g and l required to have an expected gain of zero solve the equation $\mathbb{E}(Y) = 0$ to get

$$\frac{g}{l} = \frac{1 - p}{p}. \tag{89}$$

Thus as the probability p to hit the target goes to zero, a very big $\frac{g}{l}$ is required to guarantee an expected gain of zero; viceversa $\frac{g}{l}$ becomes infinitely small as $p \to 1$. At $p = 0.5$, as one would expect, $g = l$.

Exercise 4.4 *(Smart betting?)* Let $X_n \sim \mathcal{B}(p = \frac{1}{2})$, $n = 0, 1, 2, ...$ be independent Bernoulli variables. You bet $b_n = 2^{n-1}$ on the outcome of X_n if $X_0 = X_1 = X_2 = ... = X_{n-1}$, and zero else. If you bet and you win, tou receive $g_n = 2b_n$.
1. What is the expected amount of money bet?
2. What is the expected gain?

1. Let's write down explicitely the process:

- X_0 is drawn $\to$ get nothing, and bet b_1 on X_1.
- X_1 is drawn $\to$ get g_1 with probability p. If $X_1 = X_0$, bet b_2 on X_2, else stop the game.
- X_2 is drawn $\to$ get g_2 with probability p. If $X_2 = X_0$, bet b_3 on X_3, else stop the game.
- ...
- X_{n-1} is drawn $\to$ get g_{n-1} with probability p. If $X_{n-1} = X_0$, bet b_n on X_n, else stop the game.
- ...

Then the expected amount of money bet M is

$$\begin{aligned}
\mathbb{E}(M) = b_1 \quad & P(X_1 \neq X_0) + \\
+(b_1 + b_2) \quad & P(X_1 = X_0, X_2 \neq X_0) + \\
+(b_1 + b_2 + b_3) & P(X_1 = X_0, X_2 = X_0, X_3 \neq X_0) + \\
+ ... + & \\
+(b_1 + ... + b_n) & P(X_1 = X_0, ..., X_{n-1} = X_0, X_n \neq X_0) + ...
\end{aligned} \tag{90}$$

Let $B_n = \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} 2^{k-1} = \sum_{l=0}^{n-1} 2^l = \frac{1-2^n}{1-2} = 2^n - 1$. Then

$$\mathbb{E}(M) = \sum_{n=1}^{\infty} B_n \cdot \frac{1}{2^n} = \sum_{n=1}^{\infty} \frac{2^n - 1}{2^n} = +\infty. \tag{91}$$

The idea is that the expected bet amount diverges even if one stops betting at some point.

2. *This is for the standard doubling scenario, without the constraint bet iff all previous outcomes are the same.*

After n bets the bet amount is $B_n = \sum_{k=1}^{n} b_k = 2^n - 1$. If I win at the $(n+1)$-th round the net win equals the initial bet:

$$2b_n - B_n = 2^n - (2^n - 1) = +1. \tag{92}$$

Now the probability to lose n times and win at the $(n+1)$-th time is $(1-p)^n p$, so the expected net gain is also equal to the initial bet:

$$\mathbb{E}(\text{Net gain}) = \sum_{n=0}^{\infty} (1-p)^n p \cdot (+1) = \frac{p}{p} = 1. \tag{93}$$

For more details: Martingale betting_system.

Bibliography

[1] P. Billingsley, *Probability and Measure*. John Wiley & Sons, 2012.

[2] B. Jourdain, *Probabilités et statistique pour l'ingénieur*. 2018.

[3] E. H. Lieb and M. Loss, *Analysis*, 2nd ed. in Graduate Studies in Mathematics 14. American Mathematical Society, 2001.