Probabilités

Cours donnée au profit d'étudiants

ingénieurs en 2ème année

Notes de cours

Cours en probabilités, Université Abdelmalek Essaâdi

https ://hamzaelmahjour.github.io

Ces petites notes de cours ne sont pas les meilleures que tu peux trouver sur internet mais elles ont été présentées avec amour et passion aux étudiants exceptionnels que j'ai eu pendant cette année.

Première édition, Mars 2023

Table des matières

1. Notions essentielles

1.1 Introduction

Beaucoup de notre vie est basé sur l'imprévu et l'improbable, heureusement ! Car, la vie va perdre ainsi une grande partie de sa beauté au cas où l'incertain n'existait pas et si tout ce qui se passait était prévisible certainement. Les gens utilisent des termes vaguement comme "probable" ou "aléatoire" et le but de la théorie des probabilités est de définir et bien structurer ces termes par une construction mathématique adéquate.

1.2 Espace d'événements

Dans la vie quotidienne on entend souvent : la probabilité qu'**il fasse beau demain** est **très haute** ou bien il y **une chance sur deux** que cette **opération réussisse**. Donc on annonce un fait : pluie, soleil, réussite, défaite ... et on lui associe un adjectif ou un nombre pour *mesurer* son degré de faisabilité : 40% de chance, fort probable, peu probable etc. La réalisation ou la non réalisation d'un "événement A" dépend d'une chaîne de circonstances. Par exemple, la pluie a une grande probabilité de tomber s'il y a une certaine vitesse du vent, avec une concentration des nuages, une température donnée, une pression atmosphérique et d'autres paramètres. Cette chaîne de faits s'appelle "une expérience". La donnée d'une expérience est dite un résultat. En général, dans un cadre aléatoire correct, on ne peut pas prédire "certainement" l'éventuel réalisation d'une donnée, mais on peut lister tout de même l'ensemble de toutes les issues possible. Si on prend l'exemple du lancement [1] d'un dé, on ne peut pas savoir le résultat du lancer à l'avance mais on peut lister l'ensemble des résultats possibles : { ⚀, ⚁, ⚂, ⚃, ⚄, ⚅ }. Chaque face supérieur affichée par ce dé est un "résultat". Le rôle de la théorie des probabilités est non pas seulement de dénombrer les résultats mais aussi de mesurer la chance qu'un résultat ou plutôt une famille de résultats se réalise.

Définition 1.2.1 Dans une expérience en probabilités, l'ensemble de tous ses résultas possible s'appelle un **univers** (*sample space*). Souvent on le note Ω.

---

1. Et cet exemple va nous hanter tout au long de ces chapitres.

■ **Exemple 1.1** Une pièce de monnaie est jetée. il y a deux résultats possibles : Pile := P ou Face := F. On va s'intéresse un peu plus loin aux événements suivants :

(a) A_1 = Le résultat est F
(b) A_2 = le résultat est soit P ou F
(c) A_3 = Le résultat est F et P
(d) A_4 = le résultat n'est pas F

L'univers de cette expérience est $\Omega = \{P,F\}$. ■

■ **Exemple 1.2** On reprend l'exemple du dé encore une fois. L'univers de l'expérience est $\Omega =$ {⚀,⚁,⚂,⚃,⚄,⚅} $= \{1,2,3,4,5,6\}$. Voici quelques événements à noter

(a) B_1 = le nombre affiché est 1
(b) B_2 = le résultat affiché est paire
(c) B_3 = le résultat est paire et inférieur à 3
(d) B_4 = le résultat n'est pas paire.

■

Si on revient maintenant aux événements des exemples 1.1 et 1.2, on remarque qu'on peut reformuler les A_i et les B_i. En effet,

- $A_1 = \{F\}$
- $A_2 = \{F\} \cup \{P\}$
- $A_3 = \{F\} \cap \{P\}$
- $A_4 = \{F\}^C$
- $B_1 = \{1\}$
- $B_2 = \{2,4,6\}$
- $B_3 = \{1,2,3\} \cap \{2,4,6\}$
- $B_4 = \{2,4,6\}^C$

où la notation A^C se lit comme le **complémentaire** de A. Nous constatons que les **événements** sont des sous-ensembles de l'univers Ω.

Définition 1.2.2 On dit que A est un événement si $A \in \mathscr{P}(\Omega)$.

Toujours en revenant à la manipulation des exemples 1.1 et 1.2, on peut remarquer que nous avons besoin que les événements "vivent" dans un endroit où il existe une stabilité lors de l'utilisation des opérateurs $\cup$, $\cap$ et par passage au complémentaire.

R - $\emptyset$ est appelé : événement impossible comme le cas de l'événement A_3. Il est impossible d'obtenir Pile et Face au même temps.
- Ω est appelé : l'événement certain (sûr) car forcément un résultat au moins va se réaliser parmi tous les résultats possibles. C'est le cas de l'événement A_2.

1.3 Comment mesurer la probabilité des événements ?

Approche empirique

Supposons qu'on répète une expérience N fois pour N assez grand, en essayant de préserver les conditions initiales. Supposez que A est un événement qui peut se produire (ou pas). Dans la plupart des expérimentations scientifiques, la proportion des fois où A va se réaliser se stabilisera à une certaine valeur quand N est très grand. Notons N_A le nombre de fois où A se réalise.

$$\frac{N_A}{N} \longrightarrow_{N\to\infty} \text{constante} := \mathbb{P}(A)$$

Cette constante est un bon "candidat" pour mesurer la probabilité de A. C'est une sorte de mariage entre les statistiques et les probabilités ou plutôt une justification statistique à l'intuition. Ce $\mathbb{P}(A)$

est clairement entre 0 et 1. Si $A = \emptyset$ et $N_\emptyset = 0$ donc $\mathbb{P}(\emptyset) = 0$. Si $A = \Omega$ alors $\mathbb{P}(\Omega) = 0$ si $A = \Omega$ alors $N(\Omega) = N$ donc $\mathbb{P}(\Omega) = \frac{N}{N} = 1$. Cette approche que nous avons utilisé est pour montrer qu'il existe une cohérence entre ce que prétend la théorie des probabilités et les expériences réelles.

Revenons aux événements...

Comme nous l'avons cité précédemment, nous avons besoin d'une structure ou interagissent les événements. Cette structure n'est pas forcément l'ensemble de toutes les parties de Ω et il peut être défini de la façon suivante.

Définition 1.3.1 — Tribu. Soit Ω un univers et soit $\mathscr{F} \subset \Omega$. On dit que $\mathscr{F}$ est une tribu (σ-algèbre) si

(i) $\emptyset \in \mathscr{F}$
(ii) Si $A \in \mathscr{F}$ alors $A^C \in \mathscr{F}$
(iii) Si (A_n) avec $n \in \mathbb{N}^*$ est une famille dénombrable d'événements alors $\bigcup_{n\in\mathbb{N}^*} A_n \in \mathscr{F}$

On peut citer, à titre d'exemple quelques tribus usuelles :

1. **Grossière** : $\mathscr{F} = \{\emptyset, \Omega\}$
2. **Exhaustive** : $\mathscr{F} = \mathscr{P}(\Omega)$
3. Prenons $\Omega = \{1,2,3,4\}$, on peut vérifier de manière simple que $\mathscr{F} = \{\emptyset, \Omega, \{1\}, \{2,3,4\}\}$ représente bien une tribu.
4. En général, si $A \subseteq \Omega$ alors $\mathscr{F} = \{\emptyset, \Omega, A, A^C\}$ est toujours une tribu.
5. **Borélienne** Bien que ce cours n'est pas concerné par les tribus sur les ensembles non dénombrables comme $\mathbb{R}$, il est utile de mentionner que $\mathscr{B}(\mathbb{R})$ est la tribu engendré par les "intervalles" de $\mathbb{R}$. Bien que nous n'allons pas le démontrer ni s'en servir de façon directe mais la tribu borélienne sera utile dans la construction des variables aléatoires.

On peut avoir des doutes sur l'utilité des unions ou intersections infinies d'événements dans la résolution de problèmes en probabilité mais ce doute se dissipera rapidement avec cet exemple.

Problem 1.1 Adil et Omar jouent au lancement d'un dé. On suppose que c'est Adil qui commence. La règle du jeu est très simple : celui qui obtient "cinq" en premier gagne !
Considérer les trois événements suivants : $A = \{\text{Adil gagne}\}$, $B = \{\text{Omar gagne}\}$ et $D = \{\text{aucun vainqueur}\}$.
Contemplons les événements précédents, ils ont l'air simple mais à vrai dire c'est des événements qui cache beaucoup de complexité. En essayant d'exprimer ces événements, il s'avère que ça demandera un peu d'efforts. D'où la question : comment exprimer les événements précédents en utilisant des événements moins complexes et plus clair ?

Considérons les deux événements suivants

$$\begin{aligned} F_n &= \{\text{la partie se termine au } n-\text{ème lancer}\}, \qquad \forall n \in \mathbb{N}^*. \\ C_j &= \{\text{le résultat du } n-\text{ème lancer est "cinq}\}, \qquad \forall j \in \mathbb{N}^*. \end{aligned}$$

On essaiera d'exprimer les trois événements dans le problème 1.1 en fonction de F_n et C_j. L'événement D est le moins complexes des trois. En effet, pour ne pas avoir un vainqueur il faudra avoir un résultat différent de cinq à chaque lancer. Ça veut dire qu'à chaque lancer c'est le complémentaire de C_j qui est réalisé. De plus, pour que la partie ne se termine jamais, il faudra qu'elle continue jusqu'à l'infini d'où $D = \bigcap_{j\in\mathbb{N}^*} C_j^c$. C'est un peu marrant de voir que nous partons d'un jeu banal en inventant un événement banal qui est en réalité complexe et puis on fait appel à des événements simples qui vont l'exprimer de façon plus précise en passant par une intersection infinie ! C'est une démonstration clair de l'intérêt, voire le besoin d'utiliser les unions et intersections infinies. Les autres événements A et O sont traités dans les séries d'exercices.

Définition 1.3.2 Soit Ω un ensemble et $\mathscr{F}$ une tribu sur cet ensemble. On appelle $(\Omega, \mathscr{F})$ un **espace mesurable**.

1.4 Espace probabilisé

Définition 1.4.1 Soit $(\Omega, \mathscr{F})$ un espace mesurable et $\mathbb{P} : \mathscr{F} \longrightarrow [0,1]$ une fonction qui satisfait

(i) $\mathbb{P}(\emptyset) = 0$ et $\mathbb{P}(\Omega) = 1$

(ii) Si $(A_n)_{n\in\mathbb{N}^*}$ est une famille d'événements dans Ω deux à deux disjoints alors

$$\mathbb{P}(\bigcup_{n\in\mathbb{N}^*} A_i) = \sum \mathbb{P}(A_i)$$

On dit que $\mathbb{P}$ est une mesure de probabilité.

Le triplet $(\Omega, \mathscr{F}, \mathbb{P})$ s'appelle un **espace probabilisé**. On se donne les exemples suivants pour mieux comprendre ce que représente une **mesure de probabilité**.

■ **Exemple 1.3** (i) Dans le premier exemple, une pièce de monnaie est lancée une fois. L'ensemble des résultats est $\Omega = \{F, P\}$ et on peut lui associer la tribu exhaustive $\mathscr{F} = \{\Omega, \emptyset, F, P\}$. On considère que la $\mathbb{P}_p(F) = p$ et $\mathbb{P}_p(P) = 1 - p$ où p est une constante dans l'intervalle $[0,1]$. On peut bien vérifier que $\mathbb{P}_p$ définie ainsi représente bien une mesure de probabilité. Si $p = 1/2$ alors la pièce de monnaie n'est pas biaisée [2], sinon elle le sera.

(ii) Dans le deuxième exemple, on lance un dé une fois et les résultats possibles sont $\Omega = \{1, 2, \ldots, 6\}$. La tribu associée est $\mathscr{F} = \mathscr{P}(\Omega)$. On définit la mesure $\mathbb{P}$ telle que $\mathbb{P}(\{k\}) = p_k$ pour $k = 1, 2, \ldots, 6$. Si le dé n'est pas biaisée alors $p_k = 1/6$ pour tout k. Sinon, si au moins un p_k est différent de $1/6$ et si $\sum_{k=1}^{6} p_k = 1$ alors $\mathbb{P}$ sera bien une mesure de probabilité qui modélise un dé truquée (biaisée).

■

Lemme 1.4.1 Soit A, B des événements d'un espace $(\Omega, \mathscr{F}, \mathbb{P})$:

(a) $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$.

(b) Si $A \subseteq B$ alors $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$.

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

(d) Plus généralement,

$$\mathbb{P}(\bigcup_{i=1}^{n} A_i) = \sum_i \mathbb{P}(A_i) - \sum_{i,j} \mathbb{P}(A_i \cap A_j) + \sum_{i,j,k} \mathbb{P}(A_i \cap A_j \cap A_k) + \ldots + (-1)^{n+1} \mathbb{P}(A_1 \cap \ldots \cap A_n)$$

Démonstration.

(a) $A \cup A^C = \Omega$ disjoints donc $1 = \mathbb{P}(A \cup A^C) = \mathbb{P}(A)$.

(b) $B = A \cup (B \setminus A)$ on a $A \cap B \setminus A = \emptyset$ donc $\mathbb{P}(B) = \mathbb{P}(A) + \underbrace{\mathbb{P}(B \setminus A)}_{\geq 0}$ donc $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(c) On a

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup B \setminus A \cap B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A \cap B).$$

Puisque $A \cap B \subset B$, on utilisera (b) pour obtenir

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

2. (en anglais 'biased') c'est un mot dérivé de l'ancien occitan. "

(d) Trouver la formule pour trois éléments et utilisez une récurrence.

■

Lemme 1.4.2 1. Si une famille d'événements est croissante $\{A_i\}_{i\geq 1}$ sur le même espace probabilisé alors

$$A = \bigcup_{i=1}^{\infty} A_i = \lim_{i\to\infty} A_i$$

et donc $\mathbb{P}(A) = \lim_{i\to\infty} \mathbb{P}(A_i)$.

2. De manière similaire pour une famille décroissante $\{B_i\}_{i\geq 1}$ on a

$$B = \bigcap_{i=1}^{\infty} B_i = \lim_{i\to\infty} B_i$$

et donc $\mathbb{P}(B) = \lim_{i\to\infty} \mathbb{P}(B_i)$.

■ **Exemple 1.4** On considère l'espace $\Omega = [-1,1]$ muni de la tribu borélienne $\mathscr{B}([-1,1])$. On considère la mesure de probabilité $\mathbb{P}$ qui à chaque sous-intervalle $I =]a,b[$ de Ω on a $\mathbb{P}(]a,b[) = \frac{\text{longueur}(I)}{\text{longueur}(\Omega)} = \frac{b-a}{2}$. On prend la famille $A_n =]-\frac{1}{2}+\frac{1}{n}, \frac{1}{2}-\frac{1}{n}[$ pour $n \geq 3$. Il est clair que la famille $\{A_n\}$ est croissante, on a selon le lemme précédent que

$$\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \lim_{i\to\infty} \mathbb{P}(A_i) = \lim_{i\to\infty} \mathbb{P}\left(]-\frac{1}{2}+\frac{1}{n}, \frac{1}{2}-\frac{1}{n}[\right) = \mathbb{P}\left(\left]-\frac{1}{2}, \frac{1}{2}\right[\right) = \frac{1/2-(-1/2)}{2} = 1/2.$$

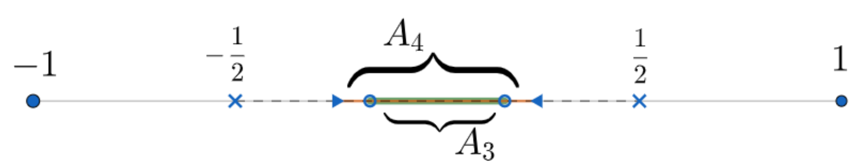


FIGURE 1.1 – La famille $\{A_i\}_{i\geq 3}$ croît vers $]-\frac{1}{2}, \frac{1}{2}[$.

1.5 Probabilité conditionnelle

Les propositions ou les déclarations suivantes : "Si B se réalise alors la probabilité que A se réalise est p" ont besoin d'être définis correctement. Pour cela nous allons emprunter à nouveau le cadre empirique utilisé dans la Section 1.3 c'est à dire on notera N_A et N_B le nombre de réalisations de deux événements A et B consécutivement sur une expérience répétée N fois. Cette fois-ci on s'intéresse aux seuls cas où B se réalise et on néglige les nombre de fois où B n'est pas réalisée. Plus précisément on cherche l'occurrence de A au cas où B se réalise. On est d'accord que dans cette situation la proportion des fois où A se réaliserait par rapport à B est $\frac{N_{A\cap B}}{N_B}$ et

$$\frac{N_{A\cap B}}{N_B} = \frac{N_{A\cap B}/N}{N_B/N} = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}.$$

On veut maintenant aboutir au même résultat en se basant sur une *intuition probabiliste*. Si en cherche la probabilité de A sachant B, on suppose implicitement que B s'est réalisé. Donc pour parler de la réalisation de A il faut absolument qu'il soit en synchronisation avec la réalisation de B. Autrement dit, A se réalise conditionnellement à B si et seulement si $A \cap B$ est accomplie c-à-d que la probabilité de A étant donné B est proportionnelle à $A \cap B$ donc $\mathbb{P}(A|B) = \alpha \mathbb{P}(A \cap B)$ où α est une constante qui dépend logiquement de B. Si on pose $A = \Omega$ On a $\mathbb{P}(\Omega|B) = \alpha \mathbb{P}(B)$. Or, quelle que soit le conditionnement appliquée à Ω cet événement va sûrement se réalisait i.e $\mathbb{P}(\Omega|B) = \mathbb{P}(\Omega) = 1$ donc $\alpha = \frac{1}{\mathbb{P}(B)}$. Finalement, on conclut que $\mathbb{P}(A|B) = \frac{1}{\mathbb{P}(B)} \mathbb{P}(A \cap B)$. Plus formellement

Définition 1.5.1 Si $\mathbb{P}(B) > 0$ alors la **probabilité conditionnelle** que A se réalise étant donné B est

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

■ **Exemple 1.5**

1. Une famille a deux enfants. Quelle est la probabilité que les deux soient des garçons sachant qu'au moins l'un des deux est garçon ? L'aîné et me benjamin peuvent être une fille ou un garçon. Il y a donc quatre résultats possibles $\Omega = \{FF, GF, FG, GG\}$ où F :=Fille et G := Garçon. On considère que nous avons une équiprobabilité entre les quatre résultats de Ω donc $\mathbb{P}(FF) = \mathbb{P}(GF) = \mathbb{P}(FG) = \mathbb{P}(GG) = 1/4$. On peut formuler les événements liées à la question comme suit

$$A = \{\text{avoir deux garçons}\} = \{GG\} \qquad B = \{\text{avoir au moins un garçon}\} = \{GG, FG, GF\}.$$

La question est de trouver $\mathbb{P}(A|B)$. On a

$$\mathbb{P}(A|B) = \mathbb{P}(GG|GF \cup FG \cup GG) = \frac{\mathbb{P}(GG \cap (GG \cup FG \cup GF))}{\mathbb{P}(GG \cup FG \cup GF)}.$$

Le lemme suivant est important en probabilités car il permet de définir un événements à partir d'une collection de probabilités conditionnelles.

Définition 1.5.2 On dit que $\{B_n\}_{n \geq 1}$ est une partition de l'espace Ω si $B_i \cap B_j = \emptyset$ pour tout $i \neq j$ et $\bigcup B_i = \Omega$

Lemme 1.5.1

(i) Soit A, B de Ω tel que $0 < \mathbb{P}(B) < 1$ alors

$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^C)\mathbb{P}(B^C)$$

(ii) Plus généralement, si $\{B_i\}_{i \geq 1}$ est une partition de Ω et $\mathbb{P}(B_i) > 0$ pour tout i alors

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

esquisse de preuve. Graphiquement, une partition de Ω est représentée par la figure 1.2. Et la figure 1.3 permet de voir que $\mathbb{P}(A) = \sum_{i1} \mathbb{P}(A \cap B_i)$ et en utilisant la définition d'une probabilité conditionnelle on peut conclure ■

■ **Exemple 1.6** Deux usines seulement au monde produisent des caméras ultra-performantes nommées "IHDI-KHAY". Un pourcentage de 20% des caméras produites par l'usine I est défectueux et seulement 5% sont défectueuses dans l'usine II. L'usine I produit le double de ce que produit l'usine II chaque semaine. Quellez est la chance de tomber sur une "IHDI-KHAY" non-défectueuse choisie aléatoirement lors d'une semaine de production ?
On voit bien que cette probabilité dépend de l'usine d'origine. On pose

$$A = \{\text{La IHDI-KHAY choisie est non défectueuse}\}, B = \{\text{La IHDI-KHAY est produite par l'usine I}\}$$

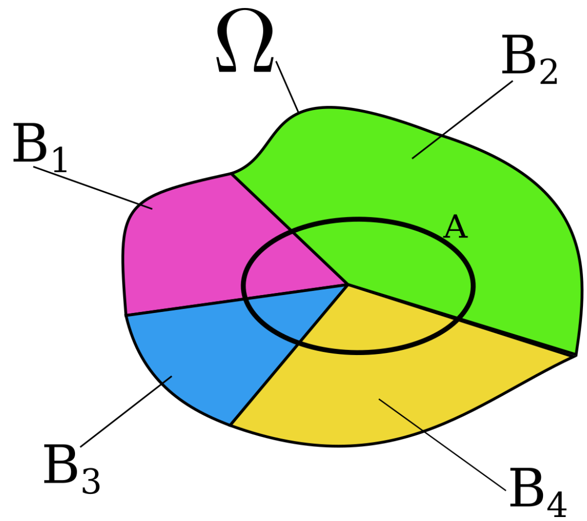


FIGURE 1.2 – La partition $\{B_i\}_{i=1}^{4}$ de Ω

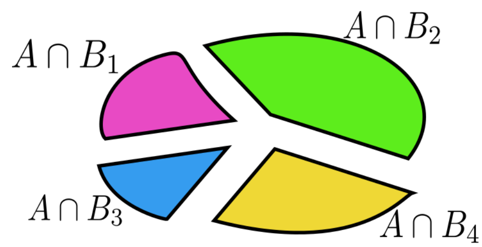


FIGURE 1.3 – A est décomposé en quatre parties disjointes.

L'espace des événements peut être partitionné en deux : B et B^C. Donc on utilise le Lemme 1.5.1 :

$$\mathbb{P}(A) = \mathbb{P}(A|B).\mathbb{P}(B) + \mathbb{P}(A|B^C).\mathbb{P}(B^C)$$

D'après les données de l'énoncé, on sait que $\mathbb{P}(A|B) = 80\% = 4/5$, $\mathbb{P}(B) = 2/3$, $\mathbb{P}(A|B^C) = 95\% = 19/20$ et $\mathbb{P}(B^C) = 1/3$ donc $\mathbb{P}(A) = 51/60 = 85\%$.

1.6 Indépendance d'événements

A et B sont indépendants si la réalisation de A n'influence pas B et inversement. On peut écrire $\mathbb{P}(A|B) = \mathbb{P}(A)$ et $\mathbb{P}(B|A) = \mathbb{P}(B)$. Dans ce cas là, on utilise $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

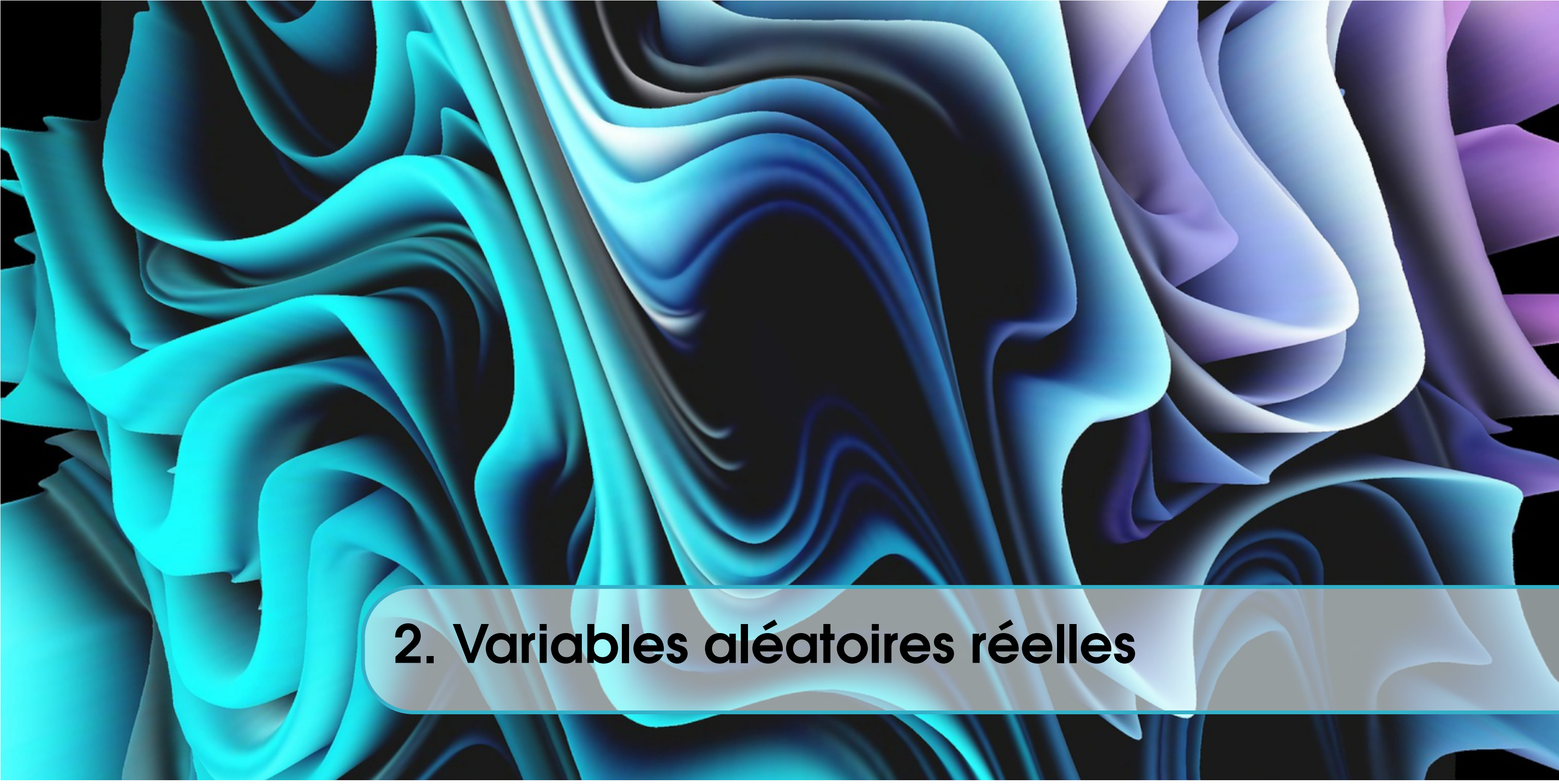
R Une erreur commune chez les étudiants est de considérer que A et B sont indépendants si $A \cap B = \emptyset$. Un exemple qui dissipe le doute est le lancement de dé une seule fois en considérant les événements $A = \{\text{pair}\}$ et $B = \{\text{impair}\}$. On a $\mathbb{P}(A) = \mathbb{P}(B) = 1/2$ et $0 = \mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0 \neq \mathbb{P}(A)\mathbb{P}(B) = \frac{1}{2} \cdot \frac{1}{2} = 1/4$.

Définition 1.6.1 Soit une famille d'indices $I \subset \mathbb{N}$. Une famille $\{A_j\}_{j \in I}$ est **indépendante** si $\mathbb{P}(\cap_{j\in J}) = \prod_{j\in J} \mathbb{P}(A_j)$ pour tout $J \subseteq I$.

Définition 1.6.2 Une famille est **deux-à-deux indépendante** si $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$, $\forall i \neq j$.

Une famille est deux à deux indépendante n'est pas forcément une famille indépendante. En voici un exemple :

■ **Exemple 1.7** Soit $\Omega = \{abc, acb, cab, cba, bca, bac, aaa, bbb, ccc\}$ avec $\mathbb{P}(\omega_k) = 1/9$ supposez $A_k = \{$k-ème lettre est 'a'$\}$. Puisque $A_1 = \{abc, acb, aaa\}$, $A_2 = \{bac, cab, aaa\}$ et $A_3 = \{bca, cba, aaa\}$. Il est clair que $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(\{aaa\}) = 1/9$ pour tout $i \neq j$ dans $\{1, 2, 3\}$ alors la famille $\{A_k\}$ est deux à deux indépendante . Mais $\mathbb{P}(A_1 \cap A_2 \cap A_3) = 1/3 \neq \mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) = 1/27$. ■



2. Variables aléatoires réelles

2.1 Introduction

Dans les expériences probabilistes effectuées, elle se peut que les résultats de l'expérience en soi ne soient pas ce qui importe le plus mais plutôt ses conséquences. Les résultats d'une expérience ne permettent pas, dans plusieurs cas, de les manipuler numériquement. Pour ces raisons, l'utilisation des variables aléatoires s'avère très utile pour répondre à ces besoins.

2.2 Exemples de motivation

On considère une expérience où on lance une pièce deux fois consécutives $\Omega = \{FF, PP, FP, PF\}$. On associe à cet expérience une fonction $X : \Omega \longrightarrow \mathbb{R}$ qui compte le nombre de faces apparues. On voit bien que

$$X(PP) = 0, \quad X(FP) = X(PF) = 1 \quad \text{et} \quad X(FF) = 2.$$

On imagine maintenant un parieur qui dispose d'une fortune de $1DH$ et mise sa fortune sur le résultat de l'expérience. Il parie de manière cumulative que sa fortune double à chaque apparition de "Face" et qu'il est ruiné si "Pile" apparaît une seule fois. On est d'accord que le parieur se soucie plutôt à éviter la ruine et non pas par les résultats de l'expérience en soi. On associera donc une fonction W :

$$W(PP) = W(FP) = W(PF) = 0 \quad \text{et} \quad W(FF) = 4.$$

Les fonctions X et W sont associés à la même expérience et prennent des valeurs dans un ensemble de valeurs inclus dans $\mathbb{R}$. Notre objectif est de décrire la distribution des probabilités des valeurs prises par ces fonctions. Dans les exemples précédents on peut voir que la fonction X a plus de probabilité de prendre la valeur 1 par rapport aux autres valeurs : 0 et 2. Et aussi, selon la définition de X, il n'y a aucune chance que cette valeur soit inférieur à 0 ou supérieur à 3 ! On peut faire des remarques similaires pour la fonction W. Plus précisément, $\mathbb{P}(X \to 0) = 1/4, \mathbb{P}(X \to 1) = 1/2$ et $\mathbb{P}(X \to 2) = 1/2$. Et $\mathbb{P}(W \to 0) = 3/4$ et $\mathbb{P}(X \to 4) = 1/4$.

2.3 Construction d'une variable aléatoire

On peut être tenté, dans un premier temps, afin d'accomplir notre objectif de détermination de la distribution des probabilités pour X et W de poser

$$f(x) = \text{"probabilité que X soit égal à la valeur x"}$$

Mais cette approche n'est pas la bonne bien qu'elle soit la plus évidente. En effet, cette déclaration peut paraître contre-intuitif car les exemples que nous discutons jusqu'à maintenant sont discrets. Mais si on considère un exemple où nous manipulons des valeurs continues plutôt que discrètes on comprendra que ce premier choix n'est pas correcte. On se donne à un jeu de tir à l'arc où on

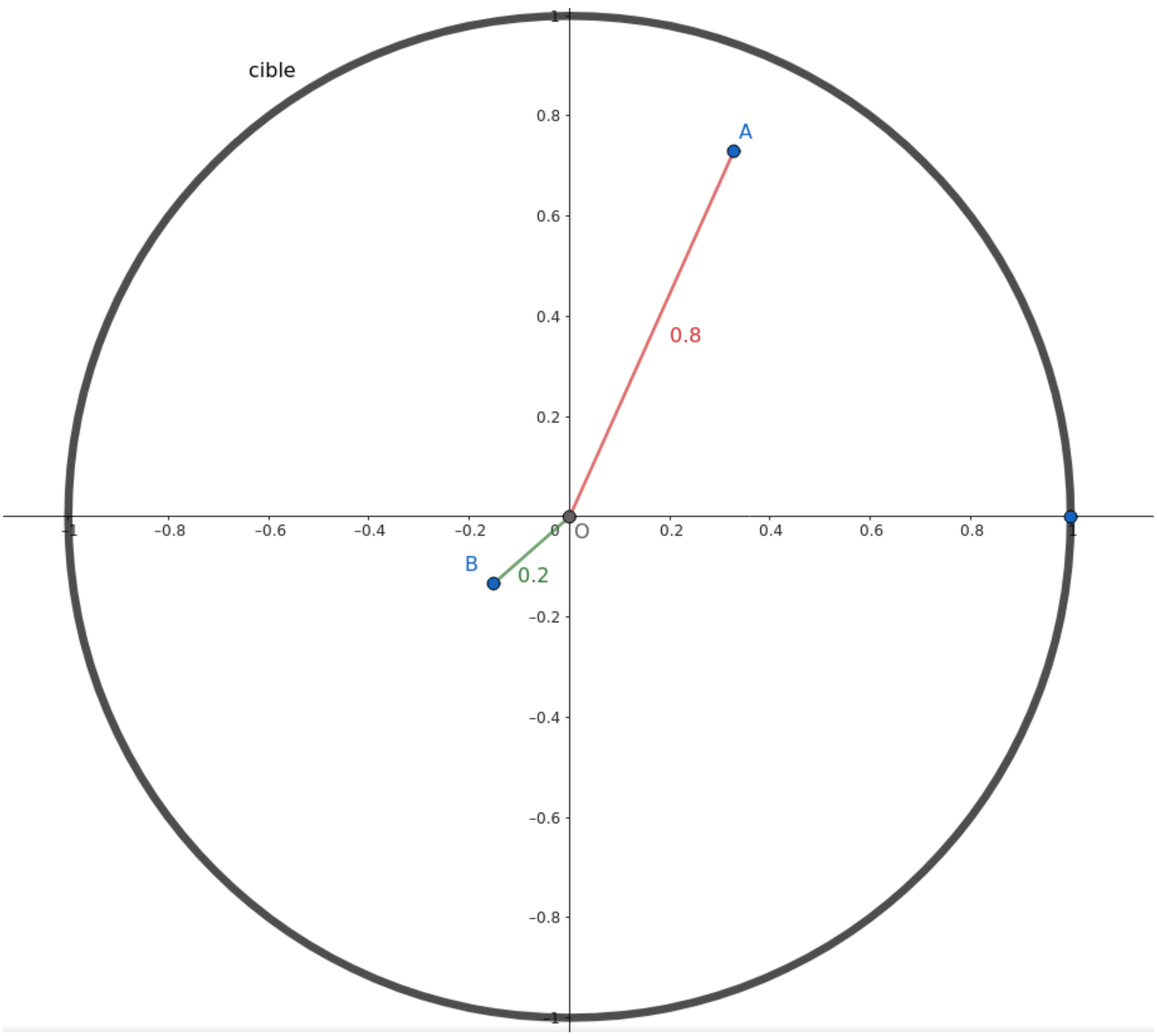


FIGURE 2.1 – Cible avec deux points d'impact A et B.

utilise des flèches très pointues sur une cible circulaire de rayon $R = 1m$. La surface de cette cible est $S = \pi\ m^2$. On supposera, afin de simplifier, que la flèche ne sort jamais en dehors de la cible et l'atteint toujours. L'univers ici est

$$\Omega = \{(x,y) \in \mathbb{R}^2, \quad x^2+y^2 \leq 1\}$$

On associe à ce jeu une fonction de score Z qui attribue au tireur une note selon l'inverse de la distance qui sépare le point d'impact et le centre du cercle. On suppose que la cible est centrée autour du point $(0,0)$ du plan réel comme le montre la figure 2.1. Le score obtenu pour l'impact A est $Z(A) = 1/0.8 = 1.25$ et pour le point B on a $Z(B) = 1/0.2 = 5$. Si maintenant on cherche à associer à cette expérience une mesure de probabilité, comme dans le cas discret, c'est à dire $P(A) = \frac{card(A)}{card(\Omega)}$, on se heurte contre une fatalité cruelle $\mathbb{P}(A) = \frac{1}{\infty} = 0$. En fait, tous les points auront une probabilité nulle, or on est sûr qu'au moins un point serait atteint ! Pour surmonter ce problème, il faut absolument trouver un autre moyen de mesurer la probabilité, et ce moyen est de tourner l'attention plutôt à la surface atteinte et non pas au point atteint. Autrement dit, on calculera la probabilité que le point d'impact soit dans une région I de surface S_I. Et vu que tous les points ont la même probabilité d'être touchés, on peut écrire que $\mathbb{P}(X \in I) = \frac{S_I}{S}$. Donc plus la région est grande plus sa probabilité est grande comme on peut le constater à partir de la figure 2.2. La probabilité d'atteindre un point dans la région I_1 est $\mathbb{P}(X \in I_1) = \frac{S_{I_1}}{S} = 0.06/\pi \simeq 0.0191$ et $\mathbb{P}(X \in I_2) = \frac{S_{I_2}}{S} = 0.36/\pi \simeq 0.1145$. Revenons maintenant à la variable de comptage de score Z. Si on veut calculer la probabilité des points pour lesquelles le score est 2 on doit définir leur ensemble qui va représenter tous les points dont la distance avec le centre est $1/2$, or nous pouvons l'exprimer comme suit

$$C_{1/2} = \{(x,y) \in \mathbb{R}^2, \quad x^2+y^2 = 1/2\}$$

Et là nous tentons encore une fois de rester ferme sur notre première intuition de définir la fonction $f(x)$ telle qu'elle a été définie et on se rend compte vite que

$$f(2) = \mathbb{P}(Z \to 2) = \frac{\text{surface } C_{1/2}}{S} = 0,$$

car la surface d'un cercle est nulle ! ! Quel malheur, quelle frustration ! Le remède efficace à cette difficulté est de changer un petit peu $f(x)$ en définissant plutôt

$$F(x) = \text{probabilité que X soit inférieur à la valeur x}$$

La fonction F mesure plus précisément $F(x) = \mathbb{P}(A(x))$ où $A(x) = \{\omega \in \Omega, X(\omega) \leq x\}$. L'écriture $\{X \leq x\}$ est juste une notation pour économiser l'encre utilisé mais surtout parce que le contexte est clair et n'embrouille pas l'idée essentielle dans les calculs. Il y a un petit tour de ruse que je suis en train d'employer depuis un bon moment et que certainement le lecteur attentif a détecté. Pour se permettre d'écrire $\mathbb{P}$(quelque chose) il faut que le "quelque chose" soit mesurable c'est à dire un élément de la tribu de l'espace probabilisé. C'est exactement ça ce qui fait la particularité des **variables aléatoires** d'où la définition suivante

Définition 2.3.1 Soit $(\Omega, \mathscr{F}, \mathbb{P})$ un espace probabilisé et $X : \Omega \longleftrightarrow \mathbb{R}$ une fonction. On dit que X est une variable aléatoire réelle (v.a.r) ssi

$$\forall B \in \mathscr{B}(\mathbb{R}), \quad X^{-1}(B) \in \mathscr{F}$$

C'est à dire que les images inverses des valeurs prises par X doivent être des éléments mesurables

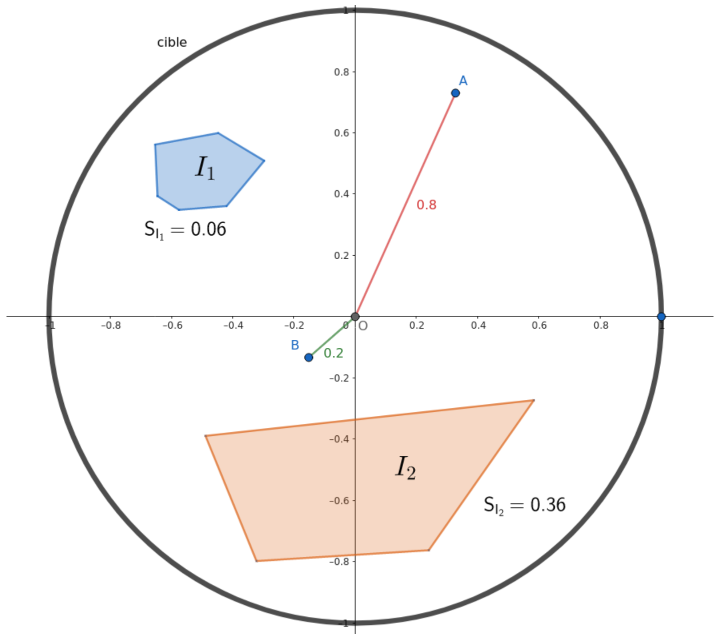


FIGURE 2.2 – Cible avec deux points d'impact A et B.

pour pouvoir leur appliquer la mesure de probabilité.

R Généralement les variables aléatoires réelles sont notées par des lettres majuscules et les valeurs possibles sont notées par des lettres minuscules de l'alphabet latin.

Définition 2.3.2 La fonction de répartition d'une v.a.r X est une application de $\mathbb{R}$ dans $[0,1]$ telle que

$$F(x) = \mathbb{P}(\{X \leq x\}) = \mathbb{P}(X^{-1}(]-\infty, x]))$$

Revenons maintenant aux exemples vus dans la Section 2.2. On établit les fonctions de répartition de la manière suivante

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ 1/4 & \text{si} \quad 0 \leq x < 1, \\ 3/4 & \text{si} \quad 1 \leq x < 2, \\ 1 & \text{si} \quad x \geq 2. \end{cases} \qquad F_W(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ 3/4 & \text{si} \quad 0 \leq x < 4, \\ 1 & \text{si} \quad x \geq 4. \end{cases}$$

Comment nous l'avons établi ? Essayons de détailler juste pour la v.a X et en déduire celle de W. Prenez une valeur réelle, disons $x = -1$. Vous êtes d'accord que $\mathbb{P}(\{x \leq -1\}) = 0$ si vous remplacez -1 par n'importe quelle autre valeur $x < 0$ car les valeurs prises par x sont soit $0, 1$ ou 2 et donc en aucun cas X ne peut être inférieur à une valeur strictement négative. On refait la même démarche, prenez une valeur comprise dans $[0,1[$, par exemple $x = 0.35$. La seule valeur de X qui est inférieur ou égal à 0.35 est la valeur $X = 0$. Donc $\mathbb{P}(\{X \leq 0.35\}) = \mathbb{P}(\{X = 0\}) = 1/4$. Remplacez 0.35 par n'importe quelle autre valeur de $[0,1[$ on aboutira à la même conclusion. Maintenant on prend $x \in [1,2[$, disons $x = 1.4$. On voit bien qu'il y a deux valeurs possible de X qui sont inférieures ou égales à 1.4, ce sont 0 ou 1. Donc, $\mathbb{P}(\{X \leq 0.35\}) = \mathbb{P}(\{X = 0\} \cup \{X = 1\})$ Et

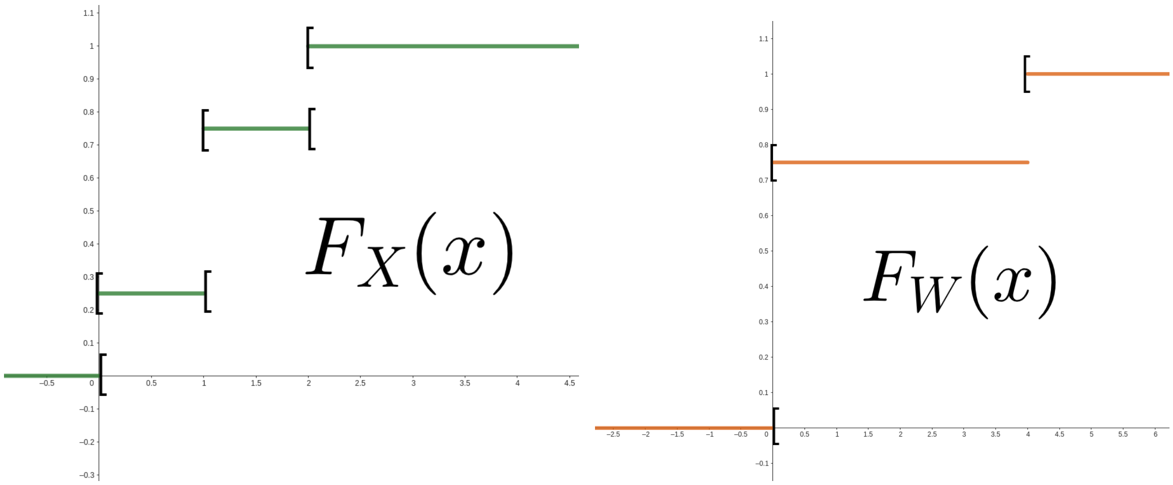


FIGURE 2.3 – Les graphes de F_X et F_W.

puisque les événements $\{X = 0\}$ et $\{X = 1\}$ sont disjoints alors $\mathbb{P}(\{X = 0\} \cup \{X = 1\}) = \mathbb{P}(\{X = 0\}) + \mathbb{P}(\{X = 1\}) = 1/4 + 1/2 = 3/4$. Par la même logique on enchaîne pour x dans $[2, +\infty[$ on trouve $\mathbb{P}(\{X \leq x\}) = \mathbb{P}(\{X = 0\}) + \mathbb{P}(\{X = 1\}) + \mathbb{P}(\{X = 2\}) = 1$. La figure 2.3 montres les graphes qui correspondent à F_X et F_W. Le lecteur attentif remarquera qu'on a dépassé l'univers des événements et qu'on s'est concentré sur le comportement et les valeurs de X et de W. Bien que nous soyons partis de la même expérience, les variables aléatoires permettent de la quantifier et d'en tirer une analyse qui dépasse l'abstraction de l'espace probabilisé. Ceci montre à quel point c'est un outil simple et efficace. À partir des graphes de fonctions de répartition nous citons le lemme suivant.

Lemme 2.3.1 Soit F_X une fonction de répartition d'une v.a.r X alors

(a) $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$

(b) si $x < y$ alors $F(x) \leq F(y)$

(c) F est continue à droite et admet toujours une limite à gauche (c'est une fonction càdlàg)

Démonstration. (a) Soit $B_n = \{\omega \in \Omega, X(\omega) \leq -n\} = \{X \leq -n\}$. Par construction, on voit que $\{B_n\}$ est une famille décroissante $B_1 \supset B_2 \supset \ldots$. Par le théorème [...], ■

2.4 Types de variables aléatoires

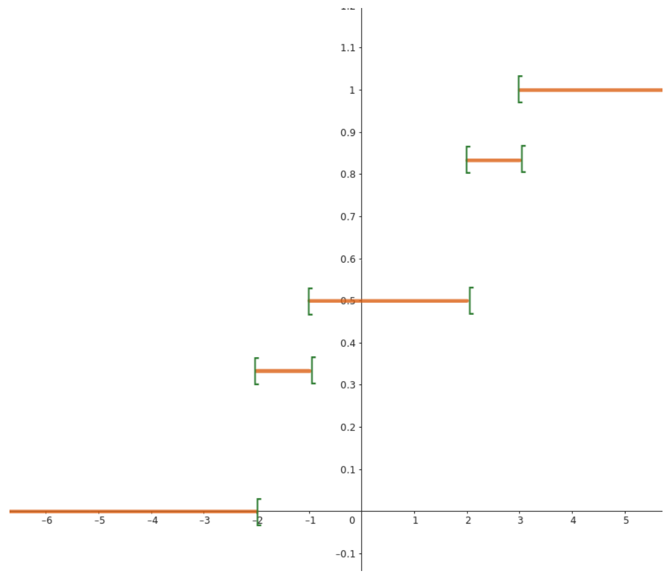
2.5 Caractéristiques d'une variable aléatoire

Il est important de noter que dans cette partie on traitera à la fois la variable aléatoire comme une fonction d'un point de vue probabiliste et comme une collection de données d'un point de vue statistique.

Tendance centrale

Définition 2.5.1 La **médiane** Med définit la valeur de la fonction de distribution F qui vérifie $F(Med) = 1/2$. La médiane partage la population en deux parties égales.

x_i	-2	-1	1	2
$\mathbb{P}(X = x_i)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$



Par exemple, dans la figure ci-contre, on illustre la fonction de répartition de la v.a.r qui suit la loi montrée sur le tableau ci-haut. La valeur médiane dans ce cas est -1 car $F(-1) = 1/2$. On se donne un exemple statistique cette fois-ci où nous avons les tableaux de données Y et Z comme suit

Y	1	8	4	3	1	0	$--$
Z	2	4	3	8	90	1	5

La première étape dans ce calcul est de ranger les valeurs dans un ordre croissant. La deuxième étape, on prend la valeur au "milieu", donc $Med(Z) = 4$ et quand le nombre de valeur est pair alors on applique une moyenne des deux valeurs centrales c'est à dire $Med(Y) = \frac{1+3}{2}$.

R Comme en statistique, compte tenu du cas d'une v.a.r discrète, on peut être amené à un calcul de moyenne lorsque aucun x_i ne fournit "exactement" $F(x_i) = 1/2$.

Le concept de médiane ne permet pas d'estimer autour de quelle valeur moyenne la variable aléatoire est centrée mais elle pourrait aider à détecter certaines valeurs aberrantes. En général, tout l'intérêt de l'étude de tendance centrale est de mieux analyser et interpréter les données mais aussi de détecter les valeurs *hors tendance*. La médiane est un cas particulier d'un élément plus général qui est introduit dans la définition suivante.

Définition 2.5.2 On appelle **quantile d'ordre α** où $0 < \alpha < 1$ d'une v.a.r (ou ensemble de données d'une variable statistique) X dont la fonction de répartition est $F(x)$, la valeur x_α pour laquelle $F(x_\alpha) = \alpha$. On dit que x_α est un quantile d'ordre α, en particulier, la médiane est un quantile d'ordre $1/2$.

Définition 2.5.3 Les **quartiles** sont des quantiles particuliers (très utilisés dans le domaine statistique) notés $Q_i (i = 1, 2, 3)$. Ils correspondent aux quantiles d'ordres : 0.25, 0.5 et 0.75 respectivement.

■ **Exemple 2.1** Soit X la variable statistique dont les valeurs sont représentées dans le tableau suivant

X	8	7	8	6	4	6	5	3	1	3

en ordonnant les valeurs de X on obtient

X	1	3	3	4	5	6	6	7	8	8
			$\uparrow$			$\uparrow$		$\uparrow$		
			Q_1			$Q_2 = 5.5$		Q_3		

Définition 2.5.4 Le **mode** est la valeur la plus dominante. C'est à dire celle qui a la plus grande probabilité ou celle qui a la plus grande fréquence si on parle d'un point de vue statistique.

Quand la variable aléatoire est continue à densité f, le mode représente le maximum de la fonction de densité. C'est à dire que nous cherchons une valeur M telle que $f'(M) = 0$ et $f"(M)) \leq 0$.

Espérance mathématique

Bien que l'espérance mathématique est liée souvent au concept d'une "moyenne" mais en réalité c'est une notion beaucoup plus profonde et dépasse les objectifs de ce cours. Vu que nous manipulons des variables aléatoires réelles, nous nous limiterons au calcul des espérances pour des variables aléatoires discrètes et des variables aléatoires continues à densité.

Définition 2.5.5 Soit X une v.a discrète prenant des valeurs dans $\{x_1, \ldots, x_n\}$ et dont les probabilités associées sont $\mathbb{P}(\{X = x_i\}) = p_i$ avec $\sum_i p_i = 1$. On définit l'espérance de X comme suit

$$\mathbb{E}[X] = \sum_i p_i x_i$$

Statistiquement parlant, les probabilités p_i sont remplacés par des fréquences $f_i = \frac{n_i}{N}$ où n_i et le nombre de fois où la valeur x_i est compté sur N répétitions. Donc si $\mathbb{E}[X]$ représente une moyenne **théorique** alors la moyenne **expérimentale** est $\bar{X} = \sum f_i x_i$.

Définition 2.5.6 Si X est une variable aléatoire continue de densité f alors l'espérance de X est

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f(x) dx$$

Proposition 2.5.1 Si X est une v.a.r et $x \mapsto y(x)$ est une fonction continue. Alors la v.a.r $Y = y(X)$ a une espérance

$$\begin{cases} \mathbb{E}[Y] = \sum_{i=1}^{N} p_i x_i & \text{si } X \text{ est discrète} \\ \mathbb{E}[Y] = \int_{-\infty}^{+\infty} y(x) f(x) dx & \text{si } X \text{ est continue de densité } f \end{cases}$$

Voici quelques propriétés essentielles de l'espérance

Proposition 2.5.2 Soit X une v.a.r intégrable

(i) Si $X = a$ est constante alors $\mathbb{E}[a] = a$.
(ii) On a aussi, $\mathbb{E}[\mathbb{E}[X]] = \mathbb{E}[X]$.
(iii) Si a et b sont deux réels alors $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$.
(iv) On a $\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$.

Paramètres de dispersion

Les paramètres de dispersion sont nombreux mais on se limitera à la variance et l'écart-type qui permettent de mesurer est-ce que les valeurs sont plutôt rapprochées autour de la moyenne ou éloignées.

Définition 2.5.7 Soit X une v.a.r d'espérance $\mathbb{E}[X]$ finie est la quantité

$$var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

X_1	10	14	8	14	8
X_2	20	0	18	10	2

TABLE 2.1 – Deux groupes d'étudiants (1 et 2) passent le même examen et obtiennent les notes ci-dessus.

En développant le calcul on obtient la formule de Koenig

$$var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

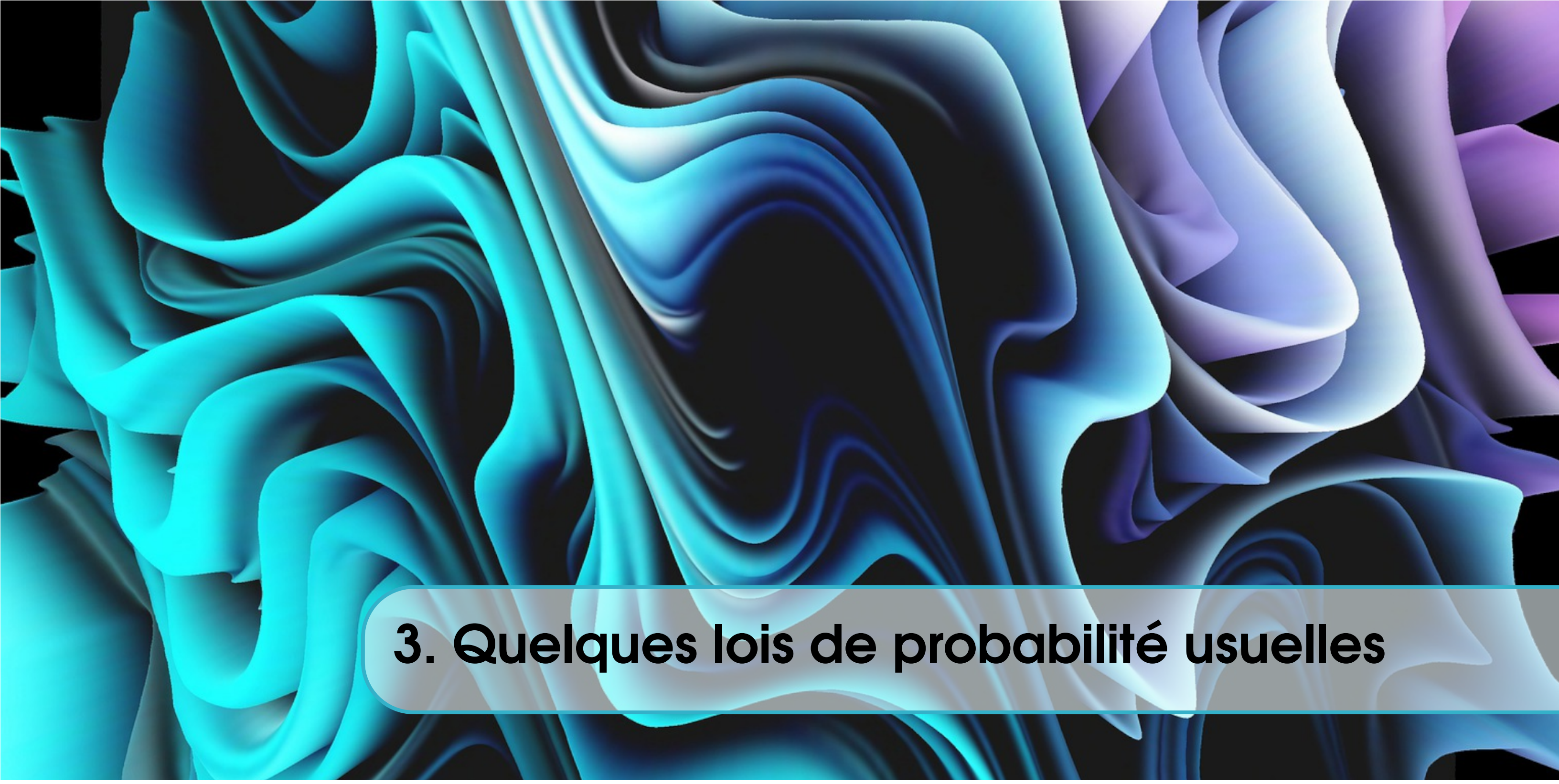
L'écart-type est dérivé de la variance et on le note souvent $\sigma(X)$ ou juste σ :

$$\sigma(X) = \sqrt{var(X)}.$$

■ **Exemple 2.2** Soit X_1 et X_2 deux variables statistiques représentant les notes sur 20 d'un examen de mathématiques dans deux classes différentes. La moyenne des deux classes est $\bar{X_1} = \bar{X_2} = \mathbb{E}[X_1] = \mathbb{E}[X_2] = 10$. Donc la tendance centrale est la même. Mais en calculant les variances nous obtenons

$$\begin{aligned} var(X_1) &= \sum p_i(x_i - \bar{X_1})^2 = (2/5)(8-10)^2 + (1/5)(10-10)^2 \\ &\quad +(2/5)(12-10)^2 + (2/5)(12-10)^2 = 3.2 \\ var(X_2) &= \sum p_i(x_i - \bar{X_2})^2 = (1/5)(0-10)^2 + (1/5)(2-10)^2 \\ &\quad +(1/5)(10-10)^2 + (1/5)(18-10)^2 + (1/5)(20-10)^2 = 65.6 \end{aligned}$$

On voit bien que la variance de X_2 est plus élevée que celle de X_1, ce qui peut être interprété comme une dispersion plus rapprochée autour de la moyenne pour X_1 et plutôt des valeurs extrêmes pour X_2 loin de leur moyenne. Vous avez certainement constaté que la valeur 65.6 n'est pas "significative" dans le cadre de ce problème qui traite une note d'examen sur 20 ! 65.6 est une valeur qui dépasse l'ordre des valeurs traitées. C'est pourquoi l'usage de l'écart-type est important et permet de redimensionner les valeurs et tomber sur des ordres de grandeurs compatible avec les données manipulées. Ici, $\sigma(X_1) \simeq 1.78$ et $\sigma(X_2) \simeq 8.09$ ce qui permet maintenant de mieux estimer que la répartition des notes est de ± 8 points autour de la moyenne (éloignées) et de ± 1.78 points par rapport à la classe 1. Donc la classe 2 souffre de grandes différences individuelles.



3. Quelques lois de probabilité usuelles

3.1 Lois discrètes

Dans cette section nous considérons que X est une variable aléatoire discrète.

Loi uniforme

On considère que l'ensemble des valeurs de X est $V = \{1,2,\dots,n\}$. On dit que la loi de X est uniforme si toute les valeurs de X ont la même probabilité c-à-d $\mathbb{P}(\{X=1\}) = \mathbb{P}(\{X=2\}) = \dots = \mathbb{P}(\{X=n\}) = 1/n$. La représentation graphique de cette loi (fonction de masse) ainsi que de sa fonction de répartition est illustrée sur la figure 3.1 On calcule l'espérance et la variance de X et on

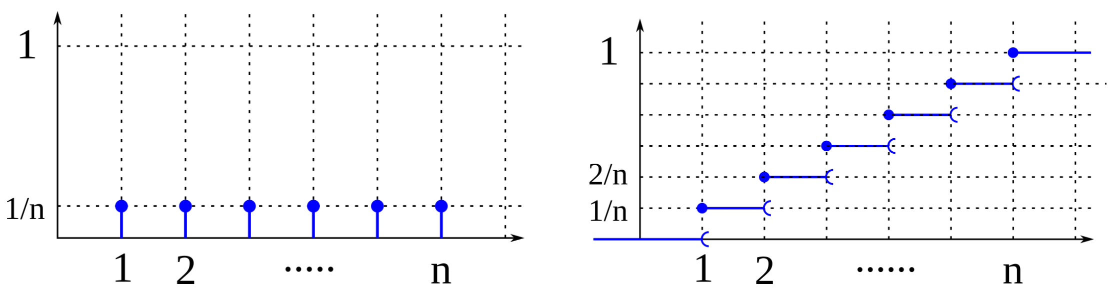


FIGURE 3.1 – À gauche la fonction de masse (loi de X). À droite la fonction de répartition F_X.

trouve

$$\mathbb{E}[X] = \frac{n+1}{2}, \qquad var(X) = \frac{n^2-1}{12}$$

Cette loi intervient dans les situations où la probabilité de chaque réalisation est la même (équiprobabilité) comme le cas d'une pièce de monnaie ou un dé non-biaisés. Dans ce cas on dit que X **suit une loi uniforme** et on note $X \sim \mathrm{unif}(n)$.

Loi de Bernoulli et Loi binomiale

Quand on modélise une expérience où on souhaite obtenir deux résultats antagonistes : "réussite" ou "échec", la loi de Bernoulli serait le candidat le plus approprié. Par exemple, on lance une pièce de monnaie équilibrée et on associe à Pile l'échec et à Face la réussite. Ou bien, un lancement de dé non-biaisé où on associe à l'apparition de ⚄ la réussite et à son non apparition l'échec. Plus précisément la variable aléatoire prend la valeur 1 en cas de "réussite" et 0 en cas d'"échec". Dans la première expérience, puisque la pièce de monnaie est non truquée $\mathbb{P}(\{X=1\})=1/2$ et $\mathbb{P}(\{X=0\})=1-1/2=1/2$. Et dans la deuxième expérience $\mathbb{P}(\{X=1\})=1/6$ et $\mathbb{P}(\{X=0\}=1-1/6=5/6$. En général, soit $p\in[0,1]$ si $\mathbb{P}(\{X=1\})=p$ et $\mathbb{P}(\{X=0\})=1-p$ alors on dit que X **suit une loi de Bernoulli de paramètre** p et on note $X\sim\mathscr{B}(p)$. L'espérance et la variance d'une variable de Bernoulli sont :

$$\mathbb{E}[X]=p, \qquad var(X)=p(1-p).$$

Quant à une loi binomiale, c'est la répétition de la loi de Bernoulli n fois. Dans ce cas, on a $\mathbb{P}(\{X=k\})=C_n^k p^k(1-p)^{n-k}$. On dit dans ce cas que la variable **suite une loi binomiale de paramètres** (n,p) et on note $X\sim\mathscr{B}(n,p)$. L'espérance et la variance sont exprimées comme suit

$$\mathbb{E}[X]=np, \qquad var(X)=np(1-p).$$

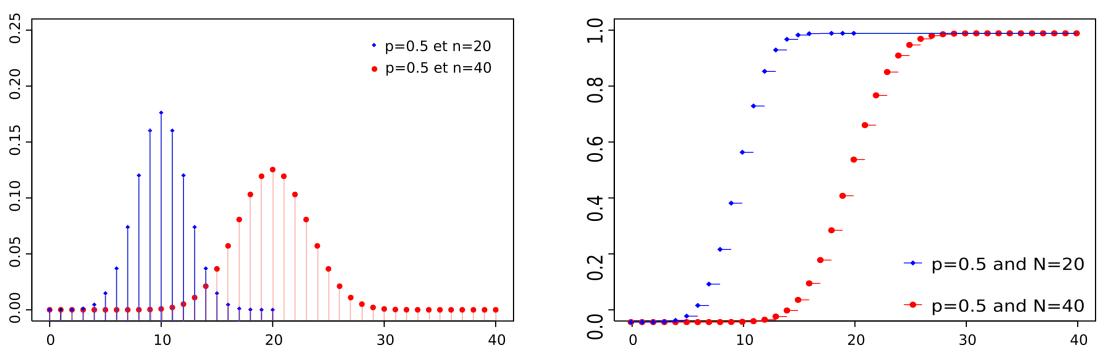


FIGURE 3.2 – À gauche deux fonctions de masse pour différents paramètres de la loi binomiale. À droite on montre leurs fonctions de répartition respectives.

Loi de Poisson

La loi de Poisson est utilisé souvent pour modéliser des phénomène "rares" comme par exemple le nombre d'accidents par jour, le nombre des défauts trouvés dans une pièce produite par une usine de fabrication etc. En effet, la loi stipule que plus la valeur de X est grande moins elle est probable. Pour mieux sentir cette loi, pensez à une variable qui compte le nombre d'accidents routiers enregistré par jour sur l'autoroute liant Tanger et Rabat. On est d'accord qu'il est fort possible d'enregistrer quatre accidents, 5 accidents , 6 accidents mais surtout pas 50 accidents. Donc la rareté vient de l'improbabilité d'être égale à certaines valeurs loin de la moyenne. Elle est exprimé par ... retenez bien votre souffle ...

$$\mathbb{P}(\{X=k\})=\frac{\lambda^k e^{-\lambda}}{k!}.$$

Certains d'entre vous seront dégoûtés, d'autres frustrés et certains émerveillés devant cette expression. Mais en contemplant profondément la formule elle contient des éléments clés de ses exigences notamment la division par le factorielle qui permet de décroître rapidement vers 0 quand k tend vers l'infini. Son espérance et sa variance sont les suivants :

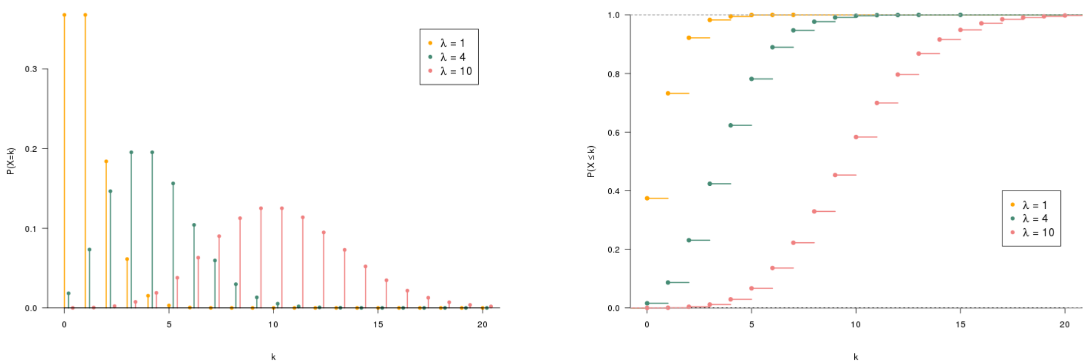


FIGURE 3.3 – À gauche les fonction de masse pour différentes paramètres de Poisson λ. À droite leurs fonctions de répartition respectives (Par Daticien — Travail personnel, CC BY-SA 4.0, https ://commons.wikimedia.org/w/index.php ?curid=55091962).

$$\mathbb{E}[X] = \lambda, \qquad var(X) = \lambda$$

On dit que la variable X **suit une loi de Poisson de paramètre** λ et on note $X \sim \text{Pois}(\lambda)$.

R La moyenne d'une variable de Poisson peut surprendre mais en réalité un simple calcul montre que la somme obtenue est un Développement Limité de la fonction $x \mapsto \exp(x)$ ce qui permet de faire des simplifications.

3.2 Lois continues

Dans cette section on considère que X est une variable aléatoire continue à densité $f(x)$.

Loi uniforme

Soit un intervalle $I = [a,b] \subset \mathbb{R}$. On suppose que la v.a.r X prend ces valeurs dans cet intervalle. La loi uniforme est donnée par la fonction de densité

$$f(x) = \frac{1}{b-a},$$

pour $x \in [a,b]$ et 0 sinon. Sa fonction de distribution est

$$F(x) = \begin{cases} 0 & \text{si} \quad x < a \\ \frac{x-a}{b-a} & \text{si} \quad a \leq x < b \\ 1 & \text{si} \quad x \geq b. \end{cases}$$

Si on prend deux sous-intervalles I_1 et I_2 de I alors cette loi exprime deux choses :
1. Plus la taille de I_1 est grande, plus sa probabilité est grande.
2. Si I_1 et I_2 ont la même taille alors leurs probabilités sont égales.

Loi normale (gaussienne)

C'est la fameuse loi de courbe en cloche [1] qui est présente dans une multitude de phénomènes sur lesquelles on applique des mesures continues : pointure de chaussure, taille d'un être humain ... Elle est adaptée aux données qui montrent une certaine symétrie autour d'une moyenne avec une probabilité de plus en plus faible pour les valeurs extrêmes à gauche et à droite de la moyenne. Si la variable X suit une loi normale de moyenne m et d'écart-type σ alors sa fonction de densité s'exprime comme suit

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

1. Bell curve

Rappels sur le dénombrement