

# PROBABILITÉS

Contrairement au programme de première année, l'univers d'une expérience aléatoire n'est plus supposé fini.

## 1 Espace probabilisé

### 1.1 Ensembles dénombrables

#### Définition 1.1 Ensemble dénombrable

Un ensemble est dit dénombrable s'il est en bijection avec  $\mathbb{N}$ .

**REMARQUE.** On dit qu'un ensemble est **au plus dénombrable** s'il est fini ou dénombrable.

**REMARQUE.** La relation «être en bijection avec» est une relation d'équivalence. Autrement dit, un ensemble en bijection avec un ensemble dénombrable est lui-même dénombrable.

#### Proposition 1.1

Les parties infinies de  $\mathbb{N}$  sont dénombrables.

#### Exemple 1.1

L'ensemble des entiers pairs et l'ensemble des entiers impairs sont dénombrables de même que l'ensemble des nombres premiers.

#### Proposition 1.2

Un ensemble est fini ou dénombrable si et seulement si il est en bijection avec une partie de  $\mathbb{N}$ .

#### Exercice 1.1

Soit  $A$  un ensemble. Montrer que les propositions suivantes sont équivalentes.

- (i)  $A$  est fini ou dénombrable ;
- (ii) il existe une injection de  $A$  dans un ensemble dénombrable ;
- (iii) il existe une surjection d'un ensemble dénombrable sur  $A$ .

#### Proposition 1.3 Opération sur les ensembles dénombrables

- Un produit cartésien fini d'ensembles dénombrables est dénombrable.
- Une réunion finie ou dénombrable d'ensembles finis ou dénombrables est finie ou dénombrable.

**Corollaire 1.1**

$\mathbb{Z}$ ,  $\mathbb{N}^p$  ( $p \in \mathbb{N}^*$ ) et  $\mathbb{Q}$  sont dénombrables.

**Exercice 1.2**

Montrer que  $\mathbb{Q}[X]$  est dénombrable.

**Proposition 1.4 Dénombrabilité du support d'une famille sommable**

Soit  $(u_i)_{i \in I} \in \mathbb{C}^I$  une famille **sommable**. Alors le support de la famille  $(u_i)_{i \in I}$ , c'est-à-dire l'ensemble  $\{i \in I, u_i \neq 0\}$ , est au plus dénombrable.

**Proposition 1.5**

$\mathbb{R}$  n'est pas dénombrable.

**REMARQUE.** Si  $X$  est un ensemble, il n'existe pas de surjection (et donc pas de bijection) de  $X$  sur  $\mathcal{P}(X)$ . Raisonnons par l'absurde et supposons qu'il existe une telle surjection. Considérons alors  $A = \{x \in X, x \notin f(x)\}$ . Par surjectivité de  $f$ , il existerait  $a \in X$  telle que  $A = f(a)$ . Mais alors  $a \in A \iff a \notin f(a) \iff a \notin A$ , ce qui est absurde.

**Exemple 1.2**

D'après la remarque précédente, il n'existe pas de bijection de  $\mathbb{N}$  sur  $\mathcal{P}(\mathbb{N})$ . Ainsi  $\mathcal{P}(\mathbb{N})$  n'est pas dénombrable.

Comme l'application  $\begin{cases} \mathcal{P}(\mathbb{N}) & \longrightarrow & \{0, 1\}^{\mathbb{N}} \\ X & \longmapsto & \mathbb{1}_X \end{cases}$  est bijective, ceci prouve également que  $\{0, 1\}^{\mathbb{N}}$  n'est pas dénombrable.

**REMARQUE.** On peut aussi prouver que  $\{0, 1\}^{\mathbb{N}}$  n'est pas dénombrable à l'aide d'un argument du type «diagonale de Cantor». Supposons que  $\{0, 1\}^{\mathbb{N}}$  soit dénombrable. Il existe donc une bijection  $f$  de  $\mathbb{N}$  sur  $\{0, 1\}^{\mathbb{N}}$ . On définit la suite  $u \in \{0, 1\}^{\mathbb{N}}$  telle que  $u_n = 1 - f(n)_n$  pour tout  $n \in \mathbb{N}$ . Alors  $u \notin f(\mathbb{N})$ . En effet, pour tout  $n \in \mathbb{N}$ ,  $u \neq f(n)$  puisque  $u_n \neq f(n)_n$ . Ceci contredit la surjectivité de  $f$ .

**1.2 Notion de tribu****Définition 1.2 Tribu**

On appelle **tribu** sur un ensemble  $\Omega$  toute partie  $\mathcal{A}$  de  $\mathcal{P}(\Omega)$  vérifiant les propriétés suivantes.

**L'univers appartient à la tribu**  $\Omega \in \mathcal{A}$ .

**Stabilité par passage au complémentaire**  $\forall A \in \mathcal{A}, \bar{A} \in \mathcal{A}$ .

**Stabilité par union dénombrable**  $\forall (A_n) \in \mathcal{A}^{\mathbb{N}}, \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ .

**REMARQUE.** Une tribu est aussi parfois appelée une  $\sigma$ -algèbre.

### Exemple 1.3

- $\mathcal{P}(\Omega)$  est une tribu sur  $\Omega$ . On l'appelle la **tribu pleine**.
- $\{\emptyset, \Omega\}$  est une tribu sur  $\Omega$ . On l'appelle la **tribu triviale**.

### Définition 1.3 Événement

- On appelle **événement** tout élément de la tribu.
- Un événement est dit **élémentaire** si c'est un singleton.
- On appelle événement **contraire** d'un événement  $A$  le complémentaire  $\bar{A}$  de cet événement dans l'univers.
- L'ensemble vide est appelé événement **impossible**.
- $\Omega$  est appelé l'événement **certain**.
- On dit que deux événements  $A$  et  $B$  sont **incompatibles** si l'événement  $A \cap B$  est impossible i.e.  $A \cap B = \emptyset$ .

Les propriétés suivantes des tribus découlent directement de la définition.

### Proposition 1.6

Soit  $\mathcal{A}$  une tribu sur un ensemble  $\Omega$ . On a alors les résultats suivants :

**Ensemble vide**  $\emptyset \in \mathcal{A}$ .

**Stabilité par intersection dénombrable**  $\forall (A_n) \in \mathcal{A}^{\mathbb{N}}, \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$ .

**Stabilité par union ou intersection finie**  $\forall (A_1, \dots, A_n) \in \mathcal{A}^n, \bigcup_{i=1}^n A_i \in \mathcal{A}$  et  $\bigcap_{i=1}^n A_i \in \mathcal{A}$ .

### Définition 1.4 Espace probabilisable

On appelle **espace probabilisable** tout couple  $(\Omega, \mathcal{A})$  où  $\mathcal{A}$  est une tribu sur  $\Omega$ .

#### Tribu engendrée

Soit  $\Omega$  un ensemble et  $\mathcal{C}$  une partie de  $\mathcal{P}(\Omega)$ . On appelle **tribu engendrée par  $\mathcal{C}$**  la plus petite tribu de  $\Omega$  contenant  $\mathcal{C}$ . C'est l'intersection des tribus contenant  $\mathcal{C}$ .

#### Jeu de pile ou face infini

On considère l'expérience aléatoire consistant à lancer une infinité de fois une pièce. On associe face à 0 et pile à 1. On peut alors considérer l'univers  $\Omega = \{0, 1\}^{\mathbb{N}^*}$ . On note  $P_n$  l'événement «obtenir pile au lancer  $n$ » autrement dit

$$P_n = \{\omega \in \Omega, \omega_n = 1\}$$

L'univers  $\Omega$  est généralement muni de la tribu engendrée par  $\{P_n, n \in \mathbb{N}^*\}$ .

### Méthode Lien entre opérations ensemblistes et quantificateurs

On peut déjà établir un lien entre union et quantificateur existentiel ainsi qu'entre intersection et quantificateur universel.

**Union**  $x \in \bigcup_{i \in I} A_i \iff \exists i \in I, x \in A_i$

**Intersection**  $x \in \bigcap_{i \in I} A_i \iff \forall i \in I, x \in A_i$

Par exemple, si  $(A_n)_{n \in \mathbb{N}}$  désigne une suite d'ensembles, on a l'équivalence suivante.

$$x \in \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k \iff \forall n \in \mathbb{N}, \exists k \geq n, x \in A_k$$

On comprend alors que  $x \in \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k$  si et seulement si  $x$  appartient à une infinité d'ensembles  $A_n$ .

De même, on a l'équivalence suivante.

$$x \in \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} A_k \iff \exists n \in \mathbb{N}, \forall k \geq n, x \in A_k$$

On comprend alors que  $x \in \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} A_k$  si et seulement si  $x$  n'appartient qu'à un nombre fini d'ensembles  $\overline{A_n}$ .

### Exemple 1.4

- L'événement «obtenir uniquement pile à partir du lancer  $n$ » est  $\bigcap_{k \geq n} P_k$ .
- L'événement «obtenir au moins un pile à partir du lancer  $n$ » est  $\bigcup_{k \geq n} P_k$ .
- L'événement «obtenir une infinité de fois pile» est  $\bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} P_k$ .
- L'événement «obtenir un nombre fini de faces» est  $\bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} P_n$ .

## 1.3 Espace probabilisé

### Définition 1.5 Probabilité

Soit  $(\Omega, \mathcal{A})$  un espace probabilisable. On appelle **probabilité** sur  $(\Omega, \mathcal{A})$  toute application  $\mathbb{P} : \mathcal{A} \rightarrow \mathbb{R}_+$  telle que

- $\mathbb{P}(\Omega) = 1$ ;
- $\mathbb{P}$  est  $\sigma$ -additive : pour toute suite d'événements  $(A_n)$  **deux à deux disjoints**,  $\mathbb{P}\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ .

### Définition 1.6 Espace probabilisé

On appelle **espace probabilisé** tout triplet  $(\Omega, \mathcal{A}, \mathbb{P})$  où  $(\Omega, \mathcal{A})$  est un espace probabilisable et  $\mathbb{P}$  une probabilité sur  $(\Omega, \mathcal{A})$ .

**Proposition 1.7 Propriétés des probabilités**

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

- Soit  $A \in \mathcal{A}$ . Alors  $\mathbb{P}(A) \in [0, 1]$ .
- $\mathbb{P}(\emptyset) = 0$ .
- Soit  $(A, B) \in \mathcal{A}^2$ . Alors  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ .
- Soit  $A \in \mathcal{A}$ . Alors  $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$ .
- Soit  $(A, B) \in \mathcal{A}^2$ . Si  $A \subset B$ , alors  $\mathbb{P}(A) \leq \mathbb{P}(B)$ .

**Définition 1.7 Événement négligeable/presque sûr**

Soit  $A$  un événement d'un univers probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .

**Événement négligeable** On dit que  $A$  est **négligeable** si  $\mathbb{P}(A) = 0$ .

**Événement presque sûr** On dit que  $A$  est **presque sûr** si  $\mathbb{P}(A) = 1$ .

**ATTENTION !**

- $A$  peut être un événement négligeable sans que  $A = \emptyset$ .
- $A$  peut être un événement presque sûr sans que  $A = \Omega$ .

**1.4 Probabilité conditionnelle****Définition 1.8 Probabilité conditionnelle**

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé et  $(A, B) \in \mathcal{A}^2$  tel que  $\mathbb{P}(B) \neq 0$ . On appelle **probabilité de  $A$  sachant  $B$**  notée  $\mathbb{P}(A | B)$  ou  $\mathbb{P}_B(A)$  le quotient  $\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ .



**ATTENTION !** La notation  $\mathbb{P}(A | B)$  peut être trompeuse. En effet,  $A | B$  ne désigne pas un événement : il n'existe pas de notion d'«événement conditionnel».

**Proposition 1.8**

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé et  $B \in \mathcal{A}$  tel que  $\mathbb{P}(B) \neq 0$ . Alors  $\mathbb{P}_B : A \in \mathcal{A} \mapsto \mathbb{P}_B(A)$  est une probabilité sur  $(\Omega, \mathcal{A})$ .

**Proposition 1.9 Formule des probabilités composées**

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé et  $(A_1, \dots, A_n) \in \mathcal{A}^n$  tel que  $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$ . Alors

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2 | A_1)\mathbb{P}(A_3 | A_1 \cap A_2) \dots \mathbb{P}(A_n | A_1 \cap \dots \cap A_{n-1})$$

**Définition 1.9** Système complet d'événements

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. On dit qu'une famille au plus dénombrable  $(A_i)_{i \in I} \in \mathcal{A}^I$  est un **système complet** d'événements si les  $A_i$  sont deux à deux disjoints et si  $\Omega = \bigsqcup_{i \in I} A_i$ .

**Proposition 1.10**

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé et  $(A_i)_{i \in I} \in \mathcal{A}^I$  un système complet d'événements. Alors  $\sum_{i \in I} \mathbb{P}(A_i) = 1$ .

**REMARQUE.** On appelle système **quasi-complet** d'événements toute famille au plus dénombrable  $(A_i)_{i \in I}$  d'événements deux à deux disjoints telle que  $\sum_{i \in I} \mathbb{P}(A_i) = 1$ .

**Proposition 1.11** Formule des probabilités totales

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé,  $(A_i)_{i \in I}$  un système complet d'événements et  $B$  un événement. Alors

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \mid A_i) \mathbb{P}(A_i)$$

**Proposition 1.12** Formule de Bayes

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soient  $A$  et  $B$  deux événements tels que  $\mathbb{P}(A) \neq 0$  et  $\mathbb{P}(B) \neq 0$ . Alors

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

Soient  $(A_i)_{i \in I}$  un système complet d'événements et  $A$  et  $B$  deux événements tels que  $\mathbb{P}(B) \neq 0$ . Alors

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A) \mathbb{P}(A)}{\sum_{i \in I} \mathbb{P}(B \mid A_i) \mathbb{P}(A_i)}$$

**REMARQUE.** La formule de Bayes et la formule des probabilités totales restent vraies si  $\mathcal{S}$  est un système quasi-complet d'événements.

## 1.5 Événements indépendants

**Définition 1.10** Événements indépendants

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. On dit que deux événements  $A$  et  $B$  sont **indépendants** si  $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$ .

**Proposition 1.13**

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. Soient  $A$  et  $B$  deux événements avec  $\mathbb{P}(B) > 0$ . Alors  $A$  et  $B$  sont indépendants si et seulement si  $\mathbb{P}(A \mid B) = \mathbb{P}(A)$



**ATTENTION !** Dire que deux événements sont indépendants ne signifie pas qu'ils sont incompatibles.

Au contraire, deux événements incompatibles ne sont généralement pas indépendants à moins que l'un d'entre eux soit de probabilité nulle.

### Définition 1.11 Famille d'événements indépendants

Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé et  $(A_i)_{i \in I}$  une famille d'événements. On dit que  $(A_i)_{i \in I}$  est une famille d'événements **mutuellement indépendants** si, pour toute partie **finie**  $J$  de  $I$

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$$

**REMARQUE.** Il découle directement de la définition que toute sous-famille d'une famille d'événements mutuellement indépendants est encore une famille d'événements mutuellement indépendants.

### Pile ou face infini

On considère un jeu de pile ou face infini modélisé par l'univers  $\Omega = \{0, 1\}^{\mathbb{N}^*}$ . On note à nouveau  $P_n = \{\omega \in \Omega, \omega_n = 1\}$  l'événement «obtenir pile au  $n^{\text{ème}}$  lancer et on considère la tribu  $\mathcal{A}$  engendrée par les  $P_n$ .

On peut prouver qu'il existe une probabilité  $\mathbb{P}$  sur  $(\Omega, \mathcal{A})$  telle que  $(P_n)_{n \in \mathbb{N}^*}$  soit une suite d'événement mutuellement indépendants et de même probabilité  $p \in [0, 1]$  ( $p = 1/2$  pour un pièce équilibrée).



**ATTENTION !** Si  $(A_i)_{i \in I}$  est une famille d'événements mutuellement indépendants alors les  $A_i$  sont indépendants deux à deux. Cependant, la réciproque est fausse.

### Exemple 1.5

On considère l'expérience aléatoire consistant à lancer deux dés. On munit l'univers de la probabilité uniforme. On considère les événements suivants :

- $A$  : «le résultat du premier dé est pair» ;
- $B$  : «le résultat du deuxième dé est pair» ;
- $C$  : «la somme des résultats des deux dés est pair».

Les événements  $A$ ,  $B$  et  $C$  sont deux à deux indépendants mais pas mutuellement indépendants.

### Proposition 1.14 Indépendance et complémentaire

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. Si  $A$  et  $B$  sont des événements indépendants, les événements  $A$  et  $\bar{B}$  sont également indépendants.

### Exercice 1.3

Soit  $(A_i)_{i \in I}$  une famille d'événements mutuellement indépendants d'un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . On considère une famille  $(B_i)_{i \in I}$  telle que pour tout  $i \in I$ ,  $B_i = A_i$  ou  $B_i = \bar{A}_i$ . Montrer que  $(B_i)_{i \in I}$  est également une famille d'événements mutuellement indépendants.

## 1.6 Continuité croissante et décroissante

### Proposition 1.15 Continuité croissante et décroissante

Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements d'un univers probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .

**Continuité croissante** Si  $(A_n)$  est **croissante** pour l'inclusion, alors  $\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right)$ .

**Continuité décroissante** Si  $(A_n)$  est **décroissante** pour l'inclusion, alors  $\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} A_n\right)$ .

**REMARQUE.** La continuité croissante et décroissante donne en particulier l'existence des limites des suites  $(\mathbb{P}(A_n))$  dans le cas où la suite d'événements  $(A_n)$  est monotone.

### Corollaire 1.2

Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements (non nécessairement monotone) d'un univers probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . Alors

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcup_{k=0}^n A_k\right) = \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right) \qquad \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcap_{k=0}^n A_k\right) = \mathbb{P}\left(\bigcap_{k=0}^{+\infty} A_k\right)$$

### Exemple 1.6

Dans un jeu de pile ou face infini, considérons l'événement  $A$  : «obtenir uniquement des faces». Alors  $A = \bigcap_{n \in \mathbb{N}} F_n$ . Posons  $E_n = \bigcap_{k=0}^n F_k$ . Alors  $\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \mathbb{P}(E_n)$ . Mais  $\mathbb{P}(E_n) = \frac{1}{2^{n+1}}$  pour tout  $n \in \mathbb{N}$ . Ainsi  $\mathbb{P}(A) = 0$ .

### Corollaire 1.3

Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements d'un univers probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . Alors

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$$

**REMARQUE.** On peut donner un sens à l'inégalité précédente même si la série  $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n)$  diverge, puisqu'étant à termes positifs, elle diverge dans ce cas vers  $+\infty$ .

### Proposition 1.16

- Une réunion **finie ou dénombrable** d'événements négligeables est un événement négligeable.
- Une intersection **finie ou dénombrable** d'événements presque sûrs est un événement presque sûr.



**Exercice 1.4**

Dans un jeu de pile ou face infini, déterminer la probabilité d'obtenir un nombre fini de piles.

**1.7 Espace probabilisé discret****Définition 1.12 Distribution de probabilités discrète**

On appelle **distribution de probabilités discrète** sur un ensemble  $\Omega$  une famille  $(p_\omega)_{\omega \in \Omega} \in (\mathbb{R}_+)^{\Omega}$  telle que  $\sum_{\omega \in \Omega} p_\omega = 1$ .

**Proposition 1.17 Support d'une distribution de probabilités discrète**

Si  $(p_\omega)_{\omega \in \Omega}$  est une distribution de probabilités discrète, alors son support  $\{\omega \in \Omega, p_\omega \neq 0\}$  est au plus dénombrable.

**Proposition 1.18 Espace probabilisé discret**

Soient  $\Omega$  un univers **au plus dénombrable**.

- Soit  $\mathbb{P}$  une probabilité sur  $(\Omega, \mathcal{P}(\Omega))$ . Alors la famille  $(\mathbb{P}(\{\omega\}))_{\omega \in \Omega}$  est sommable de somme 1.
- Soit  $(p_\omega)_{\omega \in \Omega}$  une distribution de probabilités discrète. Alors il existe une unique probabilité sur  $(\Omega, \mathcal{P}(\Omega))$  telle que  $\mathbb{P}(\{\omega\}) = p_\omega$  pour tout  $\omega \in \Omega$ .

**Exemple 1.7**

On peut définir une probabilité  $\mathbb{P}$  sur  $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$  en posant  $\mathbb{P}(\{n\}) = \frac{1}{2^{n+1}}$  puisque pour tout  $n \in \mathbb{N}$ ,  $\frac{1}{2^{n+1}} \geq 0$  et  $\sum_{n=0}^{+\infty} \frac{1}{2^{n+1}} = 1$ .

**2 Variables aléatoires discrètes****2.1 Définitions****Définition 2.1 Variable aléatoire discrète**

Soit  $(\Omega, \mathcal{A})$  un espace probabilisable. On appelle **variable aléatoire discrète** sur  $(\Omega, \mathcal{A})$  toute application  $X$  définie sur  $\Omega$  telle que

- $X(\Omega)$  est au plus dénombrable ;
- pour tout  $x \in X(\Omega)$ ,  $X^{-1}(\{x\}) \in \mathcal{A}$ .

**Notation 2.1**

Soit  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .

Pour tout  $A \in \mathcal{P}(X(\Omega))$ ,  $X^{-1}(A)$  est un élément de  $\mathcal{A}$  donc un événement.

L'événement  $X^{-1}(A)$  se note plutôt  $\{X \in A\}$ .

L'événement  $X^{-1}(\{x\})$  se note plutôt  $\{X = x\}$ .

On peut alors parler des probabilités  $\mathbb{P}(X \in A)$  et  $\mathbb{P}(X = x)$ .

Si  $X$  est une variable aléatoire réelle et  $x \in \mathbb{R}$ , alors les événements  $\{X \in ]-\infty, x]\}$  et  $\{X \in [x, +\infty[$  se notent plutôt respectivement  $\{X \leq x\}$  et  $\{X \geq x\}$ . On peut alors parler des probabilités  $\mathbb{P}(X \leq x)$  et  $\mathbb{P}(X \geq x)$ .

**REMARQUE.** Si  $X$  est une variable aléatoire discrète sur un espace probabilisable  $(\Omega, \mathcal{A})$ , pour tout  $A \in \mathcal{P}(X(\Omega))$ ,  $A$  est au plus dénombrable en tant que partie d'un ensemble au plus dénombrable et  $\{X \in A\} = \bigsqcup_{x \in A} \{X = x\}$  est un événement en tant qu'union au plus dénombrable d'événements.

On peut en fait montrer que  $\{\{X \in A\}, A \in \mathcal{P}(X(\Omega))\}$  est une tribu sur  $\Omega$  incluse dans la tribu  $\mathcal{A}$ . On l'appelle la tribu engendrée par la variable aléatoire  $X$ .

**Exemple 2.1 Temps d'arrêt**

Soit  $(X_n)$  une suite de variables aléatoires discrètes sur un même espace probabilisable  $(\Omega, \mathcal{A})$  et à valeurs dans un même ensemble. Soit  $x \in E$ . On définit  $T = \min\{n \in \mathbb{N}, X_n = x\}$  ou, peut-être de manière plus rigoureuse

$$T: \begin{cases} \Omega & \longrightarrow \mathbb{N} \cup \{+\infty\} \\ \omega & \longmapsto \min\{n \in \mathbb{N}, X_n(\omega) = x\} \end{cases}$$

**REMARQUE.** On convient que le minimum d'une partie vide de  $\mathbb{N}$  est  $+\infty$ .

Montrons que  $T$  est une variable aléatoire discrète.

Tout d'abord,  $\mathbb{N} \cup \{+\infty\}$  est bien dénombrable. Ensuite,  $\{T = +\infty\} = \bigcap_{n \in \mathbb{N}} \{X_n \neq x\}$ . Pour tout  $n \in \mathbb{N}$ ,  $\{X_n = x\} \in \mathcal{A}$

donc  $\{X_n \neq x\} = \overline{\{X_n = x\}} \in \mathcal{A}$ . Ainsi  $\{T = +\infty\} \in \mathcal{A}$  en tant qu'intersection dénombrable d'éléments de  $\mathcal{A}$ .

Enfin, soit  $n \in \mathbb{N}$ . Alors

$$\{T = n\} = \left( \bigcap_{k=0}^{n-1} \{X_k \neq x\} \right) \cap \{X_n = x\}$$

**REMARQUE.** On convient qu'une intersection indexée sur l'ensemble vide est égale à  $\Omega$  dans le cas  $n = 0$ .

A nouveau,  $\{T = n\} \in \mathcal{A}$  en tant qu'intersection finie d'éléments de  $\mathcal{A}$ .

**Définition 2.2 Loi d'une variable aléatoire**

Soit  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . On appelle **loi** de  $X$  l'application

$$\mathbb{P}_X: \begin{cases} \mathcal{P}(X(\Omega)) & \longrightarrow [0, 1] \\ A & \longmapsto \mathbb{P}(X \in A) \end{cases}$$

**Notation 2.2**

Si  $X$  et  $Y$  sont des variables aléatoires de même loi, on note  $X \sim Y$ .

Lorsque  $X$  suit une loi  $\mathcal{L}$  donnée, on notera également  $X \sim \mathcal{L}$ .

**Proposition 2.1**

Soit  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . Alors la loi  $\mathbb{P}_X$  de  $X$  est une probabilité sur l'espace probabilisable  $(X(\Omega), \mathcal{P}(X(\Omega)))$ .

**Corollaire 2.1**

Soit  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . Alors  $\mathbb{P}_X$  est entièrement déterminée par la donnée des  $\mathbb{P}(X = x)$  pour  $x \in X(\Omega)$ .

**REMARQUE.** C'est pour cela qu'en pratique, lorsque l'on demande la loi d'une variable aléatoire  $X$  sur  $\Omega$ , on ne demande pas  $\mathbb{P}(X \in A)$  pour tout  $A \in \mathcal{P}(X(\Omega))$  mais  $\mathbb{P}(X = x)$  pour tout  $x \in X(\Omega)$ .

**Proposition 2.2 Image d'une variable aléatoire**

Soit  $X$  une variable aléatoire discrète sur un espace probabilisable  $(\Omega, \mathcal{A})$ . Alors pour toute application  $f$  définie sur  $X(\Omega)$ ,  $f(X) = f \circ X$  est une variable aléatoire sur  $(\Omega, \mathcal{A})$ .

**Proposition 2.3**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes. Si  $X \sim Y$ , alors pour toute application  $f$ ,  $f(X) \sim f(Y)$ .

**Définition 2.3 Loi conditionnelle**

Soient  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  et  $B \in \mathcal{A}$  tel que  $\mathbb{P}(B) > 0$ . On appelle **loi de  $X$  conditionnée par l'événement  $B$**  l'application

$$\begin{cases} \mathcal{P}(X(\Omega)) & \longrightarrow [0, 1] \\ A & \longmapsto \mathbb{P}_B(X \in A) \end{cases}$$

**REMARQUE.** A nouveau, quand on demande en pratique la loi conditionnelle de  $X$  conditionnée par l'événement  $B$ , on se contente de donner  $\mathbb{P}_B(X = x)$  pour tout  $x \in X(\Omega)$  plutôt que  $\mathbb{P}_B(X \in A)$  pour tout  $A \in \mathcal{P}(X(\Omega))$ .

**2.2 Lois usuelles****Définition 2.4 Loi géométrique**

Soient  $X$  une variable aléatoire sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  à valeurs dans  $\mathbb{N}^*$  et  $p \in [0, 1]$ . On dit que  $X$  suit la **loi géométrique de paramètre  $p$**  si

$$\forall n \in \mathbb{N}^*, \mathbb{P}(X = n) = (1 - p)^{n-1} p$$

Pour abréger, on dit que  $X$  suit la loi  $\mathcal{G}(p)$  ou encore  $X \sim \mathcal{G}(p)$ .

**Interprétation de la loi géométrique**

On considère une succession d'épreuves de Bernoulli indépendantes de même paramètre  $p$ . Le rang d'apparition du premier succès suit une loi géométrique de paramètre  $p$ .

**Exercice 2.1 Caractérisation de la loi géométrique comme loi sans mémoire**

Soit  $X$  une variable aléatoire discrète à valeurs dans  $\mathbb{N}^*$ . Montrer que  $X$  suit une loi géométrique si et seulement si

$$\forall (m, n) \in \mathbb{N}, \mathbb{P}(X > m + n \mid X > m) = \mathbb{P}(X > n)$$

**Définition 2.5 Loi de Poisson**

Soient  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  à valeurs dans  $\mathbb{N}^*$  et  $\lambda \in \mathbb{R}_+^*$ . On dit que  $X$  suit la **loi de Poisson de paramètre**  $\lambda$  si

$$\forall n \in \mathbb{N}, \mathbb{P}(X = n) = \frac{e^{-\lambda} \lambda^n}{n!}$$

Pour abréger, on dit que  $X$  suit la loi  $\mathcal{P}(\lambda)$  ou encore  $X \sim \mathcal{P}(\lambda)$ .

**Exercice 2.2 Approximation d'une loi binomiale par une loi de Poisson**

Soit  $(X_n)_{n \in \mathbb{N}}$  une suite de variables aléatoires telle que  $X_n \sim \mathcal{B}(n, p_n)$  pour tout  $n \in \mathbb{N}$ . On suppose que  $np_n \xrightarrow{n \rightarrow +\infty} \lambda > 0$ . Montrer que pour tout  $k \in \mathbb{N}$ ,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(X_n = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

**REMARQUE.** On dit que la suite  $(X_n)$  converge en loi vers une variable aléatoire de loi  $\mathcal{P}(\lambda)$ .

**2.3 Couples et uplets de variables aléatoires**

Remarquons tout d'abord que si  $X$  et  $Y$  sont des variables aléatoires discrètes sur un espace probabilisable  $(\Omega, \mathcal{A})$ , il en est de même du couple  $(X, Y)$ , c'est-à-dire de l'application  $\omega \in \Omega \mapsto (X(\omega), Y(\omega))$ .

**Définition 2.6 Loi conjointe, lois marginales**

Soient  $X$  et  $Y$  des variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .

La loi du couple  $(X, Y)$  est appelée **loi conjointe** des variables aléatoires  $X$  et  $Y$ .

Les lois de  $X$  et  $Y$  sont appelées les **lois marginales** du couple  $(X, Y)$ .

**Notation 2.3**

Pour  $(x, y) \in X(\Omega) \times Y(\Omega)$ , on note

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}((X, Y) = (x, y))$$

**Proposition 2.4**

Soient  $X$  et  $Y$  des variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .

La loi conjointe de  $X$  et  $Y$  permet de retrouver les lois marginales. Plus précisément

$$\begin{aligned} \forall x \in X(\Omega), \mathbb{P}(X = x) &= \sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y)) \\ \forall y \in Y(\Omega), \mathbb{P}(Y = y) &= \sum_{x \in X(\Omega)} \mathbb{P}((X, Y) = (x, y)) \end{aligned}$$



**ATTENTION !** Les lois marginales ne permettent pas de retrouver la loi conjointe (à moins que les variables aléatoires soient indépendantes).

### Définition 2.7

Soient  $X_1, \dots, X_n$  sont des variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
La loi du  $n$ -uplet  $(X_1, \dots, X_n)$  est appelée **loi conjointe** des variables aléatoires  $X_1, \dots, X_n$ .  
Les lois des variables aléatoires  $X_1, \dots, X_n$  sont appelées les **lois marginales** du  $n$ -uplet  $(X_1, \dots, X_n)$ .

### Proposition 2.5

Soient  $X_1, \dots, X_n$  des variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
La loi conjointe de  $(X_1, \dots, X_n)$  permet de retrouver les lois marginales. Plus précisément

$$\forall i \in \llbracket 1, n \rrbracket, \forall x_i \in X_i(\Omega), \mathbb{P}(X = x_i) = \sum_{(x_j)_{j \in \llbracket 1, n \rrbracket \setminus \{i\}} \in \prod_{j \in \llbracket 1, n \rrbracket \setminus \{i\}} X_j(\Omega)} \mathbb{P} \left( \bigcap_{j \in \llbracket 1, n \rrbracket \setminus \{i\}} \{X_j = x_j\} \right)$$

## 2.4 Variables aléatoires indépendantes

### Définition 2.8 Couple de variables aléatoires indépendantes

Soient  $X$  et  $Y$  des variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
Les variables aléatoires  $X$  et  $Y$  sont dites **indépendantes** si

$$\forall (A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)), \mathbb{P}((X, Y) \in A \times B) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B)$$

### Proposition 2.6

Soient  $X$  et  $Y$  des variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
Les variables aléatoires  $X$  et  $Y$  sont indépendantes si et seulement si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(X = x) \mathbb{P}(Y = y)$$

### Notation 2.4

Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes, on note  $X \perp\!\!\!\perp Y$ .

### Définition 2.9 Famille de variables aléatoires mutuellement indépendantes

Soit  $(X_i)_{i \in I}$  une famille **finie** de variables aléatoires discrètes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
On dit que  $(X_i)_{i \in I}$  est une famille de variables aléatoires **mutuellement indépendantes** si

$$\forall (A_i)_{i \in I} \in \prod_{i \in I} \mathcal{P}(X_i(\Omega)), \mathbb{P} \left( \bigcap_{i \in I} X_i \in A_i \right) = \prod_{i \in I} \mathbb{P}(X_i \in A_i)$$

Une famille **infinie** de variables aléatoires est une famille de variables aléatoires mutuellement indépendantes si toute sous-famille finie l'est.

**Proposition 2.7**

Soit  $(X_i)_{i \in I}$  une famille **finie** de variables aléatoires sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
 $(X_i)_{i \in I}$  est une famille de variables aléatoires mutuellement indépendantes si et seulement si

$$\forall (x_i)_{i \in I} \in \prod_{i \in I} X_i(\Omega), \quad \mathbb{P}\left(\bigcap_{i \in I} X_i = x_i\right) = \prod_{i \in I} \mathbb{P}(X_i = x_i)$$

Soit  $(X_i)_{i \in I}$  une famille **infinie** de variables aléatoires sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
 $(X_i)_{i \in I}$  est une famille de variables aléatoires mutuellement indépendantes si et seulement si pour toute partie **finie**  $J$  de  $I$

$$\forall (x_i)_{i \in J} \in \prod_{i \in J} X_i(\Omega), \quad \mathbb{P}\left(\bigcap_{i \in J} X_i = x_i\right) = \prod_{i \in J} \mathbb{P}(X_i = x_i)$$

**REMARQUE.** Toute sous-famille d'une famille finie de variables mutuellement indépendantes est une famille de variables aléatoires mutuellement indépendantes.



**ATTENTION !** Si des variables aléatoires sont mutuellement indépendantes, alors ces variables aléatoires sont indépendantes deux à deux mais la réciproque est fausse.

**Exemple 2.2**

Soit  $X$  et  $Y$  deux variables aléatoires indépendantes de loi uniforme sur  $\{-1, 1\}$ . On pose  $Z = XY$ . Alors  $X, Y, Z$  sont deux à deux indépendantes mais pas mutuellement indépendantes.

**Exercice 2.3**

Montrer que  $(A_i)_{i \in I}$  est une famille d'événements mutuellement indépendants si et seulement si  $(\mathbb{1}_{A_i})_{i \in I}$  est une famille de variables aléatoires mutuellement indépendantes.

**Proposition 2.8**

Si  $X$  et  $Y$  sont deux variables aléatoires indépendantes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ , alors, pour toutes applications  $f$  et  $g$  définies respectivement sur  $X(\Omega)$  et  $Y(\Omega)$ ,  $f(X)$  et  $g(Y)$  sont indépendantes.

**REMARQUE.** De manière plus condensée,  $X \perp\!\!\!\perp Y \implies f(X) \perp\!\!\!\perp g(Y)$ .

**Corollaire 2.2**

Soient  $(X_i)_{i \in I}$  une famille de variables aléatoires mutuellement indépendantes et pour tout  $i \in I$ , une application  $f_i$  définie sur  $X_i(\Omega)$ . Alors  $(f_i(X_i))_{i \in I}$  est une famille de variables aléatoires mutuellement indépendantes.

**Proposition 2.9 Lemme des coalitions**

Soient  $X_1, \dots, X_n$  des variables aléatoires discrètes mutuellement indépendantes sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
 Soient  $m \in \llbracket 1, n-1 \rrbracket$ ,  $f$  une application définie sur  $\prod_{k=1}^m X_k(\Omega)$  et  $g$  une application définie sur  $\prod_{k=m+1}^n X_k(\Omega)$ . Alors,  $f(X_1, \dots, X_m)$  et  $g(X_{m+1}, \dots, X_n)$  sont des variables aléatoires indépendantes.

**Exemple 2.3**

Si  $X$ ,  $Y$  et  $Z$  sont trois variables aléatoires discrètes réelles mutuellement indépendantes, alors  $\cos(X^2 + Y^2)$  et  $e^Z$  sont indépendantes.

**Corollaire 2.3**

Soient  $(X_i)_{i \in I}$  est une famille de variables aléatoires mutuellement indépendantes,  $I = \bigsqcup_{j \in J} I_j$  une partition de  $I$  et pour tout  $j \in J$ ,  $f_j$  une application définie sur  $\prod_{i \in I_j} X_i(\Omega)$ . Alors la famille  $(f_j((X_i)_{i \in I_j}))_{j \in J}$  est une famille de variables aléatoires mutuellement indépendantes.

**Proposition 2.10**

Soit  $(\mathcal{L}_n)_{n \in \mathbb{N}}$  une suite de lois discrètes. Alors il existe un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  et une suite  $(X_n)_{n \in \mathbb{N}}$  de variables aléatoires mutuellement indépendantes sur  $(\Omega, \mathcal{A}, \mathbb{P})$  telles que  $\forall n \in \mathbb{N}, X_n \sim \mathcal{L}_n$ .

**REMARQUE.** Une suite de variables aléatoires indépendantes et de même loi est appelée une suite de variables aléatoires **indépendantes et identiquement distribuées** (en abrégé, suites de variables aléatoires i.i.d.).

**Jeu de pile ou face infini**

La proposition précédente permet de donner un cadre théorique au jeu de pile ou face infini. En effet, on peut affirmer l'existence d'un espace probabilisé pour lequel il existe une suite de variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre donné (1/2 si la pièce lancée est équilibrée).

### 3 Espérance, variance et covariance

Le corps  $\mathbb{K}$  désigne  $\mathbb{R}$  ou  $\mathbb{C}$ .

#### 3.1 Espérance

**Définition 3.1** Espérance d'une variable aléatoire discrète réelle positive

Soit  $X$  une variable aléatoire discrète sur un espace probabilisé  $(\Omega, \mathcal{A}, \mathbb{P})$ . On suppose que  $X(\Omega) \subset \mathbb{R}_+ \cup \{+\infty\}$ . L'**espérance** de  $X$  est la somme, dans  $\mathbb{R}_+ \cup \{+\infty\}$ , de la famille  $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$ . On la note  $\mathbb{E}(X)$ .

**Exercice 3.1**

Soit  $X$  une variable aléatoire discrète positive. Montrer que  $\mathbb{E}(X) = 0$  si et seulement si  $X$  est presque sûrement nulle i.e.  $\mathbb{P}(X = 0) = 1$ .

**Proposition 3.1** Formule d'antirépartition

Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N} \cup \{+\infty\}$ . Alors

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n)$$

**Définition 3.2** Espérance d'une variable aléatoire discrète réelle ou complexe

Soit  $X$  une variable aléatoire discrète sur un espace probabibilisé  $(\Omega, \mathcal{A}, \mathbb{P})$  à valeurs dans  $\mathbb{K}$ . On dit que  $X$  est **d'espérance finie** si la famille  $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$  est sommable. Dans ce cas, la somme de cette famille est l'**espérance** de  $X$ . On la note  $\mathbb{E}(X)$ .

**REMARQUE.** Il est primordial de remarquer que l'espérance d'une variable aléatoire ne dépend que de sa loi.

**REMARQUE.** Si  $X(\Omega)$  est fini,  $X$  admet toujours une espérance finie.

Si  $X(\Omega)$  est dénombrable, on peut écrire  $X(\Omega) = \{x_n, n \in \mathbb{N}\}$ . Alors  $X$  est d'espérance finie si et seulement si la série

$$\sum_{n \in \mathbb{N}} x_n \mathbb{P}(X = x_n) \text{ converge absolument et, dans ce cas, } \mathbb{E}(X) = \sum_{n=0}^{+\infty} x_n \mathbb{P}(X_n).$$

**REMARQUE.** Une variable aléatoire d'espérance finie est dite centrée si  $\mathbb{E}(X) = 0$ .

**Notation 3.1**

Si  $X$  est d'espérance finie, on note  $X \in L^1$ .

**Proposition 3.2** Espérance des lois usuelles

- Si  $X \sim \mathcal{G}(p)$ , alors  $\mathbb{E}(X) = \frac{1}{p}$ .
- Si  $X \sim \mathcal{P}(\lambda)$ , alors  $\mathbb{E}(X) = \lambda$ .

**Proposition 3.3** Formule de transfert

Soient  $X$  une variable aléatoire discrète et  $f : X(\Omega) \rightarrow \mathbb{K}$ . Alors  $f(X)$  est d'espérance finie si et seulement si la famille  $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$  est sommable et, dans ce cas,

$$\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x) \mathbb{P}(X = x)$$

**REMARQUE.** L'intérêt de la formule de transfert est qu'elle permet de calculer l'espérance de  $f(X)$  sans connaître la loi de  $f(X)$  mais uniquement la loi de  $X$ .

**REMARQUE.** D'après la formule de transfert appliquée à la valeur absolue ou au module,  $X \in L^1 \iff \mathbb{E}(|X|) < +\infty$ .

**Exercice 3.2**

Soit  $X$  une variable suivant la loi géométrique de paramètre  $p$ . Déterminer l'espérance de  $1/X$ .



**Proposition 3.4 Propriétés de l'espérance**

**Linéarité** Soient  $X$  et  $Y$  deux variables aléatoires discrètes complexes d'espérances finies. Soit  $(\lambda, \mu) \in \mathbb{K}^2$ . Alors  $\lambda X + \mu Y$  est d'espérance finie et  $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$ .

**Positivité** Soit  $X$  une variable aléatoire discrète positive. Alors  $\mathbb{E}(X) \geq 0$ .

**Croissance** Soient  $X$  et  $Y$  deux variables aléatoires discrètes réelles d'espérances finies. Si  $X \leq Y$ , alors  $\mathbb{E}(X) \leq \mathbb{E}(Y)$ .

**Inégalité triangulaire** Soit  $X$  une variable aléatoire d'espérance finie. Alors  $|X|$  est d'espérance finie et  $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$ .

**Proposition 3.5**

L'ensemble des variables aléatoires discrètes de  $L^1$  à valeurs dans  $\mathbb{K}$  forment un  $\mathbb{K}$ -espace vectoriel.

**Proposition 3.6**

Soient  $X$  une variable aléatoire à valeurs dans  $\mathbb{K}$  et  $Y$  une variable aléatoire réelle telles que  $|X| \leq Y$ . Si  $Y \in L^1$ , alors  $X \in L^1$ .

**Exemple 3.1**

Si  $X$  est une variable aléatoire bornée, il existe  $M \in \mathbb{R}_+$  tel que  $|X| \leq M$ . Il est clair qu'une variable aléatoire constante est d'espérance finie. On en déduit que  $X \in L^1$ .

**Proposition 3.7 Espérance et indépendance**

Soient  $X$  et  $Y$  deux variables aléatoires discrètes **indépendantes**, à valeurs dans  $\mathbb{K}$  et d'espérances finies. Alors  $XY$  est d'espérance finie et

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

**Corollaire 3.1 Espérance et indépendance**

Soient  $X_1, \dots, X_n$  des variables aléatoires discrètes **mutuellement indépendantes**, à valeurs dans  $\mathbb{K}$  et d'espérances finies.

Alors  $\prod_{i=1}^n X_i$  est d'espérance finie et

$$\mathbb{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbb{E}(X_i)$$

## 3.2 Variance

### Notation 3.2

Si  $X$  est une variable aléatoire réelle discrète telle que  $X^2$  est d'espérance finie, on note  $X \in L^2$ . Autrement dit,  $X \in L^2 \iff X^2 \in L^1$ .

### Proposition 3.8

Soit  $X$  une variable aléatoire réelle discrète telle que  $X^2$  est d'espérance finie. Alors  $X$  est d'espérance finie.

**REMARQUE.** De manière plus condensée,  $X \in L^2 \implies X \in L^1$ .

### Proposition 3.9

L'ensemble des variables aléatoires discrètes réelles de  $L^2$  forment un  $\mathbb{R}$ -espace vectoriel.

### Exercice 3.3

Montrer que l'application  $(X, Y) \in (L^2)^2 \mapsto \mathbb{E}(XY)$  est bien définie, bilinéaire, symétrique et positive.

### Proposition 3.10 Inégalité de Cauchy-Schwarz

Soit  $(X, Y) \in (L^2)^2$ . Alors  $XY \in L^1$  et

$$\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$$

### Exercice 3.4 Cas d'égalité dans l'inégalité de Cauchy-Schwarz

Soit  $(X, Y) \in (L^2)^2$  telle que  $\mathbb{E}(XY)^2 = \mathbb{E}(X^2)\mathbb{E}(Y^2)$ . Montrer que  $X$  et  $Y$  sont presque sûrement colinéaires, autrement dit qu'il existe  $\lambda \in \mathbb{R}$  tel que  $\mathbb{P}(X = \lambda Y) = 1$  ou  $\mathbb{P}(Y = \lambda X) = 1$ .

### Définition 3.3 Variance et écart-type

Soit  $X \in L^2$ .

- On appelle **variance** de  $X$  le réel positif

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

- On appelle **écart-type** de  $X$  le réel

$$\sigma(X) = \sqrt{\mathbb{V}(X)}$$

**REMARQUE.** La variance et l'écart-type d'une variable aléatoire ne dépendent que de sa **loi**.

**REMARQUE.** On dit qu'une variable aléatoire  $X$  est **réduite** si  $\mathbb{V}(X) = 1$ .

### Exercice 3.5

Soit  $X \in L^2$  une variable aléatoire discrète réelle. Montrer que  $\mathbb{V}(X) = 0$  si et seulement si  $X$  est presque sûrement constante.

### Proposition 3.11 Théorème de König-Huygens

Soit  $X \in L^2$ . Alors

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

### Proposition 3.12 Variance d'une fonction affine d'une variable aléatoire

Soient  $X \in L^2$  et  $(a, b) \in \mathbb{R}^2$ . Alors  $\mathbb{V}(aX + b) = a^2\mathbb{V}(X)$ .

**REMARQUE.** Si  $X$  est une variable aléatoire réelle admettant une variance finie non nulle, alors  $\frac{X - \mathbb{E}(X)}{\sigma(X)}$  est **centrée réduite**.

### Proposition 3.13 Variance des lois usuelles

- Si  $X \sim \mathcal{G}(p)$ , alors  $\mathbb{V}(X) = \frac{1-p}{p^2}$ .
- Si  $X \sim \mathcal{P}(\lambda)$ , alors  $\mathbb{V}(X) = \lambda$ .

## 3.3 Covariance

### Définition 3.4 Covariance de deux variables aléatoires

Soit  $(X, Y) \in (L^2)^2$ . On appelle **covariance de  $X$  et  $Y$**  le réel

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

**REMARQUE.**  $\mathbb{V}(X) = \text{Cov}(X, X)$

**REMARQUE.** D'après l'inégalité de Cauchy-Schwarz :

$$\text{Cov}(X, Y)^2 \leq \mathbb{V}(X)\mathbb{V}(Y)$$

ou encore

$$|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$$

**REMARQUE.** La covariance est «presque» un produit scalaire. C'est une forme bilinéaire symétrique positive mais pas nécessairement définie.

**REMARQUE.** Si  $(X, Y) \in (L^2)^2$ , alors, en appliquant l'inégalité de Cauchy-Schwarz à  $X - \mathbb{E}(X)$  et  $Y - \mathbb{E}(Y)$ , on obtient

$$\text{Cov}(X, Y)^2 \leq \mathbb{V}(X)\mathbb{V}(Y)$$

ou encore

$$|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$$

### Proposition 3.14

Soit  $(X, Y) \in (L^2)^2$ . Alors

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

### Corollaire 3.2 Covariance de variables aléatoires

Si  $X$  et  $Y$  sont deux variables aléatoires discrètes réelles **indépendantes** de  $L^2$ , alors  $\text{Cov}(X, Y) = 0$ .



**ATTENTION!** La réciproque est fausse. Il suffit par exemple de prendre  $X \sim \mathcal{U}(\{-1, 0, 1\})$  et  $Y = 1 - X^2$ . On a bien  $\text{Cov}(X, Y) = 0$  mais  $X$  et  $Y$  ne sont évidemment pas indépendantes.

### Proposition 3.15 Variance d'une somme et théorème de Pythagore

Soient  $X_1, \dots, X_n$  des variables aléatoires réelles discrètes de  $L^2$ . Alors

$$\mathbb{V}\left(\sum_{k=1}^n X_k\right) = \sum_{1 \leq i, j \leq n} \text{Cov}(X_i, X_j)$$

Si de plus,  $\text{Cov}(X_i, X_j) = 0$  pour tout  $(i, j) \in \llbracket 1, n \rrbracket^2$  tel que  $i \neq j$ ,

$$\mathbb{V}\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n \mathbb{V}(X_k)$$

**REMARQUE.** Notamment si  $X_1, \dots, X_n$  sont des variables aléatoires **deux à deux indépendantes** de  $L^2$ , alors

$$\mathbb{V}\left(\sum_{k=1}^n X_k\right) = \sum_{k=1}^n \mathbb{V}(X_k)$$

## 3.4 Inégalités classiques

### Exercice 3.6

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé. Soit  $A \in \mathcal{A}$ . Montrer que la fonction indicatrice de  $A$ , à savoir l'application

$$\mathbb{1}_A : \begin{cases} \Omega & \longrightarrow \{0, 1\} \\ \omega & \longmapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases} \end{cases}$$

suit la loi de Benoulli de paramètre  $\mathbb{P}(A)$ .

**Proposition 3.16 Inégalité de Markov**

Soient  $X$  une variable aléatoire discrète réelle **positive** d'espérance finie et  $a \in \mathbb{R}_+^*$ . Alors

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Si on ne suppose plus  $X$  positive, on peut affirmer que

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}$$

**Corollaire 3.3 Inégalité de Bienaymé-Tchebychev**

Soient  $X \in L^2$  et  $\varepsilon \in \mathbb{R}_+^*$ . Alors

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2} = \frac{\sigma(X)^2}{\varepsilon^2}$$

**Corollaire 3.4 Loi faible des grands nombres**

Soit  $(X_n)_{n \in \mathbb{N}^*}$  une suite de variables aléatoires deux à deux indépendantes, de même loi et admettant des variances finies.

Alors, en posant  $S_n = \sum_{k=1}^n X_k$  et en notant  $m$  l'espérance commune des  $X_k$ ,

$$\forall \varepsilon \in \mathbb{R}_+^*, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0$$

**REMARQUE.** On peut en fait prouver à l'aide de l'inégalité de Bienaymé-Tchebychev qu'en notant  $\sigma^2$  la variance commune des  $X_k$ ,

$$\forall \varepsilon \in \mathbb{R}_+^*, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

La loi faible des grands nombres s'en déduit immédiatement.

## 4 Fonctions génératrices

Si  $X$  est une variable aléatoire à valeurs dans  $\mathbb{N}$ , alors pour tout  $t \in [-1, 1]$ ,  $|t^X| \leq 1$  donc  $t^X \in L^1$ . On peut alors définir  $\mathbb{E}(t^X)$ .

**Définition 4.1 Fonction génératrice**

Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$ . On appelle fonction génératrice de  $X$  la fonction  $G_X$  définie par

$$G_X(t) = \mathbb{E}(t^X) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)t^k$$

**REMARQUE.** A nouveau, la fonction génératrice d'une variable aléatoire ne dépend que de sa loi.

**REMARQUE.** Puisque la série  $\sum_{k \in \mathbb{N}} \mathbb{P}(X = k)$  converge (vers 1),

- la série entière  $\sum_{k \in \mathbb{N}} \mathbb{P}(X = k)t^k$  a un rayon de convergence supérieur ou égal à 1 ;
- elle converge normalement sur le disque fermé de centre 0 et de rayon 1 ;
- sa somme  $G_X$  est continue sur ce même disque ;
- $G_X$  est de classe  $\mathcal{C}^\infty$  sur  $] -1, 1[$ .

#### Fonctions génératrices des lois usuelles

- Si  $X \sim \mathcal{B}(p)$ ,  $G_X(t) = 1 - p + pt$ .
- Si  $X \sim \mathcal{B}(n, p)$ ,  $G_X(t) = (1 - p + pt)^n$ .
- Si  $X \sim \mathcal{G}(p)$ ,  $G_X(t) = \frac{pt}{1 - (1 - p)t}$ .
- Si  $X \sim \mathcal{P}(\lambda)$ ,  $G_X(t) = e^{\lambda(t-1)}$ .

#### Proposition 4.1 Fonction génératrice et loi

Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$ . Alors

$$\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{G_X^{(k)}(0)}{k!}$$

#### Proposition 4.2

Deux variables aléatoires à valeurs dans  $\mathbb{N}$  sont égales si et seulement si elles ont la même fonction génératrice.

#### Proposition 4.3 Fonction génératrice et espérance

Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$ . Alors  $X$  admet une espérance finie si et seulement si  $G_X$  est dérivable en 1 et, dans ce cas,  $\mathbb{E}(X) = G'_X(1)$ .

**REMARQUE.** Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{N}$ . Si  $G_X$  est deux fois dérivable en 1, alors  $X \in L^2$  et  $\mathbb{V}(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$ .

#### Proposition 4.4 Fonction génératrice d'une somme finie de variables aléatoires indépendantes

Soient  $X_1, \dots, X_n$  des variables aléatoires **mutuellement indépendantes** à valeurs dans  $\mathbb{N}$ . On pose  $S = \sum_{k=1}^n X_k$ . Alors

$$G_S = \prod_{k=1}^n G_{X_k}$$

**Exemple 4.1**

Si  $(X_n)_{n \in \mathbb{N}^*}$  est une suite de variables aléatoires indépendantes et de même loi, alors pour tout  $n \in \mathbb{N}^*$  et tout  $m \in \mathbb{N}$ ,  $S = \sum_{k=1}^n X_k$  et  $T = \sum_{k=m+1}^{m+n} X_k$  ont la même loi. En effet,  $G_S = G_T = (G_{X_1})^n$ .

**Exercice 4.1**

Soient  $X_1, \dots, X_n$  des variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs  $\lambda_1, \dots, \lambda_n$ . Déterminer la loi de  $S = \sum_{k=1}^n X_k$ .