# Résumé du cours Probabilités 1 S3
# Université de Strasbourg
# L2 Mathématiques / Mathématiques et Physique Approfondies

Davide Giraudo

13 septembre 2024

## Table des matières

# 1 Espaces probabilisés

## 1.1 Terminologie

### 1.1.1 Réalisation, événements

On considère le lancer d'un dé à six faces.

- Un résultat possible est appelé une réalisation ; il est noté $\omega$.
- L'ensemble de tous les résultats possibles est appelé espace d'états et est notée $\Omega$. Ici : $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Un événement est une collection de résultats. Par exemple, si $A$ est l'événement "le résultat du lancer est pair", $A = \{2, 4, 6\} \subset \Omega$.
- L'événement contraire de $A$ est noté $A^c$ ; dans l'exemple précédent,

$$A^c = \Omega \setminus A = \{1, 2, 3, 4, 5, 6\} \setminus \{2, 4, 6\} = \{1, 3, 5\} .$$

Événements particuliers

- L'événement certain : $\Omega$.
- L'événement impossible : $\emptyset$.
- L'événement $A$ et $B$ : $A \cap B$.
- L'événement $A$ ou $B$ ("ou" non exclusif) : $A \cup B$.

**Définition 1.1.** *Les événements $A$ et $B$ sont incompatibles si $A \cap B = \emptyset$.*

## 1.2 Ensembles

### 1.2.1 Ensembles finis/ ensembles dénombrables

**Définition 1.2** (Ensemble fini)**.** *Un ensemble $\Omega$ est fini s'il existe un entier $n \in \mathbb{N}^*$ et une bijection de $\Omega$ dans $\{1, \ldots, n\}$. On dit alors que* $\operatorname{card}(\Omega) = n$.

**Définition 1.3** (Ensemble dénombrable)**.** *On dit que $\Omega$ est dénombrable s'il existe une bijection de $\Omega$ dans $\mathbb{N}$.*

**Définition 1.4** (Ensemble au plus dénombrable)**.** *Un ensemble est dit au plus dénombrable s'il est fini ou dénombrable.*

*Exemples* 1.1.
- $\mathbb{Z}$, $\mathbb{N} \times \mathbb{N}$ et $\mathbb{Q}$ sont dénombrables.
- L'intervalle $[0, 1]$ n'est pas dénombrable.

Dans la suite, on suppose $\Omega$ au plus dénombrable et on pose $\mathcal{F} = \mathcal{P}(\Omega) = \{A : A \subset \Omega\}$.

### 1.2.2 Rappel sur les opérations ensemblistes

**Définition 1.5** (Partition)**.** *On dit que* $(A_i)_{i\in I}$ *forme une partition de* $A$ *si pour tous* $i, i' \in I$ *tels que* $i \neq i'$, $A_i \cap A_{i'} = \emptyset$ *et* $\bigcup_{i\in I} A_i = A$.

**Proposition 1.6.** *Soient* $\Omega$ *un ensemble,* $I$ *un ensemble d'indices,* $(A_i)_{i\in I}$ *une collection de sous-ensembles de* $\Omega$ *et* $B \subset \Omega$. *Alors*

$$\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c, \quad \left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} A_i^c \tag{1.2.1}$$

$$B \cap \bigcap_{i\in I} A_i = \bigcap_{i\in I} (B \cap A_i)\,, \quad B \cap \bigcup_{i\in I} A_i = \bigcup_{i\in I} (B \cap A_i)\,. \tag{1.2.2}$$

## 1.3 Probabilités

**Définition 1.7** (Probabilité)**.** *Une probabilité* $\mathbb{P}$ *est une fonction définie sur une tribu* $\mathcal{F}$ *à valeurs dans* $[0,1]$ *telle que*

1. $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$ *et*
2. *si* $(A_i)_{i\in I}$ *est une collection au plus dénombrable d'éléments de* $\mathcal{F}$ *deux à deux disjoints* ($A_i \cap A_{i'} = \emptyset$ *si* $i \neq i'$)*, alors*

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mathbb{P}(A_i)\,. \tag{1.3.1}$$

   *[cette propriété est appelée* $\sigma$*-additivité]*

**Définition 1.8** (Espaces probabilisés)**.** *Soient* $\Omega$ *un ensemble,* $\mathcal{F} = \mathcal{P}(\Omega) = \{A : A \subset \Omega\}$ *et* $\mathbb{P}\colon \mathcal{F} \to [0,1]$ *une probabilité. Le triplet* $(\Omega, \mathcal{F}, \mathbb{P})$ *est appelé espace probabilisé.*

*Exemple* 1.9. Lancer d'un dé équilibré : $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ et $\mathbb{P}(\{\omega\}) = 1/6$ pour tout $\omega \in \Omega$.

**Définition 1.10.** *Un événement* $A \in \mathcal{F}$ *est :*

- *presque sûr si* $\mathbb{P}(A) = 1$ *;*
- *négligeable si* $\mathbb{P}(A) = 0$.

**Proposition 1.11** (Propriétés des mesures de probabilité)**.**

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ *;*
2. *Si* $A, B \in \mathcal{F}$ *et* $A \subset B$*, alors* $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ *;*
3. *monotonie : si* $A, B \in \mathcal{F}$ *et* $A \subset B$*, alors* $\mathbb{P}(A) \leqslant \mathbb{P}(B)$ *;*
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ *;*
5. *soit* $(A_i)_{i\in I}$ *une collection au plus dénombrable d'ensembles deux à deux disjoints et telle que* $\sum_{i\in I} \mathbb{P}(A_i) = 1$. *Alors pour* $B \in \mathcal{F}$,

$$\mathbb{P}(B) = \sum_{i\in I} \mathbb{P}(A_i \cap B)\,. \tag{1.3.2}$$

6. *Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements ($A_n \subset A_{n+1}$ pour tout $n$). Alors*

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n) \qquad \textit{convergence monotone.} \tag{1.3.3}$$

## 1.4 Probabilité sur un ensemble fini ou dénombrable

**Lemme 1.12.** *Pour tout $A \in \mathcal{F}$,*

$$\mathbb{P}(A) = \sum_{\omega\in A} \mathbb{P}(\{\omega\}). \tag{1.4.1}$$

**Définition 1.13.** *Si $\Omega$ fini et pour tout $\omega \in \Omega$, $\mathbb{P}(\{\omega\}) = 1/\operatorname{card}(\Omega)$, alors on dit que $\mathbb{P}$ est la probabilité uniforme.*

Par le Lemme 1.12, si $\mathbb{P}$ est la probabilité uniforme, alors pour tout $A \subset \Omega$,

$$\mathbb{P}(A) = \frac{\operatorname{card}(A)}{\operatorname{card}(\Omega)} = \frac{\text{nombre de cas favorables à } A}{\text{nombre de cas possibles}}. \tag{1.4.2}$$

## 1.5 Dénombrement

**Définition 1.14** (Permutations)**.** *Le nombre de permutations de $\{1,\dots,n\}$ dans $\{1,\dots,n\}$ (nombre de bijections d'un ensemble de $n$ éléments dans un ensemble de $n$ éléments) est*

$$n! = \prod_{k=1}^{n} k = n \times (n-1) \times \cdots \times 2 \times 1, \tag{1.5.1}$$

*avec la convention $0! = 1$.*

**Définition 1.15** (Arrangements)**.** *Le nombre d'arrangements de $k$ objets parmi $n$ avec $k \leqslant n$ est noté $A_n^k$. Il correspond au nombre d'injections de $\{1,\dots,k\}$ dans $\{1,\dots,n\}$ :*

$$A_n^k = \frac{n!}{(n-k)!}. \tag{1.5.2}$$

**Définition 1.16** (Coefficient binomial)**.** *Le nombre de sous-ensembles à $k$ éléments d'un ensemble à $n$ éléments, $0 \leqslant k \leqslant n$, est le coefficient binomial*

$$C_n^k = \frac{n!}{k!\,(n-k)!}. \tag{1.5.3}$$

Formule du binôme de Newton :

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}. \tag{1.5.4}$$

## 1.6 Probabilités conditionnelles

**Définition 1.17** (Probabilités conditionnelles)**.** *Soient $A$ et $B$ deux événements tels que $\mathbb{P}(B) > 0$. La probabilité conditionnelle de $A$ sachant $B$, notée $\mathbb{P}(A \mid B)$, est définie par*

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}. \tag{1.6.1}$$

Si $A$ et $B$ sont deux événements tels que $\mathbb{P}(B) > 0$ alors

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\,\mathbb{P}(B) \tag{1.6.2}$$

**Proposition 1.18.** *Si $A$ et $B$ sont deux événements tels que $0 < \mathbb{P}(B) < 1$ alors on a*

$$\mathbb{P}(A) = \mathbb{P}(A \mid B)\,\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\,\mathbb{P}(B^c) \tag{1.6.3}$$

*et la formule de Bayes*

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B)\,\mathbb{P}(B)}{\mathbb{P}(A \mid B)\,\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\,\mathbb{P}(B^c)}. \tag{1.6.4}$$

## 1.7 Indépendance

**Définition 1.19** (Indépendance de deux événements)**.** *On dit que deux événements $A$ et $B$ sont indépendants si*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B). \tag{1.7.1}$$

*Remarque* 1.20. Si $A$ et $B$ sont indépendants et $\mathbb{P}(B) > 0$, alors $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

**Définition 1.21.** *On dit qu'une collection au plus dénombrable d'événements $(A_i)_{i\in I}$ est indépendante si pour tout sous-ensemble fini $J$ de $I$,*

$$\mathbb{P}\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J} \mathbb{P}(A_j). \tag{1.7.2}$$

*Remarque* 1.22. Si $(A_i)_{i\in I}$ est indépendante et $I' \subset I$, alors $(A_i)_{i\in I'}$ est indépendante.

*Exemple* 1.23 (Lancer de dés). On lance deux dés distingables. L'espace d'états pour le premier dé est $\Omega_1 = \{1, \dots, 6\}$ avec la probabilité uniforme $\mathbb{P}_1$ et pour le second $\Omega_2 = \{1, \dots, 6\}$ avec la probabilité uniforme $\mathbb{P}_2$. L'espace d'états associé au lancer des deux dés est

$$\Omega = \Omega_1 \times \Omega_2 = \{\omega = (\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}. \tag{1.7.3}$$

Une probabilité $\mathbb{P}$ sur $\Omega$ est déterminée par $\mathbb{P}(\{(\omega_1, \omega_2)\})\,;\ \omega_1 \in \Omega_1, \omega_2 \in \Omega_2$. Comme les résultats des deux lancers sont indépendants,

$$\mathbb{P}(\{(\omega_1, \omega_2)\}) = \mathbb{P}_1(\{\omega_1\})\,\mathbb{P}_2(\{\omega_2\}). \tag{1.7.4}$$

On dit que $\mathbb{P}$ est la probabilité produit, et on note $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$.

Plus généralement, les espaces produit permettent de modéliser des expériences indépendantes.

Exemple : lancer de dés Avec les mêmes notations que dans l'exemple précédent, si $A = E_1 \times \Omega_2$ et $B = \Omega_1 \times E_2$, avec $E_1 \subset \Omega_1$, $E_2 \subset \Omega_2$, alors $A$ et $B$ sont indépendants (intuitivement, la réalisation de $A$ ne dépend que du résultat du lancer du premier dé et celle de $B$ seulement de celle du second).

Mais si $A = \{(1,1), (1,2)\}$ et $B = \{(1,1), (1,6)\}$, alors

$$\mathbb{P}(A) = \frac{2}{36} = \mathbb{P}(B) \tag{1.7.5}$$

et

$$\mathbb{P}(A \cap B) = \mathbb{P}(\{(1,1)\}) = \frac{1}{36} \neq \mathbb{P}(A)\,\mathbb{P}(B)\,. \tag{1.7.6}$$

Par conséquent, $A$ et $B$ ne sont pas indépendants.

# 2 Variables aléatoires discrètes

## 2.1 Loi d'une variable aléatoire discrète

**Définition 2.1** (Variable aléatoire). *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé où $\Omega$ est au plus dénombrable. Une variable aléatoire discrète $X$ est une application $X\colon \Omega \to E$ où $E$ est au plus dénombrable.*

**Définition 2.2** (Loi d'une variable aléatoire). *Pour tout $A \subset E$, on note*

$$\mathbb{P}_X(A) = \mathbb{P}(\{\omega \in \Omega, X(\omega) \in A\}) \tag{2.1.1}$$

*L'application $\mathbb{P}_X\colon \mathcal{P}(E) \to [0,1]$ est une probabilité sur $E$ et est appelée loi de la variable aléatoire $X$.*

- Si $E = \{x_1, \ldots, x_N\}$ est fini, soit $p_k = \mathbb{P}(X = x_k) = \mathbb{P}(\{\omega \in \Omega; X(\omega) = x_k\})$.
- Si $E$ est dénombrable, soit $\tau\colon E \to \mathbb{N}$ une bijection, $x_k = \tau^{-1}(k)$ et $p_k = \mathbb{P}(X = x_k)$ (si $E = \mathbb{N}$, on a simplement $p_k = \mathbb{P}(X = k)$).

On a

$$\mathbb{P}_X(A) = \sum_{k : k \in A} p_k. \tag{2.1.2}$$

## 2.2 Moments d'une variable aléatoire discrète

### 2.2.1 Espérance d'une variable aléatoire discrète

**Définition 2.3** (Espérance). *Soient $E \subset \mathbb{R}$ un ensemble au plus dénombrable et $X\colon \Omega \to E$ une variable aléatoire discrète. Si*

$$\sum_{x \in E} |x|\,\mathbb{P}(X = x) < \infty, \tag{2.2.1}$$

*on dit que la variable aléatoire $X$ admet un moment d'ordre un et on pose*

$$\mathbb{E}[X] = \sum_{x \in E} x\mathbb{P}(X = x)\,. \tag{2.2.2}$$

*La quantité $\mathbb{E}[X]$ est appelée espérance de $X$.*

### 2.2.2 Moments d'ordre supérieur

**Définition 2.4.** *On dit que la variable aléatoire réelle discrète $X$ admet un moment d'ordre $n \in \mathbb{N}^*$ si $X^n$ admet un moment d'ordre 1.*

*Remarque* 2.5. Si $X$ est bornée, au sens où il existe une constante $C$ telle que pour tout $\omega \in \Omega$, $|X(\omega)| \leqslant C$, alors $X$ admet un moment d'ordre $n$ pour chaque $n$.

Si $X$ prend les valeurs $x_1, \dots, x_N$, alors

$$\mathbb{E}[X^n] = \sum_{k=1}^{N} x_k^n \mathbb{P}(X = x_k). \tag{2.2.3}$$

### 2.2.3 Propriétés de l'espérance

**Proposition 2.6.** *Soient $X$ et $Y$ deux variables aléatoires discrètes réelles ayant un moment d'ordre un fini. L'espérance vérifie les propriétés suivantes :*

- *linéarité : si $\alpha, \beta \in \mathbb{R}$, alors $\alpha X + \beta Y$ admet un moment d'ordre un et*

$$\mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y]. \tag{2.2.4}$$

- *Positivité : si $X$ est positive presque sûrement, c'est-à-dire $\mathbb{P}(X \geqslant 0) = 1$, alors $\mathbb{E}[X] \geqslant 0$.*
- *Croissance : si $X \leqslant Y$ presque sûrement, c'est-à-dire $\mathbb{P}(X \leqslant Y) = 1$, alors $\mathbb{E}[X] \leqslant \mathbb{E}[Y]$.*

**Proposition 2.7** (Inégalités sur l'espérance).

- *Inégalité de Markov : pour tout $a > 0$,*

$$\mathbb{P}(|X| > a) \leqslant \frac{1}{a} \mathbb{E}[|X|]. \tag{2.2.5}$$

- *Inégalité de Tchebychev. Soit $a > 0$. On a*

$$\mathbb{P}(|X| > a) \leqslant \frac{1}{a^2} \mathbb{E}\left[X^2\right]. \tag{2.2.6}$$

- *Inégalité de Jensen : soit $\phi \colon \mathbb{R} \to \mathbb{R}$ une fonction convexe ($\phi(\alpha x + (1-\alpha) y) \leqslant \alpha \phi(x) + (1-\alpha) \phi(y)$ pour tous $x, y \in \mathbb{R}$ et tout $\alpha \in [0,1]$) et $X$ une variable aléatoire telle que $X$ et $\phi(X)$ admettent un moment d'ordre un. Alors*

$$\phi(\mathbb{E}[X]) \leqslant \mathbb{E}[\phi(X)]. \tag{2.2.7}$$

- *Inégalité de Cauchy-Schwarz : soient $X$ et $Y$ deux variables aléatoires réelle admettant un moment d'ordre deux fini. Alors*

$$\mathbb{E}[|XY|] \leqslant \sqrt{\mathbb{E}[X^2]\, \mathbb{E}[Y^2]}. \tag{2.2.8}$$

## 2.3 Variance d'une variable aléatoire

**Définition 2.8** (Variance)**.** *Soit $X$ une variable aléatoire admettant un moment d'ordre deux fini. La variance de $X$, notée* $\operatorname{Var}(X)$*, est définie par*

$$\operatorname{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]. \tag{2.3.1}$$

**Définition 2.9** (Écart-type)**.** *L'écart-type de $X$, noté $\sigma_X$, est défini par*

$$\sigma_X = \sqrt{\operatorname{Var}(X)}. \tag{2.3.2}$$

Remarquons que $\sigma_{cX} = |c|\,\sigma_X$ pour tout $c \in \mathbb{R}$ car $\operatorname{Var}(cX) = c^2 \operatorname{Var}(X)$.

### 2.3.1 Espérance d'une fonction d'une variable aléatoire

Soit $X$ une variable aléatoire discrète réelle, prenant ses valeurs dans un ensemble $E$ au plus dénombrable. Soit $g\colon \mathbb{R} \to \mathbb{R}$ une fonction telle que

$$\sum_{x \in E} |g(x)|\, \mathbb{P}(X = x) < \infty. \tag{2.3.3}$$

La variable aléatoire $g \circ X =: g(X)$, définie par $g \circ X(\omega) = g(X(\omega))$, admet un moment d'ordre 1 et

$$\mathbb{E}[g(X)] = \sum_{x \in E} g(x)\, \mathbb{P}(X = x). \tag{2.3.4}$$

## 2.4 Exemple de lois discrètes

### 2.4.1 Variable constante

**Définition 2.10.** *On dit qu'une variable aléatoire $X$ est constante s'il existe un réel $c$ tel que $\mathbb{P}(X = c) = 1$.*

S'il existe un réel $c$ tel que $\mathbb{P}(X = c) = 1$, alors $\mathbb{E}[X] = c$ et $\operatorname{Var}(X) = 0$.

### 2.4.2 Loi de Bernoulli

**Définition 2.11.** *On dit que $X$ suit une loi de Bernoulli de paramètre $p \in [0, 1]$ si*

$$\mathbb{P}(X = 0) = 1 - p, \quad \mathbb{P}(X = 1) = p. \tag{2.4.1}$$

On a alors

$$\mathbb{E}[X] = p; \quad \operatorname{Var}(X) = p(1 - p). \tag{2.4.2}$$

### 2.4.3 Loi Binomiale

**Définition 2.12.** *On dit que $X$ suit une loi binomiale de paramètres $(n, p)$ avec $n \in \mathbb{N}^*$ et $p \in [0, 1]$ si*

$$\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}, \quad k \in \{0, 1, \dots, n\}. \tag{2.4.3}$$

On a

$$\mathbb{E}[X] = np, \quad \operatorname{Var}(X) = np(1 - p). \tag{2.4.4}$$

### 2.4.4 Loi uniforme discrète

**Définition 2.13.** *On dit que $X$ suit une loi uniform discrète de paramètre $N \in \mathbb{N}^*$ si*

$$\mathbb{P}(X = k) = \frac{1}{N}, \quad k \in \{1, \ldots, N\}. \tag{2.4.5}$$

On a $\mathbb{E}[X] = (N+1)/2$ et $\operatorname{Var}(X) = (N^2 - 1)/12$.

### 2.4.5 Loi Géométrique

**Définition 2.14.** *On dit que $X$ suit une loi géométrique de paramètre $p \in ]0,1[$ si pour chaque $n \in \mathbb{N}^*$,*

$$\mathbb{P}(X = n) = (1-p)^{n-1} p. \tag{2.4.6}$$

On a $\mathrm{E}[X] = \frac{1}{p}, \quad \operatorname{Var}(X) = \frac{1-p}{p^2}$.

### 2.4.6 Loi de Poisson

**Définition 2.15.** *On dit que $X$ suit une loi de Poisson de paramètre $\lambda \in ]0, \infty[$ si pour tout $k \in \mathbb{N}$,*

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}. \tag{2.4.7}$$

On a $\mathbb{E}[X] = \operatorname{Var}(X) = \lambda$.

## 2.5 Fonction génératrice d'une v.a. à valeurs entières positives

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

**Définition 2.16.** *La fonction génératrice d'une variable aléatoire $X\colon \Omega \to \mathbb{N}$ est la fonction $G_X\colon [0,1] \to [0,1]$ définie par*

$$G_X(s) = \sum_{n=0}^{\infty} s^n \mathbb{P}(X = n), 0 < s \leqslant 1, \quad G_X(0) = \mathbb{P}(X = 0). \tag{2.5.1}$$

**Proposition 2.17.** *Soit $X$ une variable aléatoire à valeurs entières : $G_X$ caractérise la loi de $X$ ; plus précisément*

$$\mathbb{P}(X = n) = \frac{G_X^{(n)}(0)}{n!}, \tag{2.5.2}$$

*où $G_X^{(0)} = G_X$ et pour $n \geqslant 1$, $G_X^{(n)} = \left(G_X^{(n-1)}\right)'$.*

$$\mathbb{E}[X] = G'_X(1), \tag{2.5.3}$$

$$\operatorname{Var}(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2. \tag{2.5.4}$$

| | |
|---|---|
| Loi de Bernoulli de paramètre $p$ | $G_X(s) = 1 - p + ps$ |
| Loi binomiale de paramètres $(n, p)$ | $G_X(s) = (1 - p + ps)^n$ |
| Loi géométrique de paramètre $p$ | $G_X(s) = \frac{ps}{1-(1-p)s}$ |
| Loi de Poisson de paramètre $\lambda$ | $G_X(s) = e^{\lambda(s-1)}$ |

## 2.6 Loi conditionnelle et espérance conditionnelle

**Définition 2.18** (Loi conditionnelle)**.** *Soient $X$ et $Y$ deux variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et prenant leurs valeurs respectives dans les ensembles $E$ et $F$. On dit que $X$ et $Y$ sont indépendantes si*

$$\forall x \in E, y \in F, \quad \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y). \tag{2.6.1}$$

**Définition 2.19.** *Soient $X$ et $Y$ deux variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et prenant leurs valeurs respectives dans les ensembles $E$ et $F$. La loi conditionnelle de $Y$ sachant $X$ est la donnée de*

$$\mathbb{P}(Y = y \mid X = x), x \in E, y \in F. \tag{2.6.2}$$

*Remarque* 2.20. Si $X$ est indépendante de $Y$, alors la loi conditionnelle de $Y$ sachant $X$ est simplement la loi de $Y$.

**Définition 2.21** (Indépendance d'une collection de variables aléatoires discrètes)**.** *Soient $X_1, \ldots, X_n$ une suite de variables aléatoires discrètes où $X_i$ prend ses valeurs dans l'ensemble $E_i$. La suite $(X_i, 1 \leqslant i \leqslant n)$ est indépendante si pour tous $x_1 \in E_1, \ldots, x_n \in E_n$,*

$$\mathbb{P}((X_1 = x_1) \cap \cdots \cap (X_n = x_n)) = \mathbb{P}(X_1 = x_1) \ldots \mathbb{P}(X_n = x_n). \tag{2.6.3}$$

**Proposition 2.22.** *Soit $(X_i, 1 \leqslant i \leqslant n)$ une suite indépendante de variables aléatoires de Bernoulli de paramètre $p$. Alors $\sum_{i=1}^n X_i$ suit une loi binomiale de paramètres $(n, p)$.*

### 2.6.1 Loi conditionnelle de $X$ sachant un événement $E$

**Définition 2.23.** *La loi conditionnelle de $X$ sachant un événément $E$ de probabilité non nulle est la donnée des couples $(x_i, \mathbb{P}(X = x_i \mid E))_{i \in I}$, où $(x_i)_{i \in I}$ est l'ensemble des valeurs prises par $X$.*

## 2.7 Couples de variables aléatoires discrètes

### 2.7.1 Loi jointe de $(X, Y)$

**Définition 2.24** (Concept de couples de variables aléatoires)**.** *Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $X : \Omega \to I$, $Y : \Omega \to J$ des variables aléatoires, où $I, J \subset \mathbb{R}$. Le couple de variables aléatoires $(X, Y)$ est la variable aléatoire $(X, Y) : \Omega \to I \times J$ définie par*

$$(X, Y)(\omega) = (X(\omega), Y(\omega)). \tag{2.7.1}$$

**Définition 2.25** (Loi jointe d'un couple de variables aléatoires)**.** *Soit $(X, Y)$ un couple de variables aléatoires discrètes. La loi jointe de $(X, Y)$ est la donnée de*

$$\{(x_i, y_j, p_{i,j})\}_{i \in I, j \in J}, \tag{2.7.2}$$

*où*

- $\{(x_i, y_j)\}_{i \in I, j \in J}$ *sont les valeurs prises par le couple $(X, Y)$,*
- $\{(p_{i,j})\}_{i \in I, j \in J}$ *sont les probabilités correspondantes :*

$$p_{i,j} = \mathbb{P}(X = x_i, Y = y_j), i \in I, j \in J. \tag{2.7.3}$$

*Exemple* 2.26. On peut représenter la loi jointe dans un tableau.

| $y_j \backslash x_i$ | 0 | 2 | 4 |
|---|---|---|---|
| 1 | 1/8 | 0 | 1/8 |
| 2 | 1/16 | 1/16 | 1/16 |
| 8 | 1/16 | 1/4 | 1/4 |

Ici, $I = \{0, 2, 4\}$, $J = \{1, 2, 8\}$,

$$p_{0,1} = \mathbb{P}(X = 0, Y = 1) = 1/8, p_{0,2} = \mathbb{P}(X = 0, Y = 2) = 1/16, p_{0,8} = 1/16$$

$$p_{2,1} = 0, p_{2,2} = 1/16, p_{2,8} = 1/4$$

$$p_{4,1} = 1/8, p_{4,2} = 1/16, p_{4,8} = 1/4$$

*Remarque* 2.27. On doit avoir $\sum_{i \in I} \sum_{j \in J} p_{i,j} = 1$ et $p_{i,j} \geqslant 0$.

### 2.7.2 Lois marginales

**Définition 2.28.** *Soit $(X, Y)$ un couple de variables aléatoires discrètes. Les lois marginales de $X$ et $Y$ sont les lois des variables aléatoires réelles $X$ et $Y$. Si $(X, Y)$ a pour loi jointe $\{(x_i, y_j, p_{i,j})\}_{i \in I, j \in J}$, alors*

- *la loi de $X$ est donnée par $\{(x_i, p_{i,\bullet})\}_{i \in I}$, où*

$$p_{i,\bullet} = \mathbb{P}(X = x_i) = \sum_{j \in J} p_{i,j}. \tag{2.7.4}$$

- *la loi de $Y$ est donnée par $\{(y_j, p_{\bullet,j})\}_{j \in J}$, où*

$$p_{\bullet,j} = \mathbb{P}(Y = y_j) = \sum_{i \in I} p_{i,j}. \tag{2.7.5}$$

*Exemple* 2.29. Revenons à l'exemple précédent.

| $y_j \backslash x_i$ | 0 | 2 | 4 | $\mathbb{P}(Y = y_j)$ |
|---|---|---|---|---|
| 1 | 1/8 | 0 | 1/8 | 1/4 |
| 2 | 1/16 | 1/16 | 1/16 | 3/16 |
| 8 | 1/16 | 1/4 | 1/4 | 9/16 |
| $\mathbb{P}(X = x_i)$ | 1/4 | 5/16 | 7/16 | 1 |

### 2.7.3 Fonction de deux variables

Espérance d'une fonction d'un couple Soit $(X, Y)$ un couple de variables aléatoires discrètes prenant ses valeurs dans $\{(x_i, y_j), i \in I, j \in J\}$, $g\colon \mathbb{R}^2 \to \mathbb{R}$ une fonction, et $Z = g(X, Y)$. Alors $Z$ est une variable aléatoire discrète, prenant ses valeurs dans $\{g(x_i, y_j), i \in I, j \in J\}$. Par conséquent,

$$\mathbb{E}[Z] = \mathbb{E}[g(X, Y)] = \sum_{i \in I} \sum_{j \in J} g(x_i, y_j) \mathbb{P}(X = x_i, Y = y_j). \tag{2.7.6}$$

En particulier, si $X$ est indépendante de $Y$ et $g(x, y) = xy$, on a

$$\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]. \tag{2.7.7}$$

Si $X$ est indépendante de $Y$, en prenant $g(x, y) = s^{x+y}$, $s \in [0, 1]$, on trouve que $G_{X+Y}(s) = G_X(s) G_Y(s)$.

### 2.7.4 Covariance

**Définition 2.30** (Covariance)**.** *Soient $X$ et $Y$ deux variables aléatoire discrètes définies sur un même espace probabilisé $(\Omega \mathcal{F}, \mathbb{P})$. On suppose que $X$ et $Y$ admettent un moment fini d'ordre deux. La covariance de $X$ et $Y$, notée* $\mathrm{Cov}(X, Y)$*, est définie par*

$$\mathrm{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]. \tag{2.7.8}$$

On a également

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X] \mathbb{E}[Y]. \tag{2.7.9}$$

**Proposition 2.31.** *Si $X$ et $Y$ sont deux variables aléatoires indépendantes de carré intégrable, alors* $\mathrm{Cov}(X, Y) = 0$*.*

La réciproque est fausse : soit $X$ prenant les valeurs $-1$, $0$ et $1$ avec probabilité $1/3$ et $Y = X^2$.

**Proposition 2.32.** *Soient $X$, $Y$ et $Z$ trois variables aléatoires discrètes de carré intégrable et $a, b \in \mathbb{R}$. On a :*

- $\operatorname{Cov}(X, X) = \operatorname{Var}(X)$ ;
- $\operatorname{Cov}(X, Y) = \operatorname{Cov}(Y, X)$ ;
- $\operatorname{Cov}(aX + bY, Z) = a \operatorname{Cov}(X, Z) + b \operatorname{Cov}(Y, Z)$ ;
- $\operatorname{Cov}(X, a) = 0$.

**Proposition 2.33** (Covariance et variance)**.** *Si $X$ et $Y$ sont deux variables aléatoires de carré intégrable, alors*

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2 \operatorname{Cov}(X, Y). \tag{2.7.10}$$

*Si de plus $X$ et $Y$ sont indépendantes, alors*

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y). \tag{2.7.11}$$

### 2.7.5 Coefficient de corrélation linéaire

**Définition 2.34** (Coefficient de corrélation linéaire)**.** *Soient $X$ et $Y$ deux variables aléatoires de carré intégrable. Le coefficient de corrélation linéaire, noté $\rho(X, Y)$, est défini par*

$$\rho(X, Y) = \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X) \operatorname{Var}(Y)}}. \tag{2.7.12}$$

**Proposition 2.35.** *Si $X$ et $Y$ sont indépendantes, alors $\rho(X, Y) = 0$.*

**Proposition 2.36.** *Pour toutes variables aléatoires $X$ et $Y$ non constantes,*

$$-1 \leqslant \rho(X, Y) \leqslant 1. \tag{2.7.13}$$

L'inégalité (2.7.13) est une conséquence de l'inégalité de Cauchy-Schwarz (2.2.8).

**Proposition 2.37.** *Soient $X$ et $Y$ deux variables aléatoires de carré intégrable. On a*

1. *$\rho(X, Y) = 1$ si et seulement si il existe $a > 0$ et $b \in \mathbb{R}$ tels que $X = aY + b$.*
2. *$\rho(X, Y) = -1$ si et seulement si il existe $a < 0$ et $b \in \mathbb{R}$ tels que $X = aY + b$.*

# 3 Convergence et théorèmes en probabilité

## 3.1 Convergence en probabilité

**Définition 3.1.** *On dit qu'une suite de variables aléatoires $(X_n)_{n \geqslant 1}$ converge en probabilité vers une variable aléatoire $X$ si pour tout $\varepsilon > 0$,*

$$\lim_{n \to \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0. \tag{3.1.1}$$

## 3.2 Convergence en moyenne d'ordre $p$

**Définition 3.2.** *On dit qu'une suite de variables aléatoires $(X_n)_{n\geqslant 1}$ converge en moyenne d'ordre $p \in \mathbb{N}^*$ vers une variable aléatoire $X$ si*

$$\lim_{n\to\infty} \mathbb{E}\left[|X_n - X|^p\right] = 0. \tag{3.2.1}$$

**Proposition 3.3.** *Si $(X_n)_{n\geqslant 1}$ converge en moyenne d'ordre $p$ pour un certain $p \in \mathbb{N}^*$ vers une variable aléatoire $X$, alors $(X_n)_{n\geqslant 1}$ converge en probabilité vers $X$.*

## 3.3 Convergence en loi pour des variables aléatoires entières

**Définition 3.4.** *Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$. On dit que $(X_n)_{n\geqslant 1}$ converge en loi vers $X$ si pour tout $k \in \mathbb{N}$,*

$$\lim_{n\to\infty} \mathbb{P}\left(X_n = k\right) = \mathbb{P}\left(X = k\right). \tag{3.3.1}$$

**Proposition 3.5.** *Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$ qui converge en probabilité vers une variable aléatoire $X$. Alors $(X_n)_{n\geqslant 1}$ converge en loi vers $X$.*

## 3.4 Loi faible des grands nombres

**Théorème 3.6** (Loi faible des grands nombres)**.** *Soit $(X_n)_{n\geqslant 1}$ une suite indépendante de variables aléatoires de même loi. Soit $\overline{X_n} := \frac{\sum_{i=1}^n X_i}{n}$ la moyenne empirique. On suppose que $X_1$ admet un moment fini d'ordre deux. Alors $\left(\overline{X_n} - \mathbb{E}\left[X_1\right]\right)_{n\geqslant 1}$ converge en moyenne d'ordre deux vers $0$. En particulier,*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\overline{X_n} - \mathbb{E}\left[X_1\right]\right| > \varepsilon\right) = 0. \tag{3.4.1}$$

## 3.5 Théorème limite central

**Théorème 3.7.** *(Théorème limite central (TLC)] Soit $(X_n)_{n\geqslant 1}$ une suite indépendante de variables aléatoires de même loi. Soit $\overline{X_n} := \frac{\sum_{i=1}^n X_i}{n}$ la moyenne empirique et $\sigma^2 = \operatorname{Var}(X_1) > 0$. Alors pour tout réel $t$,*

$$\mathbb{P}\left(\sqrt{n}\left(\overline{X_n} - \mathbb{E}\left[X_1\right]\right) \leqslant t\sigma\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{s^2}{2}\right) \mathrm{d}s. \tag{3.5.1}$$

Soit l'intervalle aléatoire

$$I_{n,a} := \left[\overline{X_n} - \frac{a\sigma}{\sqrt{n}}, \overline{X_n} + \frac{a\sigma}{\sqrt{n}}\right]. \tag{3.5.2}$$

Alors

$$\lim_{n\to\infty} \mathbb{P}\left(\mathbb{E}\left[X_1\right] \in I_{n,a}\right) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \exp\left(-\frac{s^2}{2}\right) \mathrm{d}s. \tag{3.5.3}$$

**Définition 3.8.** *Soit $X$ une variable aléatoire. La fonction de répartition de $X$, notée $F$, est définie par*

$$F\left(t\right) := \mathbb{P}\left(X \leqslant t\right) = \mathbb{P}\left\{\omega \in \Omega : X\left(\omega\right) \leqslant t\right\}. \tag{3.5.4}$$