

Chapitre 05 – Probabilités (1)

1 Probabilités sur un univers discret

1.1 Événements et variables aléatoires

Définition 1.

- On appelle tribu (ou sigma-algèbre) sur un **univers** Ω une partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ telle que :
 - $\Omega \in \mathcal{A}$
 - $\forall A \in \mathcal{A}, \bar{A} = \Omega \setminus A \in \mathcal{A}$
 - pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{A}$, $\bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$
- Les éléments de $\mathcal{A}$ sont appelés événements (ou parties mesurables).
- Deux événements A et B sont dits incompatibles lorsque $A \cap B = \emptyset$.
- Un système complet d'événements est une suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles dont la réunion est égale à Ω .

Proposition 2. Si $\mathcal{A}$ est une tribu sur Ω , $\emptyset \in \mathcal{A}$ et $\bigcap_{n=0}^{+\infty} A_n \in \mathcal{A}$ pour toute suite $(A_n)_{n \in \mathbb{N}}$ de $\mathcal{A}^{\mathbb{N}}$.

Définition 3.

- Si $\mathcal{A}$ est une tribu sur un ensemble Ω , une variable aléatoire discrète sur $(\Omega, \mathcal{A})$ est une application X sur Ω , d'image finie ou dénombrable, telle que : $\forall x \in X(\Omega), X^{-1}(\{x\}) \in \mathcal{A}$.
- On définit, si $B \subset X(\Omega)$, l'événement $\{X \in B\} = \{\omega \in \Omega, X(\omega) \in B\}$.
- Si $X(\Omega) \subset \mathbb{R}$, X est dite réelle. On définit alors les événements $X < a$, $X \leq a$, etc.
- Si A est un événement, la **fonction indicatrice** de A , notée $\mathbb{1}_A$, est la variable aléatoire définie par $\mathbb{1}_A(\omega) = 1$ si $\omega \in A$ et 0 sinon.

Proposition 4 (et définition).

- Si X est une variable aléatoire discrète à valeurs dans E , ensemble dénombrable, si f est une fonction de E dans F , alors on note $f(X)$ la variable aléatoire définie par : $\forall \omega \in \Omega, f(X)(\omega) = f(X(\omega))$. On définit ainsi une variable aléatoire discrète.
- Si X et Y sont deux variables aléatoires discrètes, à valeurs respectives dans E_1 et E_2 , alors pour toute $g : E_1 \times E_2 \rightarrow G$, $g(X, Y)$ est une variable aléatoire discrète.
- En particulier, (X, Y) est une variable aléatoire discrète.
- On généralise au cas de n variables aléatoires.

1.2 Probabilité

Définition 5.

- Si $\mathcal{A}$ est une tribu sur Ω , une probabilité sur $(\Omega, \mathcal{A})$ est une application $\mathbb{P}$ de $\mathcal{A}$ dans $[0, 1]$ vérifiant : $\mathbb{P}(\Omega) = 1$ et pour toute suite $(A_n)_{n \geq 0}$ d'événements deux à deux incompatibles, $\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ (σ -additivité).
- On dit alors que $(\Omega, \mathcal{A}, \mathbb{P})$ est un **espace probabilisé**.
- Si $\mathbb{P}(A) = 1$ (resp 0), A est dit presque sûr (resp. négligeable).
- On dit qu'une suite d'événements $(A_n)_{n \in \mathbb{N}}$ est un système quasi-complet lorsque ces événements sont deux à deux incompatibles et que la somme de leurs probabilités vaut 1.

Proposition 6. Soient A, B , deux événements, $(A_n)_{n \in \mathbb{N}}$ une suite d'événements.

1. $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$
2. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
3. Une probabilité est croissante pour l'inclusion.
4. (Continuité croissante) Si pour tout n , $A_n \subset A_{n+1}$, alors $\mathbb{P}(A_n) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(\bigcup_{k=0}^{+\infty} A_k)$.
5. De manière générale, $\mathbb{P}(\bigcup_{k=0}^n A_k) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(\bigcup_{k=0}^{+\infty} A_k)$.
6. (Continuité décroissante) Si pour tout n , $A_{n+1} \subset A_n$, alors $\mathbb{P}(A_n) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(\bigcap_{k=0}^{+\infty} A_k)$.
7. De manière générale, $\mathbb{P}(\bigcap_{k=0}^n A_k) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(\bigcap_{k=0}^{+\infty} A_k)$.
8. (Sous-additivité) $\mathbb{P}(\bigcup_{n=0}^{+\infty} A_n) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$.
9. Une réunion dénombrable d'événements négligeables est négligeable.
10. Une intersection finie ou dénombrable d'événements presque sûrs est presque sûre.

Remarque 7. Si Ω est discret, une distribution de probabilité sur Ω est une famille $(a_\omega)_{\omega \in \Omega}$ de réels de $[0, 1]$ telle que $\sum_{\omega \in \Omega} a_\omega = 1$. On définit alors naturellement une probabilité sur $\mathcal{P}(\Omega)$.

Proposition 8 (Probabilités totales). Si $(A_n)_{n \in \mathbb{N}}$ est un sqce alors : $\forall B, \mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B \cap A_n)$.

Définition 9. Si X est une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, on appelle loi de X l'application $\mathbb{P}_X : \begin{cases} \mathcal{P}(X(\Omega)) \rightarrow [0, 1] \\ U \mapsto \mathbb{P}(X \in U) = \mathbb{P}(X^{-1}(U)) \end{cases}$.

Remarque 10.

1. $\mathbb{P}_X$ est une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$, entièrement déterminée par $\{P(X = x)\}_{x \in X(\Omega)}$.
2. En pratique, on n'utilise **JAMAIS** la notation $\mathbb{P}_X$.
3. Déterminer la **loi** d'une variable aléatoire X , c'est déterminer un ensemble fini ou dénombrable E tel que $X(\Omega) \subset E$ et, pour tout x dans E , $\mathbb{P}(X = x)$.
4. Si deux variables X et Y suivent la même loi, on note $X \sim Y$.
5. Si $X \sim Y$ et que $f : X(\Omega) \rightarrow F$, alors $f(X) \sim f(Y)$.

Définition 11 (Lois usuelles). Soit $p \in]0, 1[, \lambda > 0$.

Loi	Paramètre	Notation	$X(\Omega)$	$\mathbb{P}(X = k)$	Espérance	Variance
Uniforme		$X \sim \mathcal{U}(E)$	E	$1/\text{Card}(E)$		
Bernoulli	p	$X \sim \mathcal{B}(p)$	$\{0, 1\}$	$1 - p \ (k = 0); p \ (k = 1)$	p	$p(1 - p)$
Binomiale	(n, p)	$X \sim \mathcal{B}(n, p)$	$\llbracket 0, n \rrbracket$	$\binom{n}{k} p^k (1 - p)^{n-k}$	np	$np(1 - p)$
Géométrique	p	$X \sim \mathcal{G}(p)$	$\mathbb{N}^*$	$p(1 - p)^{k-1}$	$1/p$	$(1 - p)/p^2$
Poisson	λ	$X \sim \mathcal{P}(\lambda)$	$\mathbb{N}$	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ

Remarque 12. Si $X \sim \mathcal{G}(p)$, alors $\mathbb{P}(X > k) = (1 - p)^k$.

Définition 13.

1. La loi du couple (X, Y) est appelée loi conjointe.
Elle est déterminée par $\{P((X, Y) = (x, y)) = P((X = x) \cap (Y = y)), (x, y) \in X(\Omega) \times Y(\Omega)\}$.
2. Les lois de X et de Y sont appelées lois marginales.

Remarque 14. Il faut pouvoir retrouver les lois marginales à l'aide des probabilités totales :

$$\mathbb{P}(X = x_0) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x_0, Y = y) \text{ et } \mathbb{P}(Y = y_0) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y_0).$$

2 Conditionnement et indépendance

2.1 Conditionnement

Définition 15. Si A et B sont des événements et si $P(B) \neq 0$, la probabilité conditionnelle de A sachant B est définie par : $P_B(A) = P(A | B) = \frac{P(A \cap B)}{P(B)}$. On définit bien une probabilité.

Proposition 16 (Le trio des formules). On suppose que toutes les probabilités conditionnelles suivantes sont définissables.

1. (Probabilités composées) $\mathbb{P}(A_1 \cap \dots \cap A_n) = P(A_1) P_{A_1}(A_2) P_{A_1 \cap A_2}(A_3) \dots P_{A_1 \cap \dots \cap A_{n-1}}(A_n)$.
2. (Probabilités totales) Si $(A_n)_{n \in \mathbb{N}}$ est un s(q)ce, $\forall B \in \mathcal{A}$, $P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n)$.
3. (Formule de Bayes) $\mathbb{P}_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{P_B(A)P(B)}{P(A)}$.

Définition 17. Soient X une variable aléatoire discrète, A un événement. La loi de X conditionnée à l'événement A est la donnée, pour tout x de E , de $\mathbb{P}_A(X = x)$.

2.2 Indépendance

Définition 18. Soient $A, B, A_1, \dots, A_n$ des événements, $X, Y, (X_k)_{k \in \mathbb{N}}$ des variables aléatoires discrètes.

1. A et B sont dits indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ ou (quans $\mathbb{P}(B) > 0$ à $\mathbb{P}_B(A) = \mathbb{P}(A)$)
2. $A_1, A_2, \dots, A_n$ sont dits mutuellement indépendants si : $\forall I \subset \llbracket 1, n \rrbracket$, $\mathbb{P}(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$.
3. X et Y sont indépendantes lorsque : $\forall A \subset X(\Omega), \forall B \subset Y(\Omega), \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$.
On note $X \perp\!\!\!\perp Y$.
4. $(X_1, \dots, X_n)$ sont mutuellement indépendantes si pour tous $A_1 \subset X_1(\Omega), \dots, A_n \subset X_n(\Omega)$,
$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i).$$
5. Les variables aléatoires $(X_k)_{k \in \mathbb{N}}$ sont dites mutuellement indépendantes si pour tout n dans $\mathbb{N}$, $(X_0, \dots, X_n)$ sont mutuellement indépendantes.
6. Des variables aléatoires sont dites i.i.d. (indépendantes et identiquement distribuées) si elles sont indépendantes et suivent la même loi.

Proposition 19.

1. $X \perp\!\!\!\perp Y \iff \forall (x, y) \in X(\Omega) \times Y(\Omega), P(X = x, Y = y) = P(X = x)P(Y = y)$.
2. L'indépendance mutuelle implique l'indépendance deux à deux ; pas l'inverse.
3. Si $A_1, A_2, \dots, A_n$ sont mutuellement indépendants alors $A_1, A_2, \dots, \overline{A_n}$ le sont aussi.
4. (Lemme des coalitions) Si $X_1, \dots, X_n$ sont des v.a.d. indépendantes alors $f(X_1, \dots, X_k)$ et $g(X_{k+1}, \dots, X_n)$ sont deux v.a.d. indépendantes.

Proposition 20. Soient $X_1, \dots, X_n$ des vaia telles que $X_1 \sim \mathcal{B}(p)$. Alors $X_1 + \dots + X_n \sim \mathcal{B}(n, p)$.

3 Moments d'une variable aléatoire

3.1 Espérance d'une v.a.d. réelle ou complexe

Définition 21.

1. L'espérance d'une v.a.d. X à valeurs dans $[0, +\infty]$ est définie par : $\mathbb{E}(X) = \sum_{x \in X(\Omega)} xP(X = x)$. (avec la convention $xP(X = x) = 0$ si $x = +\infty$ et $\mathbb{P}(X = +\infty) = 0$)
2. On dit qu'une v.a.d. X à valeurs dans $\mathbb{C}$ est d'espérance finie, et on note $X \in L^1$, lorsque la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable. $\mathbb{E}(X)$ est alors la somme de cette famille.
3. On dit que X est centrée lorsque $\mathbb{E}(X) = 0$.

Proposition 22. Si $X \in L^1$ et X est réelle, presque sûrement positive, alors $\mathbb{E}(X) \geq 0$, avec égalité ssi X est presque sûrement nulle.

Proposition 23. Si X est à valeurs dans $\mathbb{N} \cup \{+\infty\}$, alors $\mathbb{E}(X) = \sum_{n \geq 1} \mathbb{P}(X \geq n)$.

Proposition 24 (Espérances usuelles). Cf. Définition 11.

Remarque 25. Il est important de remarquer que $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

Proposition 26 (Formule de transfert). $f(X) \in L^1$ si et seulement si la famille $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas, $\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X = x)$.

Proposition 27.

1. Linéarité de l'espérance : si X et Y sont deux v.a.d. d'espérance finie, si $(\lambda, \mu) \in \mathbb{C}^2$, alors $\lambda X + \mu Y$ admet une espérance et $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.
2. Si $X \leq Y$ presque sûrement, $\mathbb{E}(X) \leq \mathbb{E}(Y)$.
3. Si $|X| \leq Y$ et $\mathbb{E}(Y) < +\infty$ alors X est d'espérance finie et $|\mathbb{E}(X)| \leq \mathbb{E}(Y)$.

Remarque 28. On retrouve l'espérance de la loi binomiale.

Proposition 29. Si X et Y sont deux v.a.r. indépendantes d'espérances finies, alors XY aussi et $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. On adapte au cas de n v.a.r. mutuellement indépendantes.

Proposition 30 (Inégalité de Markov). Si X est une v.a.d. réelle **positive** (presque sûrement) d'espérance finie, alors pour tout $a > 0$, $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$.

3.2 Moments d'ordre 2 – Variance d'une v.a.d. réelle

Proposition 31 (Inégalité de Cauchy-Schwarz). Soient X et Y deux v.a.d. réelles. Si X^2 et Y^2 sont dans L^1 , alors XY aussi et $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)\mathbb{E}(Y^2)}$, avec égalité ssi X et Y sont presque sûrement colinéaires. Ainsi, $X \in L^2 \Rightarrow X \in L^1$.

Définition 32. Si X et Y sont deux v.a.d. réelles avec X^2 et Y^2 dans L^1 , on définit :

1. la variance de X , $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
2. la covariance de X et Y , $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
3. l'écart type de X , $\sigma(X) = \sqrt{\mathbb{V}(X)}$ (et l'on dit que X est réduite lorsque $\sigma(X) = 1$)

Proposition 33 (Variances usuelles). Cf. Définition 11.

Proposition 34.

1. La covariance est une forme bilinéaire symétrique positive sur l'espace vectoriel L^2 .
2. $\mathbb{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$.
3. Si X et Y sont indépendantes ou juste décorrélées, alors $\text{Cov}(X, Y) = 0$ et donc $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$. On généralise à n variables.
4. $\mathbb{V}(X) = 0$ si et seulement si X est presque sûrement constante.
5. $\forall (a, b) \in \mathbb{R}^2$, $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.
6. si $\sigma(X) > 0$ alors $\frac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée réduite.

Proposition 35.

1. (inégalité de Bienaymé-Tchebycheff) Soit X une v.a.d. réelle d'espérance et de variance finies. Alors pour tout $a > 0$, $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}$.
2. (loi faible des grands nombres) Soit $(X_k)_{k \in \mathbb{N}^*}$ une suite de v.a.d. réelles i.i.d., d'espérance finie μ et de variance finie σ^2 . On note, pour n dans $\mathbb{N}^*$, $S_n = \sum_{k=1}^n X_k$. Alors pour tout $\varepsilon > 0$, $\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0$.
On dit que la suite $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge en probabilité vers μ .