

Résumé 10 – Probabilités discrètes

Dénombrement

Soient Ω un ensemble à n éléments et $p \in \llbracket 0, n \rrbracket$.

Définition

- Un p -uplet ou une p -liste de Ω est une famille de p éléments de Ω .
- Un arrangement de p éléments de Ω est un p -uplet constitué d'éléments de Ω distincts.
- Une permutation de Ω est un arrangement de Ω à n éléments.
- Une combinaison de p éléments de Ω est un sous-ensemble de Ω contenant p éléments.

On modélise les tirages successifs avec remise à l'aide de listes, les tirages successifs sans remise avec des arrangements et les tirages simultanés avec des combinaisons.

Théorème

- Il y a n^p p -listes de Ω .
- Il y a $\frac{n!}{(n-p)!}$ arrangements de p éléments de Ω .
- Il y a $n!$ permutations de Ω .
- Il y a $\binom{n}{p}$ combinaisons de p éléments de Ω .

Soient $n, p, m \in \mathbb{N}$.

$$\begin{aligned} & \binom{n}{0} = 1, \binom{n}{1} = n & \binom{n}{p} &= \binom{n}{n-p} \\ & p \binom{n}{p} = n \binom{n-1}{p-1} & \binom{n-1}{p-1} + \binom{n-1}{p} &= \binom{n}{p} \\ & \sum_{k=0}^n \binom{n}{k} = 2^n & \sum_{k=0}^p \binom{n}{k} \binom{m}{p-k} &= \binom{n+m}{p} \end{aligned}$$

Probabilités discrètes

→ Tribus et probabilités

Définition : Tribu

Une tribu sur Ω est une partie \mathcal{A} de $\mathcal{P}(\Omega)$ qui vérifie :

- $\Omega \in \mathcal{A}$;
- Si $A \in \mathcal{A}$ alors $\bar{A} \in \mathcal{A}$
- Si $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$, alors $\bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$

La donnée d'un univers Ω (au plus dénombrable ou non) et d'une tribu \mathcal{A} définit un espace probabilisable (Ω, \mathcal{A}) ; tout élément de \mathcal{A} est appelé événement de Ω .

Soit (Ω, \mathcal{A}) un espace probabilisable.

Définition : Système complet d'événements

On appelle système complet d'événements toute famille finie ou dénombrable $(A_i)_{i \in I}$ d'événements telle que :

- Pour tous i et j distincts, $A_i \cap A_j = \emptyset$;
- $\bigcup_{i \in I} A_i = \Omega$.

Définition : Probabilité

On appelle probabilité sur (Ω, \mathcal{A}) toute application $\mathbf{P} : \mathcal{A} \rightarrow [0, 1]$ vérifiant :

- $\mathbf{P}(\Omega) = 1$
- Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles,

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n) \quad (\sigma\text{-additivité})$$

Le triplet $(\Omega, \mathcal{A}, \mathbf{P})$ est appelé espace probabilisé.

Une probabilité est une application qui opère sur les événements. La σ -additivité assure la convergence des séries manipulées.

Proposition

Soient $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé et $A, B \in \mathcal{A}$.

- $\mathbf{P}(\emptyset) = 0$ et $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$.
- Si $A \subset B$ alors $\mathbf{P}(A) \leq \mathbf{P}(B)$.
- $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

Définition

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

- Si $\mathbf{P}(A) = 0$, l'événement A est dit négligeable ou quasi-impossible.
- Si $\mathbf{P}(A) = 1$, l'événement A est dit presque sûr ou quasi-certain.

→ Distribution de probabilités

On appelle distribution de probabilités discrète sur Ω toute famille de réels positifs indexée par Ω et de somme égale à 1.

Théorème

Soit Ω un ensemble au plus dénombrable.

- Soit \mathbf{P} une probabilité définie sur $(\Omega, \mathcal{P}(\Omega))$. On pose, pour tout $\omega \in \Omega$, $p_\omega = \mathbf{P}(\{\omega\})$. Alors, $(p_\omega)_{\omega \in \Omega}$ est une distribution de probabilités.
- Si $(p_\omega)_{\omega \in \Omega}$ est une distribution de probabilités, il existe une unique probabilité \mathbf{P} définie sur $(\Omega, \mathcal{P}(\Omega))$ telle que pour tout $\omega \in \Omega$, $p_\omega = \mathbf{P}(\{\omega\})$.

Dans le cas fini, on appelle probabilité uniforme sur Ω l'unique probabilité qui prend la même valeur pour chaque événement élémentaire.

→ Propriétés

Proposition : Continuité croissante

Si $(A_n)_{n \in \mathbb{N}}$ est une suite croissante d'événements (au sens de l'inclusion), alors :

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}(A_n)$$

De même, si $(A_n)_{n \in \mathbb{N}}$ est une suite décroissante d'événements (au sens de l'inclusion), $\mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbf{P}(A_n)$.

Ainsi, pour toute suite d'événements $(A_n)_{n \in \mathbb{N}}$,

$$\lim_{p \rightarrow +\infty} \mathbf{P}\left(\bigcup_{n=0}^p A_n\right) = \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right)$$

$$\lim_{p \rightarrow +\infty} \mathbf{P}\left(\bigcap_{n=0}^p A_n\right) = \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

Proposition : Sous-additivité

• Si (A_1, \dots, A_n) est une famille d'événements, alors :

$$\mathbf{P}\left(\bigcup_{k=0}^n A_k\right) \leq \sum_{k=0}^n \mathbf{P}(A_k)$$

• Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements et si la série $\sum \mathbf{P}(A_n)$ converge, alors :

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

→ Conditionnement et indépendance

Théorème / Définition : Probabilité conditionnelle

Soit A un événement tel que $\mathbf{P}(A) \neq 0$. L'application

$$\mathbf{P}_A : \begin{cases} \mathcal{A} \longrightarrow \mathbb{R} \\ B \longmapsto \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} \end{cases}$$

est une probabilité sur Ω . On l'appelle probabilité conditionnelle relative à A (ou sachant A).

En tant que probabilité, \mathbf{P}_A vérifie toutes les propriétés énoncées précédemment.

Théorème : Formule des probabilités composées

Soient $n \geq 2$ et (A_1, A_2, \dots, A_n) une famille d'événements telle que $\mathbf{P}(A_1 \cap \dots \cap A_{n-1}) \neq 0$. Alors,

$$\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1) \mathbf{P}_{A_1}(A_2) \times \dots \times \mathbf{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n)$$

Théorème : Formule des probabilités totales

Soit $(A_n)_{n \in \mathbb{N}}$ un système complet d'événements. Pour tout événement B , la série de terme général $\mathbf{P}(B \cap A_n)$ est convergente et :

$$\mathbf{P}(B) = \sum_{n=0}^{+\infty} \mathbf{P}(B \cap A_n) = \sum_{n=0}^{+\infty} \mathbf{P}(B|A_n) \mathbf{P}(A_n)$$

Théorème : Formule de Bayes

Soient A et B deux événements,

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)}{\mathbf{P}(B)} \times \mathbf{P}(B|A)$$

Définition : Indépendance

- Deux événements A et B sont dits indépendants si $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$.
- Soit $(A_i)_{i \in I}$ une famille d'événements. Ces événements sont dits (mutuellement) indépendants si pour toute partie finie $J \subset I$,

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

L'indépendance mutuelle d'une famille d'événements implique qu'ils sont deux à deux indépendants mais la réciproque est fautive.

Variables aléatoires discrètes

Soient $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé et E un ensemble quelconque (souvent \mathbb{R} ou \mathbb{C}).

Définition : Variable aléatoire discrète

On appelle variable aléatoire réelle discrète toute application $X : \Omega \rightarrow E$ telle que :

- $X(\Omega)$ est un ensemble fini ou dénombrable ;
- Pour tout $x \in X(\Omega)$, $(X = x) = X^{-1}(\{x\}) \in \mathcal{A}$.

X désigne désormais une variable aléatoire discrète.

$$X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$$

Pour tout $A \subset X(\Omega)$, $(X \in A) \in \mathcal{A}$; la famille $((X = x_n))_{n \in \mathbb{N}}$ est un système complet d'événements.

→ Loi d'une variable aléatoire

Définition : Loi d'une variable aléatoire

On appelle loi de probabilité de X l'application :

$$\mathbf{P}_X : \begin{cases} \mathcal{P}(X(\Omega)) \longrightarrow \mathbb{R} \\ A \longmapsto \mathbf{P}(X \in A) \end{cases}$$

\mathbf{P}_X est une probabilité sur $X(\Omega)$.

La loi de X est entièrement déterminée par la distribution de probabilités $(\mathbf{P}(X = x))_{x \in X(\Omega)}$.

Notations usuelles :

- Si X suit la loi de probabilité $\mathcal{L} : X \sim \mathcal{L}$;
- Si X et Y suivent la même loi : $X \sim Y$.

Si X est à valeurs dans E et $f : E \rightarrow F$, alors $f(X)$ est une variable aléatoire, de loi donnée par :

$$\forall A \subset f(X(\Omega)), \mathbf{P}_{f(X)}(A) = \mathbf{P}(f(X) \in A) = \sum_{\substack{x \in X(\Omega) \\ f(x) \in A}} \mathbf{P}(X = x)$$

On généralise aux fonctions de plusieurs variables.

En particulier,

$$\mathbf{P}(X_1 + \dots + X_n = x) = \sum_{\substack{(x_1, \dots, x_n) \in \mathbb{R}^n \\ x_1 + \dots + x_n = x}} \mathbf{P}(X_1 = x_1, \dots, X_n = x_n)$$

→ Vecteurs aléatoires discrets

(X, Y) désigne un couple de variables aléatoires discrètes.

Définition : Lois conjointe et marginales

- La loi conjointe de X et de Y est la loi de (X, Y) . Elle est donnée par la distribution de probabilités :

$$\left(\mathbf{P}(X = x, Y = y)\right)_{(x,y) \in X(\Omega) \times Y(\Omega)}$$

- Les lois marginales de (X, Y) sont celles de X et Y .

La formule des probabilités totales permet de trouver les lois marginales à partir de la loi conjointe.

Définition : Lois conditionnelles

- On appelle loi conditionnelle de X sachant $(Y = y)$ la loi définie par $(\mathbf{P}(X = x | Y = y))_{x \in X(\Omega)}$;
- On appelle loi conditionnelle de Y sachant $(X = x)$ la loi définie par $(\mathbf{P}(Y = y | X = x))_{y \in Y(\Omega)}$.

On étend les définitions suivantes aux n -uplets de variables aléatoires (X_1, \dots, X_n) .

Définition : (Mutuelle) indépendance

Les variables X_1, \dots, X_n sont dites indépendantes si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,

$$\mathbf{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n \mathbf{P}(X_i = x_i)$$

L'indépendance se traduit de manière équivalente par : pour tout $(A_1, \dots, A_n) \subset X_1(\Omega) \times \dots \times X_n(\Omega)$, les événements $(X_1 \in A_1), \dots, (X_n \in A_n)$ sont mutuellement indépendants.

L'indépendance deux à deux ne garantit pas l'indépendance mutuelle.

Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes. Plus généralement (lemme des coalitions), si les variables X_1, \dots, X_n sont mutuellement indépendantes, $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.

→ Famille infinie de variables aléatoires

On considère une famille infinie $(X_i)_{i \in I}$ de variables aléatoires sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. Alors,

- les variables X_i sont dites indépendantes si pour toute partie finie J de I , la famille $(X_i)_{i \in J}$ est indépendante.
- si les variables X_i suivent de plus toutes la même loi, on dira que $(X_i)_{i \in I}$ est une famille de variables indépendantes et identiquement distribuées (i.i.d.).

Théorème : Théorème de Kolmogorov

Soit $(\mathcal{L}_n)_{n \geq 1}$ une suite de lois discrètes sur des ensembles E_n . Il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ et une suite $(X_n)_{n \geq 1}$ de variables aléatoires indépendantes telles que pour tout $n \geq 1$, $X_n \sim \mathcal{L}_n$.

→ Moments d'une variable aléatoire

Soit X une v.a.d. sur $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans \mathbb{C} .

Définition : Espérance

La variable X est dite d'espérance finie si la famille $(x\mathbf{P}(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas, on appelle espérance de X le nombre complexe :

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

En pratique, pour $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$, on justifiera la convergence absolue de $\sum x_n \mathbf{P}(X = x_n)$.

Si X^r admet une espérance, on appelle moment d'ordre $r \in \mathbb{N}$ le nombre complexe $\mathbf{E}(X^r)$. On note L^r l'ensemble des variables aléatoires admettant un moment d'ordre r .

Théorème : Théorème de transfert

Soit $f : X(\Omega) \rightarrow \mathbb{R}$. $f(X)$ est d'espérance finie ssi la famille $(f(x)\mathbf{P}(X = x))_{x \in X(\Omega)}$ est sommable. Et alors,

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

L'espérance est linéaire, positive, croissante et vérifie l'inégalité triangulaire. Si $|X| \leq Y$ et $Y \in L^1$, alors $X \in L^1$.

Si X est à valeurs dans \mathbb{N} et d'espérance finie,

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

Théorème : Espérance et indépendance

Soient $X, Y \in L^1$ indépendantes.

Alors, $XY \in L^1$ et $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

La réciproque est fautive.

On ne considère désormais que des variables réelles.

Définition : Variance

Si $X \in L^2$, $(X - \mathbf{E}(X))^2$ est d'espérance finie. On appelle variance de X et on note $\mathbf{V}(X)$ le réel positif :

$$\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2) = \sum_{x \in X(\Omega)} (x - \mathbf{E}(X))^2 \mathbf{P}(X = x)$$

On appelle écart type de X le réel $\sigma(X) = \sqrt{\mathbf{V}(X)}$.

Proposition : Formule de Koenig-Huygens

Si $X \in L^2$, $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

Une inégalité importante : $|XY| \leq \frac{X^2 + Y^2}{2}$.

Définition

Si $X, Y \in L^2$, alors $(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))$ est d'espérance finie. On appelle covariance de X et Y et on note $\text{cov}(X, Y)$ le réel :

$$\text{cov}(X, Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$$

Si $\text{cov}(X, Y) = 0$, les variables sont dites *décorrélées*.

On suppose par la suite que $X, Y \in L^2$.

Théorème : Formule de Kœnig-Huygens

$$\text{cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X) \cdot \mathbf{E}(Y).$$

Proposition

Pour tous $a, b \in \mathbb{R}$,

$$\mathbf{V}(aX + bY) = a^2\mathbf{V}(X) + b^2\mathbf{V}(Y) + 2ab\text{cov}(X, Y)$$

Plus généralement,

$$\mathbf{V}(X_1 + \dots + X_n) = \sum_{i=1}^n \mathbf{V}(X_i) + 2 \times \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j)$$

En cas de décorrélation, $\mathbf{V}(X_1 + \dots + X_n) = \sum_{i=1}^n \mathbf{V}(X_i)$.

Cas particulier : $\mathbf{V}(aX + b) = a^2\mathbf{V}(X)$.

Inégalité de Cauchy-Schwarz : $\mathbf{E}(XY)^2 \leq \mathbf{E}(X) \cdot \mathbf{E}(Y)$.

En particulier, $|\text{cov}(X, Y)| \leq \sigma(X) \cdot \sigma(Y)$.

→ Fonctions génératrices

Définition : Fonction génératrice

Si X est à valeurs dans \mathbb{N} , la fonction génératrice de la variable X est définie par :

$$G_X : t \mapsto \mathbf{E}(t^X) = \sum_{n=0}^{+\infty} \mathbf{P}(X = n)t^n$$

La série entière $\sum_{n \in \mathbb{N}} \mathbf{P}(X = n)t^n$ a un rayon de convergence $R \geq 1$ et converge normalement sur $D_f(0, 1)$.

Théorème : Fonction génératrice et moments

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

(i) La variable aléatoire X admet une espérance $\mathbf{E}(X)$ si et seulement si G_X est dérivable en 1. Si tel est le cas, $\mathbf{E}(X) = G'_X(1)$.

(ii) La variable aléatoire X admet une variance si et seulement si G_X est deux fois dérivable en 1.

Théorème : Somme de variables indépendantes

Si X et Y sont à valeurs dans \mathbb{N} et indépendantes, alors, pour tout $t \in]-r, r[$ où $r = \min(R_X, R_Y)$,

$$G_{X+Y}(t) = \mathbf{E}(t^{X+Y}) = \mathbf{E}(t^X)\mathbf{E}(t^Y) = G_X(t)G_Y(t)$$

→ Inégalités de concentration et convergence

Lemme : Inégalité de Markov

Si X est à valeurs positives et admet une espérance,

$$\forall a > 0, \quad \mathbf{P}(X \geq a) \leq \frac{\mathbf{E}(X)}{a}$$

Proposition : Inégalité de Bienaymé-Tchebychev

Si X admet un moment d'ordre 2,

$$\forall \varepsilon > 0, \quad \mathbf{P}(|X - \mathbf{E}(X)| \geq \varepsilon) \leq \frac{\mathbf{V}(X)}{\varepsilon^2}$$

Théorème : Loi faible des grands nombres

Soit $(X_n)_{n \geq 1}$ une suite de variables indépendantes et de même loi, admettant un moment d'ordre 2.

En notant m l'espérance commune et $S_n = \sum_{i=1}^n X_i$,

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0$$

→ Lois usuelles

Nom	Notation	$X(\Omega)$	$\mathbf{P}(X = k)$	$\mathbf{E}(X)$	$\mathbf{V}(X)$
Bernoulli	$\mathcal{B}(p)$	$\{0, 1\}$	$\begin{cases} p & \text{si } k = 1 \\ q & \text{si } k = 0 \end{cases}$	p	pq
Binomiale	$\mathcal{B}(n, p)$	$\llbracket 0, n \rrbracket$	$\binom{n}{k} p^k q^{n-k}$	np	npq
Uniforme	$\mathcal{U}(\llbracket 1, n \rrbracket)$	$\llbracket 1, n \rrbracket$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Géométr.	$\mathcal{G}(p)$	\mathbb{N}^*	$q^{k-1}p$	$\frac{1}{p}$	$\frac{q}{p^2}$
Poisson	$\mathcal{P}(\lambda)$	\mathbb{N}	$e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ

- Si $X_1, \dots, X_n \sim \mathcal{B}(m_i, p)$ sont mutuellement indépendantes, alors $X_1 + \dots + X_n \sim \mathcal{B}(m_1 + \dots + m_n, p)$.
- Si $X_1, \dots, X_n \sim \mathcal{P}(\lambda_i)$ sont mutuellement indépendantes, alors $X_1 + \dots + X_n \sim \mathcal{P}(\lambda_1 + \dots + \lambda_n)$.
- Si pour tout $n \in \mathbb{N}^*$, $X_n \sim \mathcal{B}(n, p_n)$ et $\lim_{n \rightarrow +\infty} np_n = \lambda$,

$$\forall k \in \mathbb{N}, \quad \mathbf{P}(X_n = k) \xrightarrow{n \rightarrow +\infty} e^{-\lambda} \frac{\lambda^k}{k!}$$