Chapitre 5 : Probabilités discrètes

I Espace de probabilité discret.

Un univers Ω est dit discret s'il est au plus dénombrable.

Définition : une probabilité $\mathbb{P}$ sur un univers Ω discret est une application

$$\mathbb{P} : \mathcal{P}(\Omega) \longrightarrow [0,1]$$
$$A \longmapsto \mathbb{P}(A)$$

satisfaisant

* $\mathbb{P}(\Omega) = 1$
* pour toute famille $(A_i)_{i\in I}$ d'événements deux à deux disjoints, où I est fini ou dénombrable, on a

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

Comme dans le cas fini, définir une probabilité $\mathbb{P}$ sur Ω discret revient à considérer une famille de poids $\left(p_\omega = \mathbb{P}(\{\omega\})\right)_{\omega\in\Omega}$ satisfaisant $p_\omega \geq 0$ pour tout $\omega \in \Omega$ et $\sum_{\omega\in\Omega} p_\omega = 1$.

somme infinie si Ω de cardinal infini

Dans ce cas, pour tout évènement A,

$$\mathbb{P}(A) = \sum_{\omega\in A} p_\omega$$

somme infinie si A de cardinal infini.

Exemple: $\Omega = \mathbb{N}^*$, $p_n = \frac{1}{2^n}$ pour tout $n\in\mathbb{N}^*$.

On a bien $p_n \geq 0$ pour tout $n\in\mathbb{N}^*$ et

$$\sum_{n\in\mathbb{N}^*} p_n = \sum_{n=1}^{+\infty} p_n = \frac{1}{2}\sum_{k=0}^{+\infty}\frac{1}{2^k} = \frac{1}{2}\,\frac{1}{1-1/2} = 1.$$

Si $A = \{n \in \mathbb{N}^*, \ n \text{ pair}\}$,

$$P(A) = \sum_{\substack{n \in \mathbb{N}^* \\ n \text{ pair}}} \frac{1}{2^n} = \sum_{k=1}^{+\infty} \frac{1}{2^{2k}}$$

$$= \frac{1}{4} \sum_{i=0}^{+\infty} \frac{1}{4^i} = \frac{1}{4} \frac{1}{1-\frac{1}{4}} = \frac{1}{3}.$$

Les ppt́és vues dans le cas fini (proba d'une union, passage au complémentaire, conditionnement...) restent vraies dans le cas discret.

Prop : Soit $(\Omega, \mathbb{P})$ un espace de probabilité discret et $(A_n)_{n \in \mathbb{N}}$ une suite d'événements croissante ($A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$) et $(B_n)_{n \in \mathbb{N}}$ une suite d'événements décroissante ($B_{n+1} \subset B_n$ pour tout $n \in \mathbb{N}$).

Alors 1) $\lim_{n \to +\infty} P(A_n) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right)$

2) $\lim P(B_n) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} B_n\right)$

Preuve.

1) On définit $C_0 = A_0$ et, pour tout $n \geq 1$,

$$C_n = A_n \setminus A_{n-1}.$$

Alors

* les C_i sont deux à deux disjoints

si $i < j$, $C_i \subset A_{j-1}$ et

$$C_i \cap C_j \subset \underbrace{A_{j-1} \cap C_j}_{= A_{j-1} \cap (A_j \setminus A_{j-1})} = \emptyset$$

* on a $\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \in \mathbb{N}} C_n$.

En effet, puisque pour tout $n \in \mathbb{N}$, $C_n \subset A_n$,

$$\bigcup_{n \in \mathbb{N}} C_n \subset \bigcup_{n \in \mathbb{N}} A_n.$$

Soit maintenant $\omega \in \bigcup_{n \in \mathbb{N}} A_n$, Il existe $n \in \mathbb{N}$ tel que $\omega \in A_n$. Mais

$$A_n = A_n \setminus A_{n-1} \cup A_{n-1} = C_n \cup \underbrace{(A_{n-1} \setminus A_{n-2})}_{C_{n-1}} \cup A_{n-2}$$

$$= \bigcup_{k=0}^{n} C_n.$$

Donc il existe $k \in \{0, \dots, n\}$ tel que $\omega \in C_k$

et donc $\bigcup_{n\in\mathbb{N}} A_n \subset \bigcup_{n\in\mathbb{N}} C_n$.

Ainsi,

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathbb{P}\Big(\bigcup_{n\in\mathbb{N}} \overset{\text{2 à 2 disjoints}}{C_n}\Big)$$
$$= \sum_{n\in\mathbb{N}} \mathbb{P}(C_n)$$
$$= \lim_{n\to+\infty} \sum_{k=0}^{n} \mathbb{P}(C_k)$$
$$= \lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{k=0}^{n} C_k\right)$$
$$= \lim_{n\to+\infty} \mathbb{P}(A_n)$$

2) On passe au complémentaire :
on définit $A_n = B_n^c$, $(A_n)_{n\in\mathbb{N}}$ est croissante

et
$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} B_n\right) = 1 - \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right)$$
$$= 1 - \lim_{n\to+\infty} \mathbb{P}(A_n)$$
$$= \lim_{n\to+\infty} \left(1 - \mathbb{P}(A_n)\right)$$
$$= \lim_{n\to+\infty} \mathbb{P}(B_n).$$

II Variables aléatoires discrètes

1) Définition

Déf : On appelle v.a. discrète toute application $X : \Omega \to \mathbb{R}$ où $(\Omega, \mathbb{P})$ est un espace de probabilité discret.

L'ensemble E des valeurs possibles de X est au plus dénombrable, et on appelle loi de X la probabilité $\mathbb{P}_X$ sur E associée à la famille $\left(p_x = \mathbb{P}(X=x)\right)_{x \in E}$.

2) Lois discrètes classiques.

* Loi géométrique

X est de loi géométrique de paramètre $p \in [0,1]$ si X est à valeurs dans $\mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$

$$\mathbb{P}(X=k) = p(1-p)^{k-1}.$$

On note $X \sim \mathcal{G}(p)$.

Si l'on considère un schéma de Bernoulli de n épreuves indépendants avec proba de succès p, alors la proba de l'événement "le premier succès a lieu à la kième expérience" avec $k \in \{1, \dots, n\}$ est

$$\underbrace{(1-p)\times(1-p)\times(1-p)\times\dots(1-p)}_{k-1 \text{ fois}} \times p = (1-p)^{k-1} p.$$

Ex: on lance un dé à six faces équilibré jusqu'à obtenir un 6, et on note X le nbre de lancers. Alors $X \sim \mathcal{G}(\frac{1}{6})$.

Rq: La modélisation rigoureuse de cette infinité de lancers de dés sort du cadre de ce cours ($\{0,1\}^{\mathbb{N}}$ est non dénombrable).

* Loi de Poisson.

X suit la loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$ (on note $X \sim \mathcal{P}(\lambda)$) si X est à valeurs dans $\mathbb{N}$ et pour tout $k \in \mathbb{N}$

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

Utile pour modéliser un nombre d'événements dans un intervalle de temps (arrivées, connexions à un serveur) le nombre d'impuretés dans un cable ....

Théorème (événements rares). Soient $(p_n)_{n\in\mathbb{N}}$ une suite de réels de $[0,1]$ et un réel $\lambda > 0$ tels que $np_n \xrightarrow[n\to+\infty]{} \lambda$. Alors pour tout $k \in \mathbb{N}$,

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} \xrightarrow[n\to+\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Preuve :

On a

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n!}{k!\,(n-k)!} p_n^k (1-p_n)^{n-k}$$

$$= \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k}$$

$$= \frac{n(n-1)(n-2)\dots(n-k+1)}{n \times n \times n \dots \times n} \frac{(np_n)^k}{k!} (1-p_n)^{n-k}$$

$$\underset{n\to+\infty}{\sim} \frac{\lambda^k}{k!} (1-p_n)^{n-k}$$

et

$$(1-p_n)^{n-k} = \exp\left((n-k)\ln(1-p_n)\right)$$

$$= \exp\left((n-k)\left(-p_n + \underset{n\to\infty}{o}(p_n)\right)\right)$$

$$= \exp\left(-p_n(n-k) + \underset{n\to+\infty}{o}(np_n)\right)$$

$$= \exp\left(-\lambda + \underset{n\to+\infty}{o}(1)\right)$$

On a bien

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} \xrightarrow[n\to+\infty]{} \frac{\lambda^k}{k!} e^{-\lambda}.$$

III. Espérance.

Une v.a. discrète n'admet pas toujours d'espérance

<u>Ex:</u> Si X est à valeurs dans $\mathbb{N}$ et satisfait $\mathbb{P}(X=n) = \frac{6}{\pi^2 n^2}$ pour tout $n \in \mathbb{N}$ (on a $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$), alors la famille

$$\left(n\mathbb{P}(X=n)\right)_{n\in\mathbb{N}^*} = \left(\frac{6}{\pi^2 n}\right)_{n\in\mathbb{N}}$$ n'est pas sommable.

<u>Définition</u> Soit X une v.a. discrètes ayant pour valeurs possibles $\{x_i, i\in I\}$ où les x_i sont distincts. Alors l'espérance de X est bien définie si la famille

$$\left(|x_i| \mathbb{P}(X=x_i)\right)_{i\in I}$$ est sommable

et dans ce cas on la note $\mathbb{E}[X]$ et on a

$$\mathbb{E}[X] = \sum_{i\in I} x_i \mathbb{P}(X=x_i).$$

<u>Proposition</u> Soient X une v.a. discrète ayant pour valeurs possibles $\{x_i, i \in I\}$ où les x_i sont distincts et $f: \mathbb{R} \to \mathbb{R}$ une fonction.

Alors $Y = f(X)$ admet une espérance si et seulement si la famille

$$\left(|f(x_i)| \, \mathbb{P}(X = x_i) \right)_{i \in I} \text{ est sommable}$$

et dans ce cas

$$\mathbb{E}[f(X)] = \sum_{i \in I} f(x_i) \, \mathbb{P}(X = x_i).$$

<u>Preuve</u>:

Notons $\{y_j : j \in J\}$ les valeurs possibles de Y, où les y_j sont distincts.

Alors
$$|y_j| \, \mathbb{P}(Y = y_j) = |y_j| \sum_{\substack{i \in I \\ f(x_i) = y_j}} \mathbb{P}(X = x_i)$$
$$= \sum_{i \in I} |f(x_i)| \, \mathbb{P}(X = x_i)$$

Donc, par sommation par paquet,

$\left(|g_j| \, \mathbb{P}(Y = y_j) \right)_{j \in J}$ est sommable

si et seulement si $\left(|f(x_i)| \, \mathbb{P}(X = x_i) \right)_{i \in I}$

l'est, et à nouveau par sommation par

paquet

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

<u>Rq</u> Si $X_1, \dots, X_n$ sont des v.a. discrètes

à valeurs dans $E_1, \dots, E_n$ sous-ensembles

discrets de $\mathbb{R}$, et $f : \mathbb{R}^n \to \mathbb{R}$ est une

fonction, $Y = f(X_1, \dots, X_n)$ admet une

espérance si et seulement si la famille

$$\left(|f(x_1, \dots, x_n)| \, \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) \right)_{(x_1, \dots, x_n) \in E_1 \times \dots \times E_n}$$ est sommable

et dans ce cas

$$\mathbb{E}\left[f(X_1, \dots, X_n) \right] = \sum_{(x_1, \dots, x_n) \in E_1 \times \dots \times E_n} f(x_1, \dots, x_n) \, \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$$

L'espérance dans le cas discret est linéaire, comme dans le cas fini.

<u>Définition</u> : soit X une v.a. discrète et $p \in \mathbb{R}_+^*$. On appelle, s'il existe, $\mathbb{E}[|X|^p]$ le moment d'ordre p de X.

<u>Rq</u> : si $0 < p < q$, puisque l'on a, pour tout $x \in \mathbb{R}$, $|x|^p \leq 1 + |x|^q$, alors si X admet un moment d'ordre q alors X admet un moment d'ordre p.

<u>Rq</u> : On peut en fait démontrer l'inégalité de Jensen suivante : si $\varphi : \mathbb{R} \to \mathbb{R}$ est convexe et X et $\varphi(X)$ admettent une espérance,

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)].$$

Puisque $\varphi : x \longmapsto |x|^{q/p}$ est convexe ($q/p > 1$)

$$\left(\mathbb{E}[|X|^p]\right)^{q/p} \leq \mathbb{E}\left[(|X|^p)^{q/p}\right] = \mathbb{E}[|X|^q].$$

Définition : Soit X une v.a discrète admettant un moment d'ordre 2 (i.e. $\mathbb{E}[X^2]$ existe).
On définit la variance $Var(X)$ de X par

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

Définition : Soient X et Y deux v.a. discrètes admettant un moment d'ordre 2. On définit la covariance $Cov(X,Y)$ par

$$Cov(X,Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$$
$$= \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y].$$

Rq : si X et Y admettent un moment d'ordre 2, $\mathbb{E}[XY]$ est bien défini puisque

$$|XY| \leq \frac{X^2}{2} + \frac{Y^2}{2}.$$

On peut également justifier l'existence de $\mathbb{E}[XY]$ à l'aide de l'inégalité de Cauchy-Schwarz.

<u>Proposition</u> : Si $X:\Omega\to\mathbb{R}$ et $Y:\Omega\to\mathbb{R}$ sont deux v.a. discrètes admettant un moment d'ordre 2,

$$\mathbb{E}[|XY|] \leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$$

<u>Preuve</u> (utilisant C.S. pour les sommes finies)

Supposons que X prend ses valeurs dans $E_1 \subset \mathbb{R}$ discret et Y dans $E_2 \subset \mathbb{R}$ discret.

Soit $J \subset E_1 \times E_2$ fini. Il existe $J_1 \subset E_1$, $J_2 \subset E_2$ finis tels que

$$J \subset J_1 \times J_2$$

On a

$$\sum_{(x,y)\in J} |xy|\, \mathbb{P}(X=x, Y=y)$$

$$= \sum_{(x,y)\in J} |x|\sqrt{\mathbb{P}(X=x,Y=y)}\ |y|\sqrt{\mathbb{P}(X=x,Y=y)}$$

C.S. cas fini

$$\leq \sqrt{\sum_{(x,y)\in J}\left(|x|\sqrt{\mathbb{P}(X=x,Y=y)}\right)^2}\sqrt{\sum_{(x,y)\in J}\left(|y|\sqrt{\mathbb{P}(X=x,Y=y)}\right)^2}$$

termes positifs

$$\leq \sqrt{\sum_{x\in J_1} x^2 \underbrace{\sum_{y\in J_2}\mathbb{P}(X=x,Y=y)}_{\leq \mathbb{P}(X=x)}}\sqrt{\sum_{y\in J_2} y^2 \underbrace{\sum_{x\in J_1}\mathbb{P}(X=x,Y=y)}_{\leq \mathbb{P}(Y=y)}}$$

$$\leq \sqrt{\sum_{x\in E_1} x^2\, \mathbb{P}(X=x)}\sqrt{\sum_{y\in E_2} y^2\, \mathbb{P}(Y=y)}$$

$$\leq \sqrt{\mathbb{E}[X^2]}\sqrt{\mathbb{E}[Y^2]}$$

Les propriétés démontrées pour la variance dans le cas fini restent vraies dans le cas discret. Cela repose en particulier sur la proposition suivante :

<u>Proposition</u> : Soit Ω un univers discret, $\mathbb{P}$ une probabilité sur Ω et $X_1, \dots, X_n$ des v.a sur Ω à valeurs dans $E_1, \dots, E_n$ discrets et indépendantes (vérifient, pour tout $(x_1, \dots, x_n) \in E_1 \times \dots \times E_n$
$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \dots \mathbb{P}(X_n = x_n)$).
Alors pour toutes fonctions $f_1, \dots, f_n$ de $\mathbb{R}$ dans $\mathbb{R}$ telles que $\mathbb{E}[f_1(X_1) \dots f_n(X_n)]$ et $\mathbb{E}[f_k(X_k)]$ existent pour tout $k \in \{1, \dots, n\}$,

$$\mathbb{E}[f_1(X_1) \dots f_n(X_n)] = \mathbb{E}[f_1(X_1)] \dots \mathbb{E}[f_n(X_n)].$$

IV Fonction de répartition.

Définition : Soit Ω un univers discret, $\mathbb{P}$ une probabilité sur Ω et X une v.a. sur Ω, prenant ses valeurs dans une partie discrète E de $\mathbb{R}$. La fonction de répartition de X, notée F_X, est définie par

$$F_X : \begin{array}{ccl} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \mathbb{P}(X \leq t) = \sum_{\substack{x \in E \\ x \leq t}} \mathbb{P}(X = x) \end{array}$$

Ex : si $X \sim \mathcal{B}(p)$ avec $p \in [0,1]$,

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ p & \text{si } t \in [0,1[\\ 1 & \text{si } t \geq 1 \end{cases}$$