Mathématiques

PSI

Lycée Couffignal

CHAPITRE 11

Variables aléatoires discrètes

Plan du chapitre

Introduction

Nous avons précédemment généralisé la notion d'espace probabilisé et ses propriétés au cas des univers infinis.

Dans ce chapitre, nous commencerons par étudier succinctement la notion de famille sommable en vue de manipuler des sommes indexées par un ensemble infini. Ensuite, nous étendrons la notion de variable aléatoire vue en première année au cadre d'un univers infini.

Dans tout le chapitre, on désigne par $\mathbb{K}$ le corps $\mathbb{R}$ ou le corps $\mathbb{C}$.

Partie I Familles sommables

Dans la suite de ce chapitre, nous aurons besoin de manipuler des sommes de familles d'éléments de $\mathbb{K}$ indexées par un ensemble au plus dénombrable. Afin de rendre ces opérations rigoureuses, nous allons étudier la manière d'associer à une famille au plus dénombrable $(x_i)_{i\in I}$ d'éléments de $\mathbb{K}$ un nombre correspondant intuitivement à la somme de tous les éléments de la famille. Les notions vues dans cette partie ne feront l'objet d'aucune évaluation spécifique.

Si l'ensemble I est fini, il n'y a pas de difficulté : on peut écrire $I = \{i_1, \dots, i_n\}$ où les éléments $i_1, \dots, i_n \in I$ sont deux à deux distincts et définir la somme de la famille $(x_i)_{i\in I}$ par

$$\sum_{i\in I} x_i = \sum_{k=1}^{n} x_{i_k}.$$

Il est clair que la valeur de cette somme ne dépend pas de la manière dont on a numérotée les éléments de I, i.e. de l'ordre de sommation des éléments de la famille.

Si l'ensemble I est dénombrable, il est tentant de procéder de manière analogue. On peut écrire $I = \{i_k \mid k \in \mathbb{N}\}$ où $(i_k)_{k\in\mathbb{N}}$ est une suite injective d'éléments de I et définir la somme de la famille $(x_i)_{i\in I}$ par

$$\sum_{i\in I} x_i = \sum_{k=0}^{+\infty} x_{i_k}.$$

Cependant, cette approche se heurte à deux difficultés en général :

- la convergence de la série $\sum\limits_{k\geqslant 0} x_{i_k}$ dépend de la manière dont on a numéroté les éléments de I ;
- même si la série $\sum\limits_{k\geqslant 0} x_{i_k}$ converge, sa somme dépend de la manière dont on a numéroté les éléments de I.

Nous allons voir comment éviter ces difficultés dans la suite de cette partie.

Exemple 1 : On peut démontrer qu'on a l'égalité

$$\ln(2) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k+2} + \cdots$$

Si on s'autorise à effectuer une permutation des termes de la somme ci-dessus, on obtient

$$\begin{aligned}\left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \cdots + \left(\frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k}\right) + \cdots &= \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \cdots + \left(\frac{1}{4k-2} - \frac{1}{4k}\right) + \cdots \\ &= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{1}{2k-1} - \frac{1}{2k} + \cdots\right) = \frac{1}{2}\ln(2).\end{aligned}$$

Bien que les termes additionnés soient les mêmes, nous avons obtenu deux sommes différentes.

I.A - Familles sommables de réels positifs

On commence par quelques remarques sur les propriétés de l'ensemble $[0,+\infty]$.

- La relation d'ordre sur les nombres de $[0,+\infty[$ s'étend naturellement à l'ensemble $[0,+\infty]$. De plus, toute partie A de $[0,+\infty]$ admet une borne supérieure dans $[0,+\infty]$ que l'on note $\sup(A)$.
- L'addition de $[0,+\infty[$ s'étend naturellement à l'ensemble $[0,+\infty]$ en posant

$$\forall x \in [0,+\infty], \quad x+(+\infty) = (+\infty)+x = +\infty.$$

- La multiplication de $[0,+\infty[$ s'étend naturellement à l'ensemble $[0,+\infty]$ en posant

$$0\times(+\infty) = (+\infty)\times 0 = 0 \quad \text{et} \quad \forall x \in]0,+\infty], \quad x\times(+\infty) = (+\infty)\times x = +\infty.$$

Rappelons que si $(x_j)_{j\in J}$ est une famille finie d'éléments de $[0,+\infty]$, nous avons vu dans l'introduction de cette partie que nous pouvions définir $\sum\limits_{j\in J} x_j$ sans difficulté.

Définition (Somme d'une famille de $[0,+\infty]$) : La somme d'une famille $(x_i)_{i\in I}$ au plus dénombrable de $[0,+\infty]$ est l'élément de $[0,+\infty]$ définie par

$$\sum_{i\in I} x_i = \sup\left\{\sum_{j\in J} x_j \in [0,+\infty] \;\middle|\; J\in\mathscr{P}(I) \text{ est fini}\right\}.$$

Remarques 1 :

a) Par convention, si $I=\varnothing$ alors $\sum\limits_{i\in I} x_i = 0$.

b) Si I est un ensemble fini, on retrouve la notion habituelle de somme.

c) S'il existe un indice $i_0\in I$ tel que $x_i = +\infty$, alors $\sum\limits_{i\in I} x_i = +\infty$.

d) On peut vérifier que la définition ci-dessus ne dépend pas de la manière de décrire la famille sommée. Plus précisément, si $\varphi : J\to I$ est une bijection, alors on a $\sum\limits_{i\in I} x_i = \sum\limits_{j\in J} x_{\varphi(j)}$.

Proposition 1 : Si $\sum\limits_{n\geqslant 0} u_n$ est une série de réels positifs, alors on a

$$\sum_{n\in\mathbb{N}} u_n = \begin{cases} \sum\limits_{n=0}^{+\infty} u_n & \text{si la série converge} \\ +\infty & \text{si la série diverge.}\end{cases}$$

Remarque 2 : On en déduit que la nature et la somme d'une série à termes positifs ne dépendent pas de l'ordre de sommation des termes.

Exemple 2 : On a $\sum\limits_{n\in\mathbb{N}} \dfrac{1}{2^n} = \sum\limits_{n=0}^{+\infty}\dfrac{1}{2^n} = 2$ et $\sum\limits_{n\in\mathbb{N}^*}\dfrac{1}{n} = +\infty$.

Proposition (Linéarité et croissance de la somme) : Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles au plus dénombrables de $[0,+\infty]$.

(i) Pour tout $(\lambda,\mu)\in[0,+\infty]^2$, on a $\sum\limits_{i\in I}(\lambda x_i+\mu y_i) = \lambda\left(\sum\limits_{i\in I} x_i\right)+\mu\left(\sum\limits_{i\in I} y_i\right)$.

(ii) Si $x_i \leqslant y_i$ pour tout $i\in I$, alors on a $\sum\limits_{i\in I} x_i \leqslant \sum\limits_{i\in I} y_i$.

Rappelons qu'une famille $(I_n)_{n\in\mathbb{N}}$ est une partition d'un ensemble I si

(i) les ensembles de la famille $(I_n)_{n\in\mathbb{N}}$ sont non vides et deux à deux disjoints,

(ii) les ensembles de la famille $(I_n)_{n\in\mathbb{N}}$ recouvrent I, i.e. $I = \bigcup_{n\in\mathbb{N}} I_n$.

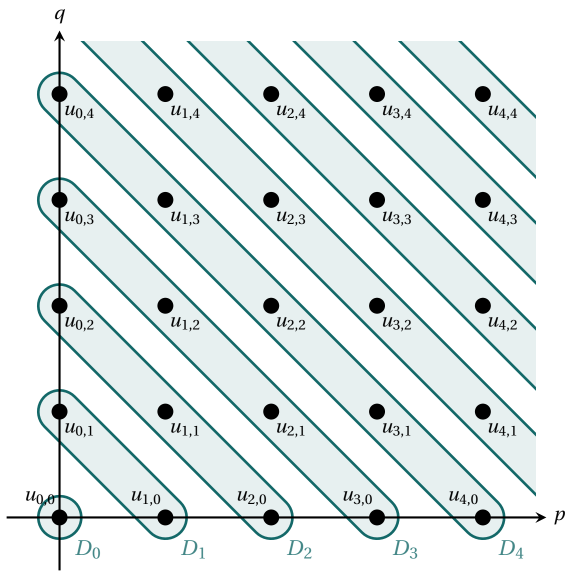
Théorème de sommation par paquet : Si $(x_i)_{i\in I}$ une famille au plus dénombrable de $[0,+\infty]$ et si $(I_n)_{n\in\mathbb{N}}$ est une partition de I, alors on a

$$\sum_{i\in I} x_i = \sum_{n\in\mathbb{N}}\left(\sum_{i\in I_n} x_i\right).$$

Exemple 3 : Notons $D_n = \left\{(p,q)\in\mathbb{N}^2 \mid p+q=n\right\}$ pour tout $n\in\mathbb{N}$.

La famille $(D_n)_{n\in\mathbb{N}}$ est une partition de $\mathbb{N}^2$, donc par le théorème ci-dessus, on a

$$\begin{aligned}\sum_{(p,q)\in\mathbb{N}^2}\underbrace{\frac{1}{(p+q+1)^2}}_{u_{p,q}} &= \sum_{n\in\mathbb{N}}\sum_{(p,q)\in D_n}\frac{1}{(p+q+1)^2}\\ &= \sum_{n\in\mathbb{N}}\frac{\mathrm{Card}(D_n)}{(n+1)^2}\\ &= \sum_{n\in\mathbb{N}}\frac{n+1}{(n+1)^2}\\ &= \sum_{n\in\mathbb{N}}\frac{1}{n+1} = +\infty.\end{aligned}$$



Théorème de Fubini : Si $(u_{i,j})_{(i,j)\in I\times J}$ est une famille au plus dénombrable de $[0,+\infty]$, alors on a

$$\sum_{(i,j)\in I\times J} u_{i,j} = \sum_{i\in I}\left(\sum_{j\in J} u_{i,j}\right) = \sum_{j\in J}\left(\sum_{i\in I} u_{i,j}\right).$$

En particulier, si $(x_i)_{i\in I}$ et $(y_j)_{j\in J}$ sont deux familles au plus dénombrables de $[0,+\infty]$, alors on a

$$\sum_{(i,j)\in I\times J} x_i y_j = \left(\sum_{i\in I} x_i\right)\times\left(\sum_{j\in J} y_j\right).$$

Exemple 4 : En utilisant le théorème de Fubini, on a

$$\sum_{(p,q)\in\mathbb{N}^2}\frac{p^q}{\mathrm{e}^{2p}q!} = \sum_{p\in\mathbb{N}}\left(\sum_{q\in\mathbb{N}}\frac{p^q}{\mathrm{e}^{2p}q!}\right) = \sum_{p\in\mathbb{N}}\mathrm{e}^{-2p}\left(\sum_{q\in\mathbb{N}}\frac{p^q}{q!}\right) = \sum_{p\in\mathbb{N}}\mathrm{e}^{-2p}\mathrm{e}^{p} = \sum_{p\in\mathbb{N}}\mathrm{e}^{-p} = \frac{\mathrm{e}}{\mathrm{e}-1}.$$

Bilan : En résumé, les règles pour manipuler la somme d'une famille de $[0,+\infty]$ sont les mêmes que celles pour manipuler les sommes finies.

I.B - Familles sommables de réels ou de complexes

Définition (Famille sommable de réels ou complexes) : Une famille $(x_i)_{i\in I}$ au plus dénombrable d'éléments de $\mathbb{K}$ est dite sommable si $\sum_{i\in I} |x_i| < +\infty$.

Proposition 2 : Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles au plus dénombrables d'éléments de $\mathbb{K}$. Si $|x_i| \leqslant |y_i|$ pour tout $i \in I$ et que la famille $(y_i)_{i\in I}$ est sommable, alors la famille $(x_i)_{i\in I}$ est sommable.

Remarques 3 :

a) Si $(x_i)_{i\in I}$ est une famille au plus dénombrable de nombres réels, alors on note

$$\forall i \in I, \quad x_i^+ = \max(x_i, 0) \quad \text{et} \quad x_i^- = \max(-x_i, 0).$$

Les familles $\left(x_i^+\right)_{i\in I}$ et $\left(x_i^-\right)_{i\in I}$ sont des familles au plus dénombrables de $[0, +\infty]$. De plus, on a

$$\forall i \in I, \quad |x_i| = x_i^+ + x_i^- \quad \text{et} \quad x_i = x_i^+ - x_i^-,$$

donc on a en particulier que

$$\forall i \in I, \quad 0 \leqslant x_i^+ \leqslant |x_i| \quad \text{et} \quad 0 \leqslant x_i^- \leqslant |x_i|.$$

On conclut par croissance de la somme que si $(x_i)_{i\in I}$ est sommable, alors $\left(x_i^+\right)_{i\in I}$ et $\left(x_i^-\right)_{i\in I}$ sont sommables, donc leur somme $\sum_{i\in I} x_i^+$ et $\sum_{i\in I} x_i^-$ sont des nombres réels.

b) Si $(x_i)_{i\in I}$ est une famille au plus dénombrable de nombres complexes, alors on a

$$\forall i \in I, \quad 0 \leqslant |\mathrm{Re}(x_i)| \leqslant |x_i| \quad \text{et} \quad 0 \leqslant |\mathrm{Im}(x_i)| \leqslant |x_i|.$$

On conclut par croissance de la somme que si $(x_i)_{i\in I}$ est sommable, alors $\big(\mathrm{Re}(x_i)\big)_{i\in I}$ et $\big(\mathrm{Im}(x_i)\big)_{i\in I}$ sont des familles sommables de nombres réels.

Définition (Somme d'une famille sommable) : Soit $(x_i)_{i\in I}$ une famille sommable au plus dénombrable d'éléments de $\mathbb{K}$.

(i) Si $(x_i)_{i\in I}$ est une famille de nombres réels, alors la somme de $(x_i)_{i\in I}$ est le nombre réel

$$\sum_{i\in I} x_i = \left(\sum_{i\in I} x_i^+\right) - \left(\sum_{i\in I} x_i^-\right).$$

(ii) Si $(x_i)_{i\in I}$ est une famille de nombres complexes, alors la somme de $(x_i)_{i\in I}$ est le nombre complexe

$$\sum_{i\in I} x_i = \left(\sum_{i\in I} \mathrm{Re}(x_i)\right) + \mathrm{i}\left(\sum_{i\in I} \mathrm{Im}(x_i)\right).$$

Remarques 4 :

a) Si $I = \varnothing$, alors $(x_i)_{i\in I}$ est sommable et on a $\sum_{i\in I} x_i = 0$.

b) Si I est un ensemble fini, alors $(x_i)_{i\in I}$ est sommable et on retrouve la notion habituelle de somme.

Proposition 3 : Soient $(x_i)_{i\in I}$ une famille au plus dénombrable d'éléments de $\mathbb{K}$ et $(i_n)_{n\in\mathbb{N}}$ une suite injective d'éléments de I telle que $I = \{i_n \mid n \in \mathbb{N}\}$. La famille $(x_i)_{i\in I}$ est sommable si et seulement si la série $\sum_{n\geqslant 0} x_{i_n}$ converge absolument. Dans ce cas, on a

$$\sum_{i\in I} x_i = \sum_{n=0}^{+\infty} x_{i_n}.$$

Remarque 5 : On en déduit que la nature et la somme d'une série absolument convergente ne dépendent pas de l'ordre de sommation des termes.

Proposition (Linéarité et croissance de la somme) : Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles sommables au plus dénombrables d'éléments de $\mathbb{K}$.

(i) Pour tout $(\lambda,\mu)\in\mathbb{C}^2$, la famille $(\lambda x_i+\mu y_i)_{i\in I}$ est sommable et on a $\displaystyle\sum_{i\in I}(\lambda x_i+\mu y_i)=\lambda\left(\sum_{i\in I}x_i\right)+\mu\left(\sum_{i\in I}y_i\right)$.

(ii) Si $x_i\leqslant y_i$ pour tout $i\in I$, alors on a $\displaystyle\sum_{i\in I}x_i\leqslant\sum_{i\in I}y_i$.

Théorème de sommation par paquet : Si $(x_i)_{i\in I}$ une famille sommable au plus dénombrable d'éléments de $\mathbb{K}$. et si $(I_n)_{n\in\mathbb{N}}$ est une partition de I, alors la famille $\displaystyle\left(\sum_{i\in I_n}x_i\right)_{n\in\mathbb{N}}$ est sommable et on a

$$\sum_{i\in I}x_i=\sum_{n\in\mathbb{N}}\left(\sum_{i\in I_n}x_i\right).$$

Théorème de Fubini : Si $(u_{i,j})_{(i,j)\in I\times J}$ est une famille sommable au plus dénombrable d'éléments de $\mathbb{K}$, alors les familles $\displaystyle\left(\sum_{j\in J}u_{i,j}\right)_{i\in I}$ et $\displaystyle\left(\sum_{i\in I}u_{i,j}\right)_{j\in J}$ sont sommables et on a

$$\sum_{(i,j)\in I\times J}u_{i,j}=\sum_{i\in I}\left(\sum_{j\in J}u_{i,j}\right)=\sum_{j\in J}\left(\sum_{i\in I}u_{i,j}\right).$$

En particulier, si $(x_i)_{i\in I}$ et $(y_j)_{j\in J}$ sont deux familles sommables au plus dénombrables d'éléments de $\mathbb{K}$, alors la famille $(x_iy_j)_{(i,j)\in I\times J}$ est une famille sommable et on a

$$\sum_{(i,j)\in I\times J}x_iy_j=\left(\sum_{i\in I}x_i\right)\times\left(\sum_{j\in J}y_j\right).$$

Exemple 5 : Soit $z\in\mathbb{C}$ avec $|z|<1$. La famille $\left(z^{ij}\right)_{(i,j)\in\mathbb{N}^*\times\mathbb{N}^*}$ est sommable, car on a

$$\sum_{(i,j)\in\mathbb{N}^*\times\mathbb{N}^*}|z|^{ij}=\sum_{i\in\mathbb{N}^*}\sum_{j\in\mathbb{N}^*}\left(|z|^i\right)^j=\sum_{i\in\mathbb{N}^*}\frac{|z|^i}{1-|z|^i}\leqslant\sum_{i\in\mathbb{N}^*}\frac{|z|^i}{1-|z|}=\frac{1}{(1-|z|)^2}<+\infty.$$

On en déduit que le théorème précédent s'applique, donc on a

$$\sum_{(i,j)\in\mathbb{N}^*\times\mathbb{N}^*}z^{ij}=\sum_{i\in\mathbb{N}^*}\sum_{j\in\mathbb{N}^*}\left(z^i\right)^j=\sum_{i\in\mathbb{N}^*}\frac{z^i}{1-z^i}.$$

Bilan : En résumé, une fois qu'on a vérifié la sommabilité d'une famille d'éléments de $\mathbb{K}$, les règles pour manipuler la somme de cette famille sont les mêmes que celles pour manipuler les sommes finies.

Partie II Variables aléatoires discrètes

II.A - Généralités

Définition (Variable aléatoire discrète) : Soit $(\Omega, \mathscr{A})$ un espace probabilisable. Une variable aléatoire discrète sur $(\Omega, \mathscr{A})$ est une application $X : \Omega \to E$ où E est un ensemble telle que

(i) l'ensemble $X(\Omega)$ est au plus dénombrable;

(ii) pour tout $x \in X(\Omega)$, on a $X^{-1}(\{x\}) \in \mathscr{A}$.

Définition (Variable aléatoire réelle) : Une variable aléatoire X sur $(\Omega, \mathscr{A})$ est dite réelle si $X(\Omega) \subset \mathbb{R}$.

Définition (Variable aléatoire complexe) : Une variable aléatoire X sur $(\Omega, \mathscr{A})$ est dite complexe si $X(\Omega) \subset \mathbb{C}$.

Remarque 6 : L'ensemble $X(\Omega)$ est l'ensemble des valeurs prises par la variable aléatoire X.

Notation : Si $X : \Omega \to E$ est une variable aléatoire discrète sur $(\Omega, \mathscr{A})$, alors pour tout $A \subset E$ et tout $x \in E$, on note

$$(X \in A) = X^{-1}(A) = \{\omega \in \Omega \mid X(\omega) \in A\}, \qquad (X = x) = X^{-1}(\{x\}) = \{\omega \in \Omega \mid X(\omega) = x\}.$$

De plus, si X est une variable aléatoire réelle ou à valeurs dans $[0, +\infty]$, alors pour tout $x \in \mathbb{R}$, on note

$$(X \leqslant x) = X^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leqslant x\}, \qquad (X \geqslant x) = X^{-1}([x, +\infty[) = \{\omega \in \Omega \mid X(\omega) \geqslant x\},$$
$$(X < x) = X^{-1}(]-\infty, x[) = \{\omega \in \Omega \mid X(\omega) < x\}, \qquad (X > x) = X^{-1}(]x, +\infty[) = \{\omega \in \Omega \mid X(\omega) > x\}.$$

Proposition 4 : Si $X : \Omega \to E$ est une variable aléatoire discrète sur un espace probabilisable $(\Omega, \mathscr{A})$ et si $U \subset E$, alors l'ensemble $(X \in U)$ est un évènement, i.e. $(X \in U)$ est un élément de $\mathscr{A}$.

Remarque 7 : On en déduit que les six ensembles ci-dessus sont des évènements de $(\Omega, \mathscr{A})$.

Exemple 6 : On lance un dé cubique rouge et un dé cubique bleu. Un univers permettant de modéliser cette expérience aléatoire est $\Omega = [\![1, 6]\!]^2$. Comme l'univers est fini, on peut considérer la tribu $\mathscr{A} = \mathscr{P}(\Omega)$.
Les applications définies sur Ω par $X : (k, \ell) \mapsto k$ et $Y : (k, \ell) \mapsto \ell$ sont des variables aléatoires sur $(\Omega, \mathscr{A})$. Elles correspondent respectivement au résultat du dé rouge et au résultat du dé bleu. On a $X(\Omega) = Y(\Omega) = [\![1, 6]\!]$.

Proposition 5 : Soit X une variable aléatoire discrète sur un espace probabilisable $(\Omega, \mathscr{A})$. Si f est une fonction définie sur $X(\Omega)$, alors l'application $f \circ X$ est une variable aléatoire discrète sur $(\Omega, \mathscr{A})$.

Notation : La variable aléatoire $f \circ X$ se note simplement $f(X)$.

Remarque 8 : Si $X : \Omega \to E$ et $Y : \Omega \to F$ sont deux variables aléatoires discrètes sur l'espace probabilisable $(\Omega, \mathscr{A})$, alors l'application $Z = (X, Y) : \Omega \to E \times F$ est une variable aléatoire discrète sur $(\Omega, \mathscr{A})$.

Corollaire 1 : Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{K}$ sur un espace probabilisable $(\Omega, \mathscr{A})$.

(i) Pour tout $(\lambda, \mu) \in \mathbb{K}^2$, l'application $\lambda X + \mu Y$ est une variable aléatoire réelle sur $(\Omega, \mathscr{A})$.

(ii) L'application XY est une variable aléatoire réelle sur $(\Omega, \mathscr{A})$.

II.B - Loi d'une variable aléatoire discrète

Définition (Loi d'une variable aléatoire discrète) : Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. La loi de la variable aléatoire X est l'application $P_X : \mathscr{P}(X(\Omega)) \to \mathbb{R}$ par $P_X : A \mapsto P_X(A) = P(X \in A)$.

Remarque 9 : On peut vérifier que P_X est une probabilité sur l'espace probabilisable $\big(X(\Omega), \mathscr{P}(X(\Omega))\big)$.

Proposition 6 : Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. La loi P_X est déterminée par la famille des nombres $P(X = x)$ pour $x \in X(\Omega)$.

Remarques 10 :

a) La proposition est une conséquence de la remarque suivante : comme l'ensemble $X(\Omega)$ est au plus dénombrable, il en est de même pour toute partie $A \in \mathscr{P}(X(\Omega))$, donc on peut écrire

$$P_X(A) = P(X \in A) = P\left(\bigcup_{x \in A} (X = x)\right) = \sum_{x \in A} P(X = x).$$

b) On en déduit que pour donner la loi d'une variable aléatoire discrète X, il suffit de spécifier l'ensemble $X(\Omega)$ et de calculer le nombre $P(X = x)$ pour tout $x \in X(\Omega)$.

Exemple 7 : Si X désigne le résultat du lancer d'un dé cubique équilibré, alors la loi de X est donnée par

$$X(\Omega) = [\![1, 6]\!] \quad \text{et} \quad \forall n \in X(\Omega), \quad P(X = n) = \frac{1}{6}.$$

Par définition, deux variables aléatoires X et Y définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ ont la même loi si et seulement si $X(\Omega) = Y(\Omega)$ et si

$$\forall a \in X(\Omega), \quad P(X = a) = P(Y = a).$$

Notation : Si deux variables aléatoires X et Y suivent la même loi, on note $X \sim Y$. De plus, si une variable aléatoire X suit une loi $\mathscr{L}$, on note $X \sim \mathscr{L}$.

Exemples 8 :

a) Si X suit une loi uniforme sur un ensemble fini non vide E, on note $X \sim \mathscr{U}(E)$.

b) Si X suit une loi de Bernoulli de paramètre $p \in [0, 1]$, on note $X \sim \mathscr{B}(p)$.

c) Si X suit une loi binomiale de paramètre $n \in \mathbb{N}^*$ et $p \in [0, 1]$, on note $X \sim \mathscr{B}(n, p)$.

Proposition 7 : Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. Si f est une application définie sur $X(\Omega)$ et si X et Y suivent la même loi, alors $f(X)$ et $f(Y)$ suivent la même loi.

Exercice 1 : On considère une variable aléatoire X telle que $X(\Omega) = \mathbb{N}$. On suppose qu'il existe un élément $a \in \mathbb{R}$ tel que $P(X = n) = a2^{-n}$ pour tout $n \in \mathbb{N}$.

1. Déterminer la valeur de a.
2. Calculer $P(X \geqslant n)$ pour tout $n \in \mathbb{N}$.

Définition (Loi conditionnelle) : Soit X une variable aléatoire discrète sur $(\Omega, \mathscr{A}, P)$.
Si B est un évènement de probabilité non nulle, la loi conditionnelle de X sachant B est la loi de X dans l'espace probabilisé $(\Omega, \mathscr{A}, P_B)$, i.e. l'application $P_{B,X} : \mathscr{P}(X(\Omega)) \to \mathbb{R}$ définie par $P_{B,X} : A \mapsto P_{B,X}(A) = P_B(X \in A)$.

Remarque 11 : D'après la proposition 6, on en déduit que pour donner la loi conditionnelle d'une variable aléatoire discrète X sachant un évènement B, il suffit de spécifier l'ensemble $X(\Omega)$ et de calculer le nombre $P_B(X = x)$ pour tout $x \in X(\Omega)$.

Exemple 9 : On lance un dé cubique équilibrée et on note X le résultat.

- La variable aléatoire X suit une loi uniforme sur l'ensemble $[\![1, 6]\!]$.
- On note B l'évènement « le résultat du dé est pair ». Déterminons la loi conditionnelle de X sachant l'évènement B. Pour tout $k \in \{2, 4, 6\}$, on a
$$P_B(X = k) = \frac{P\big((X = k) \cap B\big)}{P(B)} = \frac{P(X = k)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}$$
tandis que pour tout $k \in \{1, 3, 5\}$, on a
$$P_B(X = k) = \frac{P\big((X = k) \cap B\big)}{P(B)} = \frac{P(\varnothing)}{P(B)} = 0.$$

II.C - Couples de variables aléatoires discrètes

Définition (Couple de variables aléatoires discrètes) : On appelle couple de variables aléatoires discrètes sur un espace probabilisable $(\Omega, \mathscr{A})$ tout couple (X, Y) où $X : \Omega \to E$ et $Y : \Omega \to F$ sont deux variables aléatoires discrètes sur $(\Omega, \mathscr{A})$.

Remarque 12 : En reprenant les notations ci-dessus, on rappelle que l'application $Z = (X, Y) : \Omega \to E \times F$ est une variable aléatoire discrète sur $(\Omega, \mathscr{A})$.

Définition (Loi conjointe) : Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathscr{A}, P)$. La loi conjointe des variables aléatoires X et Y est la loi $P_{(X,Y)}$ du couple (X, Y).

Remarque 13 : D'après la proposition 6, on en déduit que pour donner la loi d'un couple de variables aléatoires discrète (X, Y), il suffit de spécifier les ensembles $X(\Omega)$ et $Y(\Omega)$, puis de calculer le nombre
$$P\big((X, Y) = (x, y)\big) = P\big((X = x) \cap (Y = y)\big)$$
pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$.

Notation : Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on note $P(X = x, Y = y) = P\big((X = x) \cap (Y = x)\big)$.

Définition (Lois marginales) : Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathscr{A}, P)$.
La loi de X et la loi de Y sont appelées les lois marginales du couple (X, Y).

Remarque 14 : La loi conjointe de deux variables aléatoires discrètes détermine entièrement les lois marginales. En effet, par la formule des probabilités totales, on a les relations
$$\forall x \in X(\Omega), \quad P(X = x) = \sum_{y \in Y(\Omega)} P\big((X = x) \cap (Y = y)\big),$$
$$\forall y \in Y(\Omega), \quad P(Y = y) = \sum_{x \in X(\Omega)} P\big((X = x) \cap (Y = y)\big).$$

Exemple 10 : On considère le couple de variables aléatoires (X, Y) dont la loi est donnée par $X(\Omega) = Y(\Omega) = \mathbb{N}$ et

$$\forall (i, j) \in \mathbb{N}^2, \quad P\big((X = i) \cap (Y = j)\big) = \frac{\mathrm{e}^{-1}}{i!2^{j+1}}.$$

Déterminons les lois marginales du couple (X, Y).

- On a $X(\Omega) = \mathbb{N}$. De plus, pour tout élément $i \in \mathbb{N}$, en appliquant la formule des probabilités totales avec le système complet d'évènements $\big((Y = j)\big)_{j\in\mathbb{N}}$, on a

$$P(X = i) = \sum_{j=0}^{+\infty} P\big((X = i) \cap (Y = j)\big) = \frac{\mathrm{e}^{-1}}{i!} \sum_{j=0}^{+\infty} \frac{1}{2^{j+1}} = \frac{\mathrm{e}^{-1}}{i!} \times \frac{1/2}{1-1/2} = \frac{\mathrm{e}^{-1}}{i!}.$$

- On a $Y(\Omega) = \mathbb{N}$. De plus, pour tout élément $j \in \mathbb{N}$, en appliquant la formule des probabilités totales avec le système complet d'évènements $\big((X = i)\big)_{i\in\mathbb{N}}$, on a

$$P(Y = j) = \sum_{i=0}^{+\infty} P\big((X = i) \cap (Y = j)\big) = \frac{\mathrm{e}^{-1}}{2^{j+1}} \sum_{i=0}^{+\infty} \frac{1}{i!} = \frac{\mathrm{e}^{-1}}{2^{j+1}} \times \mathrm{e} = \frac{1}{2^{j+1}}.$$

ATTENTION : La réciproque est fausse en général : les lois marginales ne permettent pas de déterminer la loi conjointe de deux variables aléatoires. Par exemple, les tableaux ci-dessous donnent deux lois conjointes différentes possibles pour des variables aléatoires X et Y avec des lois marginales identiques.

$P\big((X = x) \cap (Y = y)\big)$	$y = 0$	$y = 1$	$P(X = x)$
$x = 0$	$1/2$	0	$1/2$
$x = 1$	0	$1/2$	$1/2$
$P(Y = y)$	$1/2$	$1/2$	

$P\big((X = x) \cap (Y = y)\big)$	$y = 0$	$y = 1$	$P(X = x)$
$x = 0$	$1/4$	$1/4$	$1/2$
$x = 1$	$1/4$	$1/4$	$1/2$
$P(Y = y)$	$1/2$	$1/2$	

Remarque 15 : Toutes les définitions précédentes s'étendent naturellement aux n-uplets de variables aléatoires.

Exercice 2 : Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$ et $p \in]0, 1[$. On suppose qu'il existe $a \in \mathbb{R}_+$ tel que

$$\forall (i, j) \in \mathbb{N} \times \mathbb{N}, \quad P\big((X = i) \cap (Y = j)\big) = \begin{cases} \binom{j}{i} a^j p(1-p)^j & \text{si} \quad i \leqslant j \\ 0 & \text{si} \quad i > j. \end{cases}$$

1. Déterminer la valeur de a.
2. Déterminer la loi de Y.
3. Pour tout $i \in \mathbb{N}$, démontrer la relation

$$\forall x \in]-1, 1[, \quad \frac{1}{(1-x)^{i+1}} = \sum_{j=i}^{+\infty} \binom{j}{i} x^{j-i}.$$

4. En déduire la loi de X.

II.D - Indépendance de variables aléatoires discrètes

Dans cette sous-partie, on étend la notion d'indépendance de variables aléatoires vues en première année dans le cadre des univers finis.

Définition (Indépendance de deux variables aléatoires discrètes) : Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$.
Les variables aléatoires X et Y sont dites indépendantes si pour toute partie $A \subset X(\Omega)$ et toute partie $B \subset Y(\Omega)$, les évènements $(X \in A)$ et $(X \in B)$ sont indépendants, i.e.

$$P\big((X \in A) \cap (Y \in B)\big) = P(X \in A)P(Y \in B).$$

Remarque 16 : Dans le cas où X et Y sont indépendantes, la loi du couple (X, Y) est déterminée par ses lois marginales.

Exemple 11 : On lance un dé cubique rouge et un dé cubique bleu. Les variables aléatoires X et Y donnant respectivement le résultat du dé rouge et le résultat du dé bleu sont indépendantes. La probabilité que les deux dés donnent un résultat d'au moins 5 est

$$P\big((X \geqslant 5) \cap (Y \geqslant 5)\big) = P(X \geqslant 5)P(Y \geqslant 5) = \frac{2}{6} \times \frac{2}{6} = \frac{1}{9}.$$

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

Théorème de caractérisation de l'indépendance de deux variables aléatoires : Deux variables aléatoires discrètes X et Y sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ sont indépendantes si et seulement si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P\big((X = x) \cap (Y = y)\big) = P(X = x)P(Y = y).$$

Proposition 8 : Soient X et Y deux variables aléatoires discrètes indépendantes définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. Si f est une application définie sur $X(\Omega)$ et g est une application définie sur $Y(\Omega)$, alors les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Exemple 12 : On lance deux fois un dé cubique : on note X le résultat du premier lancer et Y le résultat du second lancer. Les variables aléatoires $U = X^2$ et $V = Y^3$ sont indépendantes.

On peut généraliser la notion d'indépendance à plus de deux variables aléatoires.

Définition (Indépendance de n variables aléatoires discrètes) : Soient $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ des variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$.
Les variables aléatoires $X_1, \ldots, X_n$ sont dites indépendantes si pour tout $(A_1, \ldots, A_n) \in \prod_{i=1}^{n} \mathscr{P}(X_i(\Omega))$, les évènements $(X_1 \in A_1), \ldots, (X_n \in A_n)$ sont indépendants.

Théorème de caractérisation de l'indépendance de n variables aléatoires : Soit $n \in \mathbb{N}^*$. Des variables aléatoires discrètes $X_1, \ldots, X_n$ sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ sont indépendantes si et seulement si

$$\forall (x_1, \ldots, x_n) \in \prod_{i=1}^{n} X_i(\Omega), \quad P\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = \prod_{i=1}^{n} P(X_i = x_i).$$

Lemme des coalitions : Soient $(m,n) \in \mathbb{N}^* \times \mathbb{N}^*$ avec $m < n$ et $X_1, \ldots X_n$ des variables aléatoires indépendantes définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. Si f et g sont deux fonctions définies respectivement sur les ensembles $X_1(\Omega) \times \cdots \times X_m(\Omega)$ et $X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$, alors les variables aléatoires $f(X_1, \ldots, X_m)$ et $g(X_{m+1}, \ldots, X_n)$ sont indépendantes.

Remarque 17 : Le résultat ci-dessus se généralise si l'on considère $k \in \mathbb{N}^*$ fonctions $f_1, \ldots, f_k$ à la place de f et g.

Exemple 13 : On lance sept fois un dé cubique et on note X_k le résultat du k-ième lancer pour tout $k \in [\![1, 7]\!]$. Les variables aléatoires $S = X_1 + X_2$, $P = X_3 X_4 X_5$ et $D = \cos(X_6) - \sin(X_7)$ sont indépendantes.

Définition (Suite de variables aléatoires indépendantes / i.i.d.) : Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$.

(i) On dit que $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires indépendantes si les variables aléatoires $X_0, \ldots, X_N$ sont indépendantes pour tout $N \in \mathbb{N}$.

(ii) On dit que $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires indépendantes et identiquement distribuées si les variables aléatoires $X_0, \ldots, X_N$ sont indépendantes et de même loi pour tout $N \in \mathbb{N}$.

Notation : Dans la suite, on abrège « indépendantes et identiquement distribuées » par « i.i.d. ».

Exemple 14 : Pour modéliser une succession infinie de lancers d'une pièce, on considère une suite i.i.d. de variables aléatoires de Bernoulli.

Pour terminer, rappelons le résultat suivant de première année.

Théorème 1 : Soit $n \in \mathbb{N}^*$. Si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes définies sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ et suivant la loi de Bernoulli de paramètre $p \in [0, 1]$, alors $X_1 + \cdots + X_n$ suit la loi binomiale de paramètre (n, p).

Exercice 3 : Les variables aléatoires X et Y de l'exercice 2 sont-elles indépendantes ?

Partie III Moments d'une variable aléatoire discrète réelle

Dans cette partie, on considère un espace probabilisé $(\Omega, \mathscr{A}, P)$.

III.A - Espérance d'une variable aléatoire discrète réelle ou complexe

L'objectif principal de cette sous-partie est d'étendre la notion d'espérance d'une variable aléatoire vue en première dans le cadre des univers finis.

Définition (Espérance d'une variable aléatoire discrète) : Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.

(i) Si X est à valeurs dans $[0,+\infty]$, l'espérance de X est l'élément de $[0,+\infty]$ définie par

$$\mathrm{E}(X) = \sum_{x\in X(\Omega)} xP(X=x).$$

(ii) Si X est une variable aléatoire à valeurs dans $\mathbb{K}$, on dit que X est d'espérance finie si $\mathrm{E}(|X|) < +\infty$. Dans ce cas, l'espérance de X est l'élément de $\mathbb{K}$ définie par

$$\mathrm{E}(X) = \sum_{x\in X(\Omega)} xP(X=x).$$

Remarques 18 :

a) Lorsque $x = +\infty$, on rappelle que $xP(X=x) = 0$ si et seulement si $P(X=+\infty) = 0$.

b) Si $X(\Omega)$ est fini, alors la variable aléatoire X est d'espérance finie. Dans ce cas, on retrouve la définition de l'espérance vue en première année dans le cadre d'un univers fini.

c) Si $X(\Omega)$ est dénombrable, alors on peut écrire $X(\Omega) = \{x_n \in \mathbb{R} \mid n \in \mathbb{N}\}$ où $(x_n)_{n\in\mathbb{N}}$ est une suite injective. Dans ce cas, la variable aléatoire X est d'espérance finie si et seulement si la série $\sum\limits_{n\geqslant 0} x_n P(X=x_n)$ converge absolument. Dans ce cas, on a
$$\mathrm{E}(X) = \sum_{n=0}^{+\infty} x_n P(X=x_n).$$

d) Si X n'est pas à valeurs dans $[0,+\infty]$ et que $\mathrm{E}(|X|) = +\infty$, alors l'espérance de X n'est pas définie.

e) L'espérance correspond à la moyenne théorique des valeurs prises par X.

f) On dit qu'une variable aléatoire réelle X est centrée si elle est d'espérance finie et si $\mathrm{E}(X) = 0$.

Exemples 15 :

a) Si X désigne le résultat du lancer d'un dé cubique équilibré, alors la loi de X est donnée par

$$X(\Omega) = [\![1,6]\!] \quad \text{et} \quad \forall n \in X(\Omega), \quad P(X=n) = \frac{1}{6}.$$

Comme $X(\Omega)$ est fini, la variable aléatoire X est d'espérance finie et on a

$$\mathrm{E}(X) = \sum_{n=1}^{6} nP(X=n) = \sum_{n=1}^{6} \frac{n}{6} = \frac{21}{6}.$$

b) Si X est une variable aléatoire dont la loi est donnée par

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall n \in \mathbb{N}^*, \quad P(X=n) = \frac{6}{\pi^2 n^2},$$

alors X n'est pas d'espérance finie, car la série ci-dessous n'est pas absolument convergente.

$$\sum_{n\geqslant 1} nP(X=n) = \sum_{n\geqslant 1} \frac{6}{\pi^2 n}.$$

c) Si X est une variable aléatoire dont la loi est donnée par

$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall n \in \mathbb{N}, \quad P(X = n) = \frac{1}{2^{n+1}},$$

alors X est d'espérance finie, car la série

$$\sum_{n \geqslant 0} nP(X = n) = \sum_{n \geqslant 0} \frac{n}{2^{n+1}}$$

converge absolument par la règle de d'Alembert. De plus, on peut vérifier que

$$\mathrm{E}(X) = \sum_{n=0}^{+\infty} nP(X = n) = \sum_{n=0}^{+\infty} \frac{n}{2^{n+1}} = 1.$$

Exercice 4 : Soit $a \in \mathbb{R}$. Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ dont la loi est donnée par

$$\forall n \in \mathbb{N}^*, \quad P(X = n) = \frac{a}{n(n+1)}.$$

1. En utilisant une décomposition en éléments simples, déterminer a.
2. La variable aléatoire X est-elle d'espérance finie ? Si oui, la calculer.

Exercice 5 : Soit $a \in \mathbb{R}$. Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ dont la loi est donnée par

$$\forall n \in \mathbb{N}^*, \quad P(X = n) = \frac{a}{n(n+1)(n+2)}.$$

1. En utilisant une décomposition en éléments simples, déterminer a.
2. La variable aléatoire X est-elle d'espérance finie ? Si oui, la calculer.

Proposition 9 : Si X est une variable aléatoire sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N} \cup \{+\infty\}$, alors on a

$$\mathrm{E}(X) = \sum_{n=1}^{+\infty} P(X \geqslant n).$$

Exemple 16 : En reprenant l'exemple 15.c, on a avec la somme de la série géométrique que

$$\forall n \in \mathbb{N}, \quad P(X \geqslant n) = \sum_{k=n}^{+\infty} P(X = k) = \sum_{k=n}^{+\infty} \frac{1}{2^{k+1}} = \frac{1}{2^{n+1}} \frac{1}{1 - 2^{-1}} = \frac{1}{2^n}.$$

Comme on avait déjà justifié que X est d'espérance finie, on obtient avec la proposition précédente que

$$\mathrm{E}(X) = \sum_{n=1}^{+\infty} P(X \geqslant n) = \sum_{n=1}^{+\infty} \frac{1}{2^n} = \frac{1}{2} \frac{1}{1 - 2^{-1}} = 1.$$

Théorème du transfert : Soient X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$ et f une fonction définie sur $X(\Omega)$ à valeurs dans $\mathbb{K}$. La variable aléatoire $f(X)$ est d'espérance finie si et seulement si la famille $\big(f(x)P(X = x)\big)_{x \in X(\Omega)}$ est sommable. Dans ce cas, on a

$$\mathrm{E}\big(f(X)\big) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$

Remarques 19 :

a) Si $X(\Omega)$ est fini, alors la variable aléatoire $f(X)$ est d'espérance finie et la formule du théorème s'applique. Dans ce cas, on retrouve le théorème du transfert vu en première année dans le cadre d'un univers fini.

b) Si $X(\Omega)$ est dénombrable, alors on peut écrire $X(\Omega) = \{x_n \in \mathbb{R} \mid n \in \mathbb{N}\}$ où $(x_n)_{n\in\mathbb{N}}$ est une suite injective. Dans ce cas, la variable aléatoire $f(X)$ est d'espérance finie si et seulement si la série $\sum\limits_{n\geqslant 0} f(x_n)P(X = x_n)$ converge absolument. Dans ce cas, on a

$$\mathrm{E}\big(f(X)\big) = \sum_{n=0}^{+\infty} f(x_n)P(X = x_n).$$

c) En particulier, le théorème s'applique également au cas où X est un n-uplet de variables aléatoires. Par exemple, en appliquant le théorème avec la fonction $f : (x, y) \mapsto xy$ et avec deux variables aléatoires discrètes X et Y à valeurs dans $\mathbb{K}$, on obtient que la variable aléatoire XY est d'espérance finie si et seulement si la famille $\big(xyP((X,Y) = (x, y))\big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable. Dans ce cas, on a

$$\mathrm{E}\big(XY\big) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP((X,Y) = (x, y)).$$

Exemples 17 :

a) En reprenant l'exemple 15.a, comme $X(\Omega)$ est fini, la variable aléatoire X^2 est d'espérance finie par le théorème du transfert et on a

$$E\big(X^2\big) = \sum_{n=1}^{6} n^2 P(X = n) = \sum_{n=1}^{6} \frac{n^2}{6} = \frac{91}{6}.$$

b) En reprenant l'exemple 15.c, on trouve que X^2 est d'espérance finie par le théorème du transfert, car la série

$$\sum_{n\geqslant 0} n^2 P(X = n) = \sum_{n\geqslant 0} \frac{n^2}{2^{n+1}}$$

converge absolument par la règle de d'Alembert. De plus, on peut vérifier que

$$\mathrm{E}\big(X^2\big) = \sum_{n=0}^{+\infty} n^2 P(X = n) = \sum_{n=0}^{+\infty} \frac{n^2}{2^{n+1}} = 3.$$

Proposition (Linéarité de l'espérance) : Soient X et Y deux variables aléatoires d'espérance finie sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{K}$. Pour tout $(\lambda,\mu) \in \mathbb{K}^2$, la variable aléatoire $\lambda X + \mu Y$ est d'espérance finie et on a

$$\mathrm{E}(\lambda X + \mu Y) = \lambda\,\mathrm{E}(X) + \mu\,\mathrm{E}(Y).$$

Proposition (Positivité et croissance de l'espérance) : Soient X et Y deux variables aléatoires réelle d'espérance finie sur $(\Omega, \mathscr{A}, P)$.

(i) Positivité de l'espérance : si X est positive, alors $\mathrm{E}(X) \geqslant 0$.

(ii) Croissance de l'espérance : si $X \leqslant Y$, alors $\mathrm{E}(X) \leqslant \mathrm{E}(Y)$.

(iii) Si X est positive et d'espérance nulle, alors $(X = 0)$ est presque sûr.

Proposition 10 : Soient X et Y deux variables aléatoires sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{K}$. Si $|X| \leqslant |Y|$ et si Y est d'espérance finie, alors X est d'espérance finie.

Proposition 11 : Soient X et Y deux variables aléatoires d'espérance finie sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{K}$.
Si X et Y sont indépendantes, alors la variable aléatoire XY est d'espérance finie et on a $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$.

III.B - Variance et écart type d'une variable aléatoire réelle

Proposition (Inégalité de Cauchy-Schwarz) : Soient X et Y deux variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. Si X^2 et Y^2 sont d'espérance finie, alors XY est d'espérance finie et on a

$$\mathrm{E}(XY)^2 \leqslant \mathrm{E}\left(X^2\right)\mathrm{E}\left(Y^2\right).$$

De plus, on a égalité si et seulement si l'évènement « X et Y sont colinéaires » est presque sûr.

Remarques 20 :

a) En particulier, en prenant $Y = 1$, on en déduit que si X^2 est d'espérance finie, alors X est d'espérance finie.

b) On en déduit que si X^2 est d'espérance finie, alors $\left(X - \mathrm{E}(X)\right)^2$ est aussi d'espérance finie.

Définition (Variance et écart type d'une variable aléatoire réelle) : Soit X une variable aléatoire réelle sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.

(i) Si X^2 est d'espérance finie, on appelle variance de X le nombre réel $\mathrm{V}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right)$.

(ii) Si X^2 est d'espérance finie, on appelle écart type de X le nombre réel $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarques 21 :

a) Si X^2 est d'espérance finie, on dit que X est de variance finie ou que X admet une variance.

b) Si $X(\Omega)$ est fini, alors X^2 est d'espérance finie, donc X est de variance finie.

c) L'écart type est aussi appelée l'écart quadratique moyen à la moyenne. Il permet d'évaluer la dispersion de la variable aléatoire X autour de son espérance $\mathrm{E}(X)$.

d) Une variable aléatoire réelle de variance finie est presque sûrement constante si et seulement si sa variance est nulle.

e) On dit qu'une variable aléatoire réelle X est réduite si elle est de variance finie et si $\mathrm{V}(X) = 1$.

Théorème de Koenig-Huyghens : Soit X une variable aléatoire réelle sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.
Si X est de variance finie, alors on a

$$\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2.$$

Exemple 18 : En reprenant l'exemple 15.c, on obtient que $\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 = 3 - 1^2 = 2$.

Proposition 12 : Soit X une variable aléatoire réelle discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.
Si X est de variance finie, alors pour tout $(a, b) \in \mathbb{R}^2$, la variable aléatoire $aX + b$ est de variance finie et on a

$$\mathrm{V}(aX + b) = a^2\,\mathrm{V}(X).$$

Remarque 22 : Si X est une variable aléatoire réelle de variance finie tel que $\sigma(X) > 0$, alors $\dfrac{X - \mathrm{E}(X)}{\sigma(X)}$ est une variable aléatoire centrée et réduite.

ATTENTION : Si X et Y sont des variables aléatoires réelles, on n'a pas $\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y)$ en général.

III.C - Covariance de deux variables aléatoires réelles

Définition (Covariance de deux variables aléatoires réelles) : Soient X et Y deux variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. Si X et Y sont de variance finie, on définit la covariance de X et Y comme le nombre réel

$$\mathrm{Cov}(X,Y) = \mathrm{E}\big((X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\big).$$

Remarques 23 :

a) La covariance existe d'après l'inégalité de Cauchy-Schwarz.

b) La covariance est une application bilinéaire, symétrique et positive.

c) Deux variables aléatoires dont la covariance est nulle sont dites décorrélées.

Remarque 24 : On peut interpréter la covariance de la façon suivante.

a) Si $\mathrm{Cov}(X,Y) > 0$, alors Y a tendance à augmenter lorsque X augmente.

b) Si $\mathrm{Cov}(X,Y) < 0$, alors Y a tendance à diminuer lorsque X augmente.

c) Si $\mathrm{Cov}(X,Y) = 0$, alors il n'y a pas de lien entre les variations des variables aléatoires X et Y.

Théorème 2 : Soient X et Y deux variables aléatoires réelles de variance finie sur $(\Omega, \mathscr{A}, P)$.

(i) On a $\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$.

(ii) Si X et Y sont indépendantes, alors on a $\mathrm{Cov}(X,Y) = 0$.

ATTENTION : La réciproque est fausse en général. Par exemple, si X et Y sont deux variables aléatoires indépendantes dont les lois sont données par

$$X(\Omega) = \{0,1\}, \quad Y(\Omega) = \{-1,1\} \quad \text{et} \quad P(X=0) = P(X=1) = P(Y=1) = P(Y=-1) = \frac{1}{2},$$

alors les variables aléatoires X et $Z = XY$ ne sont pas indépendantes car

$$P\big((X=1)\cap(Z=1)\big) \neq P(X=1)P(Z=1).$$

Cependant, on a $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0$.

Proposition 13 : Soient X et Y deux variables aléatoires réelles de variance finie sur $(\Omega, \mathscr{A}, P)$.

(i) Pour tout $(a,b) \in \mathbb{R}^2$, la variable aléatoire $aX + bY$ est de variance finie et on a

$$\mathrm{V}(aX+bY) = a^2\,\mathrm{V}(X) + 2ab\,\mathrm{Cov}(X,Y) + b^2\,\mathrm{V}(Y).$$

(ii) Si X et Y sont indépendantes, alors $X+Y$ est de variance finie et on a $\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y)$.

Proposition 14 : Soient $n \in \mathbb{N}^*$ et $X_1, \ldots, X_n$ des variables aléatoires réelles de variance finie sur $(\Omega, \mathscr{A}, P)$.

(i) La variable aléatoire $X_1 + \cdots + X_n$ est de variance finie et on a

$$\mathrm{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathrm{V}(X_k) + 2\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j).$$

(ii) Si $X_1, \ldots, X_n$ sont deux à deux indépendants, alors on a

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

Exercice 6 : Soient $n \in \mathbb{N}^*$ et $N \in \mathbb{N}^*$ avec $N \geqslant 2$. À un péage autoroutier n voitures franchissent équiprobablement et indépendamment l'une des N barrières de péage mises à leur disposition. Pour chaque $k \in [\![1, N]\!]$, on note X_k le nombre de voitures passant par la k-ième barrière.

1. Déterminer la loi, l'espérance et la variance de X_k pour tout $k \in [\![1, N]\!]$.
2. Calculer la variance de la variable aléatoire $X_1 + \cdots + X_N$.
3. Soit $(i, j) \in [\![1, N]\!]^2$ avec $i \neq j$. Calculer la covariance $\mathrm{Cov}(X_i, X_j)$.

III.D - Fonctions génératrices d'une variable aléatoire à valeurs entières

Dans cette partie, on considère un espace probabilisé $(\Omega, \mathscr{A}, P)$.

Définition (Fonction génératrice) : Soit X une variable aléatoire définie sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$.
La fonction génératrice de la variable aléatoire X est la fonction G_X d'une variable réelle définie par

$$G_X : t \mapsto \mathrm{E}\left(t^X\right) = \sum_{n=0}^{+\infty} P(X = n)\, t^n.$$

Remarques 25 :

a) L'égalité dans la définition provient du théorème du transfert.

b) Par définition, la fonction génératrice G_X est la somme d'une série entière. En particulier, si on désigne par R_X le rayon de convergence de cette dernière, alors l'ensemble de définition de G_X est de la forme

$$]-R_X, R_X[, \quad [-R_X, R_X[, \quad]-R_X, R_X] \quad \text{ou} \quad [-R_X, R_X].$$

c) Comme la série définissant le nombre $G_X(1) = 1$ converge absolument, on en déduit que $R_X \geqslant 1$.

d) Comme pour tout $n \in \mathbb{N}$ et tout $t \in [-1, 1]$, on a $|P(X = n)\, t^n| \leqslant P(X = n)$ et que la série $\sum_{n \geqslant 0} P(X = n)$ converge, on obtient que la série entière définissant G_X converge normalement sur $[-1, 1]$. On en déduit par le théorème de continuité pour les séries de fonctions que la fonction G_X est continue sur $[-1, 1]$.

Exemples 19 :

a) Si X suit une loi binomiale de paramètre $(n, p) \in \mathbb{N} \times [0, 1]$, alors X prend un nombre fini de valeurs, donc la fonction génératrice G_X est définie sur $\mathbb{R}$. De plus, en utilisant le binôme de Newton, on a pour tout $t \in \mathbb{R}$ que

$$G_X(t) = \sum_{k=0}^{+\infty} P(X = k)\, t^k = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} t^k = \sum_{k=0}^{n} \binom{n}{k} (pt)^k (1-p)^{n-k} = (1 - p + pt)^n.$$

b) Si X suit une loi uniforme sur $[\![1, n]\!]$ où $n \in \mathbb{N}^*$, alors X prend un nombre fini de valeurs, donc la fonction génératrice G_X est définie sur $\mathbb{R}$. De plus, on a pour tout $t \in \mathbb{R}$ que

$$G_X(t) = \sum_{k=0}^{+\infty} P(X = k)\, t^k = \sum_{k=1}^{n} \frac{1}{n} t^k = \frac{1}{n} \sum_{k=1}^{n} t^k = \begin{cases} \dfrac{t - t^{n+1}}{n(1-t)} & \text{si} \quad t \neq 1 \\ 1 & \text{si} \quad t = 1. \end{cases}$$

Proposition 15 : Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$.
La loi de X est entièrement déterminée par sa fonction génératrice G_X sur $]-1, 1[$.

Remarque 26 : Cette propriété est une conséquence directe de la relation

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

Théorème 3 : Soit X une variable aléatoire définie sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$.
La variable aléatoire X est d'espérance finie si et seulement si la fonction G_X est dérivable en 1.
Dans ce cas, on a $\mathrm{E}(X) = G_X'(1)$.

Remarque 27 : Lorsque $R_X > 1$, on sait par les propriétés sur les séries entières que

$$G_X'(1) = \sum_{n=0}^{+\infty} nP(X=n) \quad \text{et} \quad G_X''(1) = \sum_{n=0}^{+\infty} n(n-1)P(X=n).$$

En particulier, on en déduit avec le théorème du transfert que X est de variance finie et que

$$G_X'(1) = \mathrm{E}(X) \quad \text{et} \quad G_X''(1) = \mathrm{E}\big(X(X-1)\big) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X).$$

On en déduit une expression de la variance en fonction de G_X avec

$$\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 = G_X''(1) + G_X'(1) - G_X'(1)^2.$$

Exemple 20 : Si X suit une loi binomiale de paramètre $(n,p) \in \mathbb{N} \times [0,1]$, alors sa fonction génératrice G_X est polynomiale, donc elle est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$. De plus, on a

$$\forall t \in \mathbb{R}, \quad G_X'(t) = np(1-p+pt)^{n-1} \quad \text{et} \quad G_X''(t) = n(n-1)p^2(1-p+pt)^{n-2}.$$

D'après le théorème précédent, la variable aléatoire X est d'espérance finie et de variance finie. De plus, on a

$$\mathrm{E}(X) = G_X'(1) = np \quad \text{et} \quad \mathrm{V}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2 = n(n-1)p^2 + np - (np)^2 = np(1-p).$$

Théorème 4 : Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$.
Si les variables aléatoires X et Y sont indépendantes, alors pour tout $t \in]-1,1[$, on a $G_{X+Y}(t) = G_X(t)G_Y(t)$.

Exemple 21 : Soient X et Y deux variables aléatoires indépendantes suivant des lois binomiales de paramètres respectifs (n_1, p) et (n_2, p) avec $(n_1, n_2, p) \in \mathbb{N}^2 \times [0,1]$. Déterminons la loi de $Z = X + Y$.
Comme les variables X et Y sont indépendantes, on a d'après le théorème ci-dessus que

$$\forall t \in]-1,1[, \quad G_Z(t) = G_X(t)G_Y(t) = (1-p+pt)^{n_1}(1-p+pt)^{n_2} = (1-p+pt)^{n_1+n_2}.$$

On reconnait la fonction génératrice de la loi binomiale de paramètre (n_1+n_2, p), donc la variable aléatoire Z suit la loi binomiale de paramètre (n_1+n_2, p).

Remarque 28 : Le théorème se généralise directement à plusieurs variables aléatoires. Si $X_1, \dots, X_n$ sont des variables aléatoires indépendantes définies sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$, alors on a

$$\forall t \in]-1,1[, \quad G_{X_1+\cdots+X_n}(t) = G_{X_1}(t)\cdots G_{X_n}(t).$$

Exercice 7 : Une urne contient 4 boules rapportant respectivement 0, 1, 1 et 2 points. On y effectue $n \in \mathbb{N}^*$ tirages avec remise et l'on note S le score total obtenu. On désigne par X_k le nombre de points obtenus au tirage de rang $k \in [\![1,n]\!]$.

1. Déterminer la fonction génératrice de X_k pour tout $k \in [\![1,n]\!]$.
2. Déterminer la fonction génératrice de S.
3. En déduire la loi de S.

Partie IV Lois usuelles

Dans cette partie, on considère un espace probabilisé $(\Omega, \mathscr{A}, P)$.

IV.A - Loi géométrique

Définition (Loi géométrique) : Soit $p \in]0,1[$. On dit qu'une variable aléatoire X définie sur $(\Omega, \mathscr{A}, P)$ suit la loi géométrique de paramètre p si

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall k \in \mathbb{N}^*, \quad P(X = k) = p(1-p)^{k-1}.$$

Notation : Si une variable aléatoire X suit une loi géométrique de paramètre $p \in]0,1[$, on note $X \sim \mathscr{G}(p)$.

Remarques 29 :

a) La loi géométrique apparait naturellement lorsque l'on considère une succession d'épreuves de Bernoulli indépendantes de paramètre p jusqu'à obtenir un premier succès.

Si $(\Omega, \mathscr{A}, P)$ est un espace probabilisé contenant une suite $(A_n)_{n \in \mathbb{N}^*}$ d'évènements indépendants de même probabilité $p \in]0,1[$ et si on désigne par X le rang du premier évènement réalisé, alors on a par indépendance que

$$\forall k \in \mathbb{N}^*, \quad P(X = k) = P\left(\overline{A_1} \cap \cdots \cap \overline{A_{k-1}} \cap A_k\right) = P\left(\overline{A_1}\right) \cdots P\left(\overline{A_{k-1}}\right) P(A_k) = p(1-p)^{k-1}.$$

On en déduit que X suit une loi géométrique de paramètre p.

b) Remarquons que formellement nous n'avons pas défini X lorsque aucun des évènements A_n pour $n \in \mathbb{N}^*$ n'est réalisé. Cependant, il s'agit d'un détail sans importance, car par la propriété de continuité monotone, la probabilité de cet évènement est

$$P\left(\bigcap_{k \in \mathbb{N}^*} \overline{A_k}\right) = \lim_{n \to +\infty} P\left(\bigcap_{k=1}^{n} \overline{A_k}\right) = \lim_{n \to +\infty} (1-p)^n = 0.$$

Exemple 22 : On lance un dé équilibré autant de fois que nécessaire jusqu'à obtenir un 6. Si on note X le nombre de lancers effectués, alors X suit une loi géométrique de paramètre $1/6$.

Proposition 16 : Soit X une variable aléatoire suivant la loi géométrique de paramètre $p \in]0,1[$.

(i) La fonction génératrice de X est définie sur $\left]-\frac{1}{1-p}, \frac{1}{1-p}\right[$ et on a

$$\forall t \in \left]-\frac{1}{1-p}, \frac{1}{1-p}\right[, \quad G_X(t) = \frac{pt}{1-(1-p)t}.$$

(ii) La variable aléatoire X est d'espérance finie et de variance finie. De plus, on a

$$\mathrm{E}(X) = \frac{1}{p} \quad \text{et} \quad \mathrm{V}(X) = \frac{1-p}{p^2}.$$

Théorème de caractérisation de la loi géométrique (⋆) : Soit X une variable aléatoire définie sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}^*$. La variable aléatoire X suit une loi géométrique si et seulement si

$$\forall (k,n) \in \mathbb{N}^2, \quad P(X > n + k \mid X > n) = P(X > k).$$

Remarque 30 : On dit que la loi géométrique est sans mémoire.

Exercice 8 : Soit X une variable aléatoire réelle suivant une loi géométrique de paramètre $p \in]0,1[$.

1. Calculer les probabilités $P(X \leqslant n)$ et $P(X \geqslant n)$ pour tout $n \in \mathbb{N}^*$.
2. Calculer la probabilité que X soit pair.

Exercice 9 : Soient X et Y deux variables aléatoires indépendantes suivant des lois géométriques de paramètre respectif $p \in]0,1[$ et $q \in]0,1[$. Déterminer la loi de la variable aléatoire $S = X + Y$.

IV.B - Loi de Poisson

Définition (Loi de Poisson) : Soit $\lambda \in \mathbb{R}_+^*$. On dit qu'une variable aléatoire X définie sur $(\Omega, \mathscr{A}, P)$ suit la loi poisson de paramètre λ si

$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall k \in \mathbb{N}, \quad P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Notation : Si une variable aléatoire X suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$, on note $X \sim \mathscr{P}(\lambda)$.

Proposition 17 : Soit X une variable aléatoire suivant la loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$.

(i) La fonction génératrice de X est définie sur $\mathbb{R}$ et on a

$$\forall t \in \mathbb{R}, \quad G_X(t) = \exp\big(\lambda(t-1)\big).$$

(ii) La variable aléatoire X est d'espérance finie et de variance finie. De plus, on a

$$\mathrm{E}(X) = \lambda \quad \text{et} \quad \mathrm{V}(X) = \lambda.$$

Exercice 10 : Soit X une variable aléatoire réelle suivant une loi de poisson de paramètre $\lambda > 0$.

1. Calculer la probabilité que X soit pair.
2. Montrer que pour tout $n \in \mathbb{N}$, on a
$$P(X \leqslant n) = \frac{1}{n!} \int_\lambda^{+\infty} e^{-t} t^n \, dt.$$

Exercice 11 : Soit $n \in \mathbb{N}^*$. Montrer que si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes suivants des lois de Poisson de paramètre respectif $\lambda_1, \ldots, \lambda_n$, alors $S = X_1 + \cdots + X_n$ est une variable aléatoire suivant une loi de Poisson de paramètre $\lambda_1 + \cdots + \lambda_n$.

Partie V Inégalités probabilistes

Dans cette partie, on considère un espace probabilisé $(\Omega, \mathscr{A}, P)$.

Inégalité de Markov : Soit X une variable aléatoire complexe sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.
Si X est d'espérance finie, alors on a l'inégalité

$$\forall \varepsilon > 0, \quad P\big(|X| \geqslant \varepsilon\big) \leqslant \frac{\mathrm{E}\big(|X|\big)}{\varepsilon}.$$

Inégalité de Bienaymé-Tchebychev : Soit X une variable aléatoire réelle sur un l'espace probabilisé $(\Omega, \mathscr{A}, P)$.
Si X est de variance finie, alors on a l'inégalité

$$\forall \varepsilon > 0, \quad P\big(|X - \mathrm{E}(X)| \geqslant \varepsilon\big) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}.$$

Remarque 31 : On retrouve que la variance est un indicateur de dispersion : plus la variance de X est faible, plus les valeurs de X sont concentrées avec une probabilité élevée autour de son espérance $\mathrm{E}(X)$.

Pour terminer, nous allons utiliser l'inégalité précédente pour faire le lien entre l'approche des probabilités introduite au collège et le modèle plus formel que nous avons développé cette année.
Dans la suite, on considère l'expérience aléatoire du lancer d'une pièce équilibrée. La définition de la probabilité d'un évènement vue au collège est la suivante.

« Si on répète une expérience aléatoire un très grand nombre de fois, la fréquence de n'importe quel événement de cette expérience finit par se stabiliser autour d'un nombre qui est la probabilité de cet événement. »

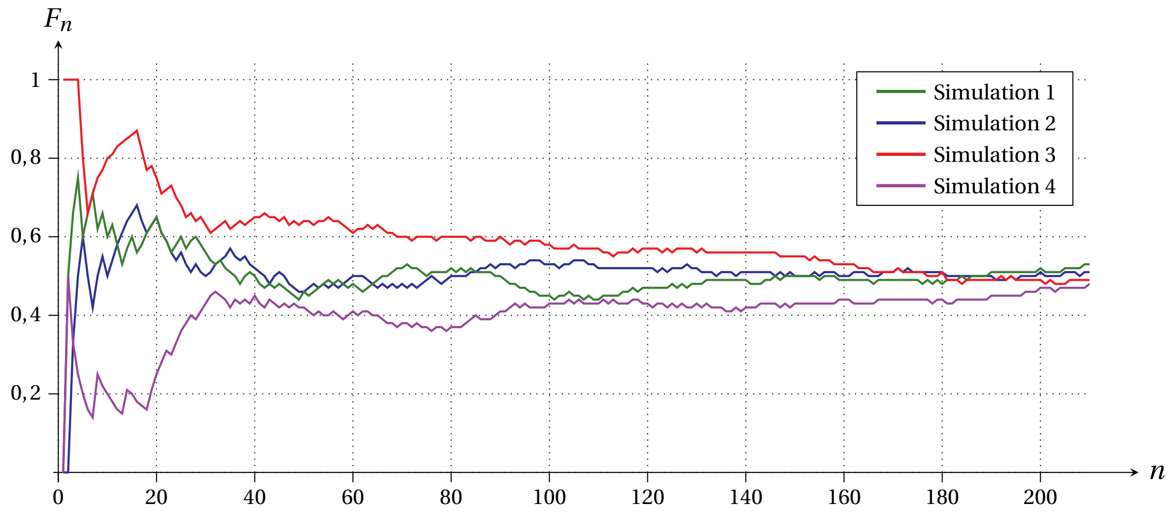
Reformulons cette définition avec notre modèle : on considère une suite de variable aléatoire $(X_n)_{n\in\mathbb{N}^*}$ définie par

$$\forall n \in \mathbb{N}^*, \quad X_n = \begin{cases} 1 & \text{si le } n\text{-ième lancer donne face.} \\ 0 & \text{si le } n\text{-ième lancer donne pile.} \end{cases}$$

Pour tout $n \in \mathbb{N}^*$, la fréquence du nombre de faces durant les n premiers lancers est

$$F_n = \frac{X_1 + \cdots + X_n}{n} = \frac{S_n}{n} \quad \text{avec} \quad S_n = X_1 + \cdots + X_n.$$

D'après l'intuition, lorsque $n \to +\infty$ la variable aléatoire F_n doit se stabiliser autour de la probabilité de l'évènement « faire face », c'est à dire $1/2$. On peut observer cette affirmation sur le graphique ci-dessus où l'on a tracé la suite $n \mapsto F_n$ pour quatre simulations différentes des variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$.



Plus généralement, si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires réelles i.i.d. sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, on peut définir

$$\forall n \in \mathbb{N}^*, \quad M_n = \frac{X_1 + \cdots + X_n}{n} = \frac{S_n}{n} \quad \text{avec} \quad S_n = X_1 + \cdots + X_n.$$

On peut interpréter M_n comme la moyenne de n réalisations d'une unique variable aléatoire X. Si l'on suppose que X est d'espérance finie, on s'attend intuitivement à ce que M_n soit proche de $\mathrm{E}(X)$ lorsque $n \to +\infty$.

Le théorème suivant formalise les observations et intuitions précédentes.

Loi faible des grands nombres : Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles i.i.d de variance finie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. Dans ce cas, en notant $\mu = \mathrm{E}(X_1)$ et $S_n = X_1 + \cdots + X_n$ pour tout $n \in \mathbb{N}^*$, on a

$$\forall \varepsilon > 0, \quad \lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right) = 0.$$

Remarques 32 :

a) La démonstration est une conséquence directe de l'inégalité de Bienaymé-Tchebytchev qui implique que

$$\forall \varepsilon > 0, \quad \forall n \in \mathbb{N}^*, \quad P\left(\left|\frac{S_n}{n} - \mu\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

b) Soit $n \in \mathbb{N}^*$. Supposons que les variables aléatoires $X_1, \ldots, X_n$ suivent une loi de Bernoulli d'un paramètre inconnu p. La relation précédente peut se réécrire sous la forme

$$P\left(\frac{S_n}{n} - \varepsilon < p < \frac{S_n}{n} + \varepsilon\right) \geqslant 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

Cette inégalité permet de construire un intervalle de confiance pour le nombre p : pour chaque réalisation des variables aléatoires $X_1, \ldots, X_n$, on obtient une estimation de p.

Signalons néanmoins que cette intervalle est de mauvaise qualité et ce n'est pas celui-ci qui est utilisé en général par les instituts de sondage.

c) La loi forte des grands nombres (hors-programme) est un théorème avec une conséquence plus forte : la suite de terme général $\dfrac{S_n}{n}$ converge presque surement vers le nombre m, i.e.

$$P\left(\left\{\omega \in \Omega \;\middle|\; \lim_{n\to+\infty} \frac{S_n(\omega)}{n} = \mu\right\}\right) = 1.$$

Partie VI Synthèse des lois usuelles

Nom	**Paramètres**	**Notation**	**$X(\Omega)$**	**Loi**	**Espérance**	**Variance**	**Fonction génératrice**
Uniforme	$n \in \mathbb{N}^*$	$\mathscr{U}(\llbracket 1, n \rrbracket)$	$\llbracket 1, n \rrbracket$	$P(X=k) = \dfrac{1}{n}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$	$\dfrac{t-t^{n+1}}{n(1-t)}$
Bernoulli	$p \in [0,1]$	$\mathscr{B}(p)$	$\{0,1\}$	$P(X=0) = 1-p$ $P(X=1) = p$	p	$p(1-p)$	$1-p+pt$
Binomiale	$n \in \mathbb{N}^*$ $p \in [0,1]$	$\mathscr{B}(n,p)$	$\llbracket 0, n \rrbracket$	$P(X=k) = \dbinom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	$(1-p+pt)^n$
Géométrique	$p \in]0,1[$	$\mathscr{G}(p)$	$\mathbb{N}^*$	$P(X=k) = p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$\dfrac{pt}{1-(1-p)t}$
Poisson	$\lambda \in \mathbb{R}_+^*$	$\mathscr{P}(\lambda)$	$\mathbb{N}$	$P(X=k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$	λ	λ	$\exp(\lambda(t-1))$