

Chapitre 14. Variables aléatoires discrètes

Plan du chapitre

1 Variables aléatoires discrètes	page 2
1.1 Définition d'une variable aléatoire discrète	page 2
1.2 Loi de probabilité d'une variable aléatoire	page 2
1.3 Fonctions de variables aléatoires	page 4
2 Couples de variables aléatoires discrètes. n-uplets de variables aléatoires discrètes	page 4
2.1 Couples de variables aléatoires discrètes	page 4
2.1.1 Définition d'un couple de variables aléatoires discrètes	page 4
2.1.2 Loi conjointe d'un couple. Lois marginales	page 5
2.1.3 Lois conditionnelles	page 8
2.2 Généralisation à un n-uplet de variables aléatoires discrètes	page 8
3 Variables aléatoires indépendantes	page 9
3.1 Cas de deux variables	page 9
3.1.1 Définition de l'indépendance de deux variables aléatoires	page 9
3.1.2 Fonctions de deux variables indépendantes	page 9
3.1.3 Somme de deux variables indépendantes à valeurs entières	page 9
3.2 Généralisation à n variables	page 10
3.3 Suites de variables indépendantes	page 12
4 Lois usuelles	page 13
4.1 Rappels de sup	page 13
4.1.1 La loi uniforme	page 13
4.1.2 La loi de BERNOULLI. La loi binomiale	page 13
4.2 Loi géométrique de paramètre p	page 13
4.3 Loi de POISSON de paramètre λ	page 14
4.3.1 Approximation de la loi binomiale par la loi de POISSON	page 14
4.3.2 Définition de la loi de POISSON	page 15
5 Espérance, variance	page 15
5.1 Espérance d'une variable aléatoire	page 15
5.1.1 Définition de l'espérance	page 15
5.1.2 Propriétés de l'espérance	page 16
5.1.3 Inégalité de MARKOV	page 19
5.2 Moments d'ordre r d'une variable aléatoire	page 19
5.3 Variance et écart-type d'une variable aléatoire	page 20
5.3.1 Définition de la variance et de l'écart-type	page 20
5.3.2 Variables centrées réduites	page 21
5.3.3 Inégalité de BIENAYMÉ-TCHEBYCHEV	page 21
5.4 Covariance d'un couple de variables aléatoires	page 22
5.4.1 Définition et propriétés de la covariance	page 22
5.4.2 Variance d'une somme finie de variables	page 22
5.5 Espérances et variances des lois usuelles	page 23
5.6 Loi faible des grands nombres	page 24
6 Fonctions génératrices	page 25
6.1 Fonction génératrice d'une variable aléatoire	page 25
6.1.1 Définition	page 25
6.1.2 Dérivées et application aux calculs des moments	page 25
6.1.3 Fonction génératrice d'une somme finie de variables indépendantes	page 27
6.2 Fonctions génératrices des variables usuelles	page 28
6.2.1 Fonction génératrice d'un variable aléatoire de BERNOULLI	page 28
6.2.2 Fonction génératrice d'un variable aléatoire suivant une loi binomiale	page 28
6.2.3 Fonction génératrice d'un variable aléatoire suivant une loi géométrique	page 28
6.2.4 Fonction génératrice d'un variable aléatoire suivant une loi de POISSON	page 29

1 Variables aléatoires discrètes

1.1 Définition d'une variable aléatoire discrète

DÉFINITION 1. Soient (Ω, \mathcal{A}) un espace probabilisable et E un ensemble non vide quelconque.

Une **variable aléatoire discrète** sur Ω à valeurs dans E est une application X de Ω dans E telle que $X(\Omega)$ soit une partie au plus dénombrable de E et telle que, pour tout x de E , $X^{-1}(\{x\}) \in \mathcal{A}$.

Si E est une partie de \mathbb{R} , la variable aléatoire X est dite **réelle**.

⇒ **Commentaire.**

◇ Dans la pratique de maths sup et maths spé, E est le plus souvent \mathbb{N} ou une partie de \mathbb{N} . C'est le cas des lois usuelles (uniforme, de BERNOULLI, géométrique, de POISSON) qui sont au programme.

◇ On rappelle la signification de $X^{-1}(\{x\})$. Si x est un élément de l'ensemble d'arrivée E , $X^{-1}(\{x\})$ est l'image réciproque du singleton $\{x\}$ par l'application X , c'est-à-dire l'ensemble des antécédents de l'élément x de E par l'application X ou encore

$$X^{-1}(\{x\}) = \{\omega \in \Omega / X(\omega) = x\}.$$

$X^{-1}(\{x\})$ est donc un élément de $\mathcal{P}(\Omega)$ et plus précisément un élément de \mathcal{A} d'après la définition. Dit autrement, $X^{-1}(\{x\})$ est un événement que l'on écrit en pratique $\{X = x\}$ ou $(X = x)$ pour simplifier la notation.

Plus généralement, si A est une partie de E , l'ensemble $X^{-1}(A) = \{\omega \in \Omega / X(\omega) \in A\}$ se note plus simplement $X \in A$. Si A est vide, $X^{-1}(A)$ est l'événement \emptyset et si A n'est pas vide,

$$X^{-1}(A) = \bigcup_{x \in A \cap X(\Omega)} \{X = x\}$$

est une réunion au plus dénombrable (car $X(\Omega)$ est au plus dénombrable) d'événements et donc $X^{-1}(A)$ est un événement.

◇ Si X est réelle et x est un élément de E , on note plus simplement les événements $X^{-1}(]x, +\infty[)$, $X^{-1}([x, +\infty[)$, $X^{-1}(]-\infty, x[)$, $X^{-1}(]-\infty, x])$: ces événements sont respectivement notés, $\{X > x\}$, $\{X \geq x\}$, $\{X < x\}$ et $\{X \leq x\}$ ou $(X > x)$, $(X \geq x)$, $(X < x)$ et $(X \leq x)$.

◇ On admettra (la démonstration étant fastidieuse) qu'une combinaison linéaire ou un produit fini de variables aléatoires discrètes est une variable aléatoire discrète.

Exemples.

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

Si a est un réel donné, l'application $X : \begin{matrix} \Omega & \rightarrow & \mathbb{R} \\ \omega & \mapsto & a \end{matrix}$ est une variable aléatoire réelle discrète sur l'espace (Ω, \mathcal{A}) . En

effet, pour $x \in \mathbb{R}$, si $x \neq a$, $X^{-1}(\{x\}) = \emptyset \in \mathcal{A}$ et si $x = a$, $X^{-1}(\{x\}) = \Omega \in \mathcal{A}$. X est une **variable constante**.

• Soit Ω un univers non vide au plus dénombrable. Soit $A \in \mathcal{P}(\Omega)$. On sait que $\mathcal{A} = \{\emptyset, A, \bar{A}, \Omega\}$ est une tribu sur Ω . La **variable indicatrice de l'événement** A est $1_A : \begin{matrix} \Omega & \rightarrow & \mathbb{R} \\ \omega & \mapsto & \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{cases} \end{matrix}$. C'est une variable aléatoire sur

l'espace (Ω, \mathcal{A}) . En effet, $X(\Omega) = \{0, 1\}$ puis $X^{-1}(\{1\}) = A \in \mathcal{A}$ et $X^{-1}(\{0\}) = \bar{A} \in \mathcal{A}$. □

1.2 Loi de probabilité d'une variable aléatoire discrète

Théorème 1. Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable et E un ensemble non vide. Soit X une variable aléatoire sur cet espace à valeurs dans E .

L'application $\mathbb{P}_X : \begin{matrix} \mathcal{P}(E) & \rightarrow & [0, 1] \\ A & \mapsto & \mathbb{P}(X \in A) \end{matrix}$ est une probabilité sur l'espace probabilisable $(E, \mathcal{P}(E))$.

DÉMONSTRATION.

• Pour chaque A de $\mathcal{P}(E)$, $\{X \in A\} = X^{-1}(A)$ est un événement, c'est-à-dire un élément de la tribu \mathcal{A} . Donc, \mathbb{P}_X est bien une application de $\mathcal{P}(E)$ dans $[0, 1]$.

• $\mathbb{P}_X(E) = \mathbb{P}(X \in E) = \mathbb{P}(\Omega) = 1$.

• Soit $(A_n)_{n \in \mathbb{N}} \in (\mathcal{P}(E))^{\mathbb{N}}$, les A_n étant deux à deux disjoints. Alors, $(X \in A_n)_{n \in \mathbb{N}}$ est une suite d'événements (éléments de \mathcal{A}) deux à deux disjoints (car X est une application). Donc, par σ -additivité de \mathbb{P} ,

$$\mathbb{P}_X \left(\bigcup_{n=0}^{+\infty} A_n \right) = \mathbb{P} \left(X \in \bigcup_{n=0}^{+\infty} A_n \right) = \mathbb{P} \left(\bigcup_{n=0}^{+\infty} (X \in A_n) \right) = \sum_{n=0}^{+\infty} \mathbb{P}(X \in A_n) = \sum_{n=0}^{+\infty} \mathbb{P}_X(A_n).$$

On a montré que \mathbb{P}_X est une probabilité sur l'espace probabilisable $(E, \mathcal{P}(E))$. □

DÉFINITION 2. L'application \mathbb{P}_X du théorème 1 s'appelle la **loi de probabilité** de la variable aléatoire X .

⇒ **Commentaire .**

◇ Dans le théorème 1, on peut remplacer E par n'importe quelle partie de E contenant $X(\Omega)$ ($X(\Omega)$ n'étant pas nécessairement E tout entier). Par exemple, on peut aussi prendre en définition de la loi de probabilité de X l'application

$$\begin{array}{ccc} \mathbb{P}_X : \mathcal{P}(X(\Omega)) & \rightarrow & [0, 1] \\ A & \mapsto & \mathbb{P}(X \in A) \end{array}$$

qui est une probabilité sur l'espace probabilisable $(X(\Omega), \mathcal{P}(X(\Omega)))$.

◇ La loi \mathbb{P}_X est entièrement déterminée par les atomes de probabilité $\mathbb{P}_X(\{x\}) = \mathbb{P}(X \in \{x\}) = \mathbb{P}(X = x)$, $x \in E$:

$$\forall A \in X(\Omega), \mathbb{P}(A) = \sum_{x \in A} \mathbb{P}(X = x) = \sum_{x \in A} \mathbb{P}_X(\{x\}).$$

Dans la pratique, pour répondre à une question du type « déterminer la loi de probabilité de la variable X », ce sont les $\mathbb{P}(X = x)$, $x \in E$ ou plus simplement encore $x \in X(\Omega)$, que l'on fournit.

Notation. Deux variables aléatoires X et Y ont la même loi de probabilité si et seulement si $X(\Omega) = Y(\Omega)$ et $\forall a \in X(\Omega)$, $\mathbb{P}(X = a) = \mathbb{P}(Y = a)$. Quand deux variables aléatoires X et Y ont la même loi de probabilité, on écrit $X \sim Y$. Quand une variable aléatoire suit une loi de probabilité \mathcal{L} donnée, on écrit $X \sim \mathcal{L}$. □

Vocabulaire. Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soit x un réel. La partie $A = \{X \leq x\}$ de Ω est un événement car réunion au plus dénombrable d'événements du type $\{X = a\}$, $a \in \mathbb{R}$. Pour $x \in \mathbb{R}$, on peut donc poser $F(x) = \mathbb{P}(X \leq x)$. F est la **fonction de répartition** de la variable aléatoire X . Cette notion n'est pas à proprement parler au programme de maths spé.

Exercice 1. On jette autant de fois que nécessaire une pièce de monnaie, chaque lancer de pièce étant indépendant des autres. Cette pièce de monnaie est truquée de sorte qu'à chaque lancer, la probabilité d'obtenir pile est $\frac{2}{3}$. On admet qu'il existe un espace probabilisé modélisant cette situation. On note X la variable aléatoire dont la valeur est le numéro du jet pour lequel on a obtenu pile pour la première fois.

Déterminer la loi de probabilité de X .

Solution 1. L'ensemble des valeurs prises par X est \mathbb{N}^* . $\mathbb{P}(X = 1) = \frac{2}{3}$.

Soit $n \geq 2$. L'événement $\{X = n\}$ est l'événement « on obtient face à chacun des $n - 1$ premiers lancers et pile au n -ème ».

Sa probabilité est $\underbrace{\frac{1}{3} \times \dots \times \frac{1}{3}}_{n-1} \times \frac{2}{3} = \frac{2}{3^n}$, ce qui reste vrai pour $n = 1$.

En résumé, $\forall n \in \mathbb{N}^*, \mathbb{P}(X = n) = \frac{2}{3^n}$.

Exercice 2. Soit $\lambda \in]0, +\infty[$. Soit X une variable aléatoire réelle sur \mathbb{N} telle que $\forall n \in \mathbb{N}, \mathbb{P}(X = n) = \frac{\lambda}{n!}$.

Déterminer λ .

Solution 2. $\sum_{n=0}^{+\infty} \mathbb{P}(X = n) = 1 \Leftrightarrow \lambda \sum_{n=0}^{+\infty} \frac{1}{n!} = 1 \Leftrightarrow \lambda e = 1 \Leftrightarrow \lambda = \frac{1}{e}$.

⇒ **Commentaire .** Dans l'énoncé ci-dessus, on ne s'est pas préoccupé de l'espace probabilisé sur lequel on travaillait de sorte que l'énoncé est à priori très imprécis. Ceci « n'est pas grave » en classe préparatoire et se produira souvent.

Exercice 3. Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Déterminer $\lim_{x \rightarrow +\infty} \mathbb{P}(X \leq x)$.

Solution 3. Soit $(a, b) \in \mathbb{R}^2$ tel que $a \leq b$. Alors, $\{X \leq a\} \subset \{X \leq b\}$ puis, par croissance de \mathbb{P} , $\mathbb{P}(X \leq a) \leq \mathbb{P}(X \leq b)$. La fonction $F : x \mapsto \mathbb{P}(X \leq x)$ est donc une fonction croissante sur \mathbb{R} . On en déduit que $\lim_{x \rightarrow +\infty} \mathbb{P}(X \leq x)$ existe dans $[0, 1]$. De plus, $\lim_{x \rightarrow +\infty} \mathbb{P}(X \leq x) = \lim_{\substack{n \rightarrow +\infty \\ n \in \mathbb{N}}} \mathbb{P}(X \leq n)$.

Pour $n \in \mathbb{N}$, posons $A_n = \{X \leq n\}$. $(A_n)_{n \in \mathbb{N}}$ est une suite croissante d'événements. Par continuité croissante,

$$\lim_{x \rightarrow +\infty} \mathbb{P}(X \leq x) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \mathbb{P}(\Omega) = 1.$$

1.3 Fonctions de variables aléatoires

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Si X est une variable aléatoire discrète sur Ω à valeurs dans un ensemble E et f est une application de E vers un ensemble F , on peut définir la composée $f \circ X$ que l'on note abusivement $f(X)$. L'ensemble des valeurs prises par $f(X)$ est $f(X(\Omega))$. C'est un sous-ensemble de F . Avec ces notations :

Théorème 2. $f(X)$ est une variable aléatoire discrète sur l'espace $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans F .

DÉMONSTRATION. Par hypothèse $X(\Omega)$ est au plus dénombrable et donc $f(X(\Omega))$ est au plus dénombrable. Soit $y \in F$. On a $(f(X))^{-1}(\{y\}) = X^{-1}(f^{-1}(\{y\}))$. $f^{-1}(\{y\})$ est une partie de E puis $f^{-1}(\{y\}) \cap X(\Omega)$ est une partie de E au plus dénombrable. Ainsi, $(f(X))^{-1}(\{y\})$ est une réunion au plus dénombrable d'éléments de \mathcal{A} et est donc un élément de \mathcal{A} . □

Ainsi, si X est une variable aléatoire réelle, on peut définir les variables aléatoires $X^2, |X|, e^X \dots$

On s'intéresse maintenant à la loi de probabilité de $f(X)$.

Théorème 3. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soient X une variable aléatoire discrète sur Ω à valeurs dans un ensemble non vide E et f une application de E vers un ensemble non vide F .

On pose $Y = f(X)$. La loi de probabilité \mathbb{P}_Y de Y est donnée par

$$\forall y \in F, \mathbb{P}_Y(\{y\}) = \mathbb{P}(Y = y) = \sum_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x).$$

DÉMONSTRATION. L'événement $\{Y = y\}$ est la réunion disjointe des $\{X = x\}$, $x \in f^{-1}(\{y\})$. □

Sinon, il est clair que si X et Y sont deux variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un ensemble E et si f est une application de l'ensemble E vers un ensemble F :

Théorème 4. Si $X \sim Y$, alors $f(X) \sim f(Y)$.

2 Couples de variables aléatoires discrètes. n-uplets de variables aléatoires discrètes

2.1 Couples de variables aléatoires discrètes

2.1.1 Définition d'un couple de variables aléatoires discrètes

DÉFINITION 3. Soient X et Y deux variables aléatoires discrètes sur un espace probabilisable (Ω, \mathcal{A}) à valeurs dans des ensembles non vides E et E' respectivement.

Le couple de variables aléatoires (X, Y) est l'application $(X, Y) : \Omega \rightarrow E \times E'$.
 $\omega \mapsto (X(\omega), Y(\omega))$

Avec les notations de la définition ci-dessus :

Théorème 5. Le couple (X, Y) est une variable aléatoire discrète sur l'espace probabilisable (Ω, \mathcal{A}) à valeurs dans $E \times E'$.

DÉMONSTRATION. $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables. Ensuite,

$$(X, Y)(\Omega) = \{(X(\omega), Y(\omega)), \omega \in \Omega\} \subset \{(X(\omega), Y(\omega')), (\omega, \omega') \in \Omega^2\} = X(\Omega) \times Y(\Omega).$$

On sait que $X(\Omega) \times Y(\Omega)$ est une partie au plus dénombrable de $E \times E'$ et donc $(X, Y)(\Omega)$ est une partie au plus dénombrable de $E \times E'$.

Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. L'ensemble $\{(X, Y) = (x, y)\}$ est l'ensemble

$$\{\omega \in \Omega / (X(\omega), Y(\omega)) = (x, y)\} = \{\omega \in \Omega / X(\omega) = x\} \cap \{\omega \in \Omega / Y(\omega) = y\} = X^{-1}(x) \cap Y^{-1}(y).$$

Donc, $\{(X, Y) = (x, y)\}$ est bien un élément de \mathcal{A} (c'est-à-dire un événement) en tant qu'intersection de deux éléments de \mathcal{A} . Ceci montre que (X, Y) est une variable aléatoire sur (Ω, \mathcal{A}) à valeurs dans $E \times E'$. □

2.1.2 Loi conjointe d'un couple. Loïs marginales

Il s'agit normalement de donner toutes les probabilités du type $\mathbb{P}((X, Y) \in A \times B)$ où $(A, B) \in E \times E'$. On a déjà signalé qu'il suffisait de connaître les atomes de probabilités $\mathbb{P}((X, Y) = (x, y))$, $(x, y) \in E \times E'$ pour calculer ce type de probabilités. Donc :

DÉFINITION 4. Soient X et Y deux variables aléatoires discrètes sur une espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans des ensembles non vides E et E' respectivement.

La **loi conjointe** des variables X et Y est la loi du couple (X, Y) .

Cette loi est l'application
$$\begin{aligned} (X, Y)(\Omega) &\rightarrow [0, 1] \\ (x, y) &\mapsto \mathbb{P}((X, Y) = (x, y)) \end{aligned}$$

⇒ **Commentaire.** On peut aussi prendre en définition de la loi conjointe de X et Y l'application

$$\begin{aligned} X(\Omega) \times Y(\Omega) &\rightarrow [0, 1] \\ (x, y) &\mapsto \mathbb{P}((X, Y) = (x, y)) \end{aligned}$$

étant entendu qu'un couple (x, y) de $X(\Omega) \times Y(\Omega)$ n'est pas toujours une valeur prise par le couple (X, Y) . Un couple (x, y) de $X(\Omega) \times Y(\Omega)$ est un couple de la forme $(X(\omega), Y(\omega'))$ où $(\omega, \omega') \in \Omega^2$ alors qu'un couple de $(X, Y)(\Omega)$ est un couple de la forme $(X(\omega), Y(\omega))$ où $\omega \in \Omega$. Quelque soit la définition adoptée, l'événement $(X, Y) = (x, y)$ est l'événement $\{X = x\} \cap \{Y = y\}$. Cet événement est vide si $(x, y) \notin (X, Y)(\Omega)$ et sa probabilité est donc nulle. En résumé, donner la loi conjointe des variables X ou Y , c'est donner tous les

$$\mathbb{P}(\{X = x\} \cap \{Y = y\}), (x, y) \in X(\Omega) \times Y(\Omega).$$

La probabilité $\mathbb{P}((X, Y) = (x, y))$ peut se aussi se noter $\mathbb{P}(X = x, Y = y)$ (le plus fréquent) ou $\mathbb{P}(\{X = x\} \cap \{Y = y\})$ ou $\mathbb{P}(X = x \text{ et } Y = y)$ (le moins fréquent).

DÉFINITION 5. Soient X et Y deux variables aléatoires discrètes sur une espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans des ensembles non vides E et E' respectivement.

La **première loi marginale** (resp. **deuxième loi marginale**) du couple (X, Y) est la loi de X (resp. la loi de Y).

Puisque $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables, on peut poser $X(\Omega) = \{x_i, i \in I\}$ et $Y(\Omega) = \{y_j, j \in J\}$ où I est un ensemble de la forme $\llbracket 1, p \rrbracket$ ($p \in \mathbb{N}^*$) ou \mathbb{N} , et J est un ensemble de la forme $\llbracket 1, q \rrbracket$ ($q \in \mathbb{N}^*$) ou \mathbb{N} . Pour $(i, j) \in I \times J$, on pose alors $p_{i,j} = \mathbb{P}(\{X = x_i\} \cap \{Y = y_j\})$.

Théorème 6.

1) a) La famille d'événements $(\{X = x_i\} \cap \{Y = y_j\})_{(i,j) \in I \times J}$ est un système complet d'événements.

b) La famille de réels $(p_{i,j})_{(i,j) \in I \times J}$ est sommable et de plus,
$$\sum_{(i,j) \in I \times J} p_{i,j} = 1.$$

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

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

DÉMONSTRATION .

- Vérifions que $(\{X = x_i\} \cap \{Y = y_j\})_{(i,j) \in I \times J}$ est un système complet d'événements.

Soit $((i, j), (k, l)) \in (I \times J)^2$ tel que $(i, j) \neq (k, l)$. On a ou bien $i \neq k$ et dans ce cas, $\{X = x_i\} \cap \{X = x_k\} = \emptyset$ (car $x_k \neq x_i$ et car X est une application) puis $(\{X = x_i\} \cap \{Y = y_j\}) \cap (\{X = x_k\} \cap \{Y = y_l\}) = \emptyset$, ou bien $j \neq l$ et dans ce cas, $\{Y = y_j\} \cap \{Y = y_l\} = \emptyset$ puis $(\{X = x_i\} \cap \{Y = y_j\}) \cap (\{X = x_k\} \cap \{Y = y_l\}) = \emptyset$. Donc, les événements $\{X = x_i\} \cap \{Y = y_j\}$, $((i, j), (k, l)) \in (I \times J)^2$, sont deux à deux disjoints.

Soit $\omega \in \Omega$. Il existe $(i_0, j_0) \in I \times J$ tel que $X(\omega) = x_{i_0}$ et $Y(\omega) = y_{j_0}$ et donc tel que $\omega \in \{X = x_{i_0}\} \cap \{Y = y_{j_0}\}$. Ceci montre que tout $\omega \in \Omega$ appartient à $\bigcup_{(i,j) \in I \times J} \{X = x_i\} \cap \{Y = y_j\}$ puis que $\Omega = \bigcup_{(i,j) \in I \times J} \{X = x_i\} \cap \{Y = y_j\}$.

On a montré que $(\{X = x_i\} \cap \{Y = y_j\})_{(i,j) \in I \times J}$ est un système complet d'événements.

- Soit $i \in I$. $\{X = x_i\} = \bigcup_{j \in J} (\{X = x_i\} \cap \{Y = y_j\})$ et ces événements sont deux à deux disjoints.

$$\text{Donc, } \mathbb{P}(X = x_i) = \sum_{j \in J} \mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) = \sum_{j \in J} p_{i,j}.$$

$$\text{De même, pour tout } j \in J, \mathbb{P}(Y = y_j) = \sum_{i \in I} \mathbb{P}(\{X = x_i\} \cap \{Y = y_j\}) = \sum_{i \in I} p_{i,j}.$$

- D'après ce qui précède, pour tout $i \in I$, $\sum_{j \in J} |p_{i,j}| = \sum_{j \in J} p_{i,j} = \mathbb{P}(X = x_i) \leq 1 < +\infty$ puis $\sum_{i \in I} \left(\sum_{j \in J} |p_{i,j}| \right) = \sum_{i \in I} \mathbb{P}(X = x_i) = 1 < +\infty$.

$$\text{Donc, la famille } (p_{i,j})_{(i,j) \in I \times J} \text{ est sommable et } \sum_{(i,j) \in I \times J} p_{i,j} = \sum_{i \in I} \left(\sum_{j \in J} p_{i,j} \right) = 1.$$

□

⇒ **Commentaire .** Le théorème précédent fournit le calcul des lois de X et Y à partir de la loi du couple (X, Y) . Par exemple, si $X(\Omega) = \{x_1, x_2\}$ et $Y(\Omega) = \{y_1, y_2, y_3\}$, on peut représenter la loi du couple (X, Y) dans un tableau à double entrée. On comprend alors l'expression **lois marginales** : les lois de X et Y sont écrites **en marge** (basse ou droite) de ce tableau.

	y_1	y_2	y_3	Loi de X
x_1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$\mathbb{P}(X = x_1) = p_{1,1} + p_{1,2} + p_{1,3}$
x_2	$p_{2,1}$	$p_{2,2}$	$p_{2,3}$	$\mathbb{P}(X = x_2) = p_{2,1} + p_{2,2} + p_{2,3}$
Loi de Y	$\mathbb{P}(Y = y_1)$	$\mathbb{P}(Y = y_2)$	$\mathbb{P}(Y = y_3)$	1

Exercice 4. (d'après un exercice d'oral de la banque CCP)

Soit (X, Y) un couple de variables aléatoires à valeurs dans \mathbb{N}^2 dont la loi est donnée par :

$$\forall (j, k) \in \mathbb{N}^2, \mathbb{P}(X = j, Y = k) = \frac{(j+k) \left(\frac{1}{2}\right)^{j+k}}{e^j k!}.$$

- 1) Vérifier qu'il s'agit bien d'une loi de couple.
- 2) Déterminer les lois marginales de ce couple.

$$\text{Solution 4. Pour } (j, k) \in \mathbb{N}^2, \text{ on pose } p_{j,k} = \mathbb{P}(X = j, Y = k) = \frac{(j+k) \left(\frac{1}{2}\right)^{j+k}}{e^j k!}.$$

- 1) Il s'agit de vérifier que la famille $(p_{j,k})_{(j,k) \in \mathbb{N}^2}$ est sommable et que $\sum_{j=0}^{+\infty} \left(\sum_{k=0}^{+\infty} p_{j,k} \right) = 1$.

Premier calcul. Soit $j \in \mathbb{N}$.

$$\begin{aligned}\sum_{k=0}^{+\infty} |p_{j,k}| &= \sum_{k=0}^{+\infty} \frac{(j+k) \left(\frac{1}{2}\right)^{j+k}}{e j! k!} = \frac{1}{e} \left(\frac{j}{2^j j!} \sum_{k=0}^{+\infty} \frac{\left(\frac{1}{2}\right)^k}{k!} + \frac{1}{2^{j+1} j!} \sum_{k=1}^{+\infty} \frac{\left(\frac{1}{2}\right)^{k-1}}{(k-1)!} \right) \\ &= \frac{1}{e} \left(\frac{j}{2^j j!} e^{\frac{1}{2}} + \frac{1}{2^{j+1} j!} e^{\frac{1}{2}} \right) = \frac{2j+1}{e^{\frac{1}{2}} 2^{j+1} j!} < +\infty.\end{aligned}$$

Ensuite,

$$\begin{aligned}\sum_{j=0}^{+\infty} \left(\sum_{k=0}^{+\infty} |p_{j,k}| \right) &= \sum_{j=0}^{+\infty} \frac{2j+1}{e^{\frac{1}{2}} 2^{j+1} j!} = \frac{1}{e^{\frac{1}{2}}} \left(\frac{2}{2^2} \sum_{j=1}^{+\infty} \frac{\left(\frac{1}{2}\right)^{j-1}}{(j-1)!} + \frac{1}{2} \sum_{j=0}^{+\infty} \frac{\left(\frac{1}{2}\right)^j}{j!} \right) \\ &= \frac{1}{e^{\frac{1}{2}}} \left(\frac{1}{2} e^{\frac{1}{2}} + \frac{1}{2} e^{\frac{1}{2}} \right) = 1 < +\infty.\end{aligned}$$

Donc, la famille $(p_{j,k})_{(j,k) \in \mathbb{N}^2}$ est sommable et $= \sum_{(j,k) \in \mathbb{N}^2} p_{j,k} = \sum_{j=0}^{+\infty} \left(\sum_{k=0}^{+\infty} p_{j,k} \right) = 1$. Ceci montre que la loi fournie est effectivement une loi de couple.

Deuxième calcul. Supposons acquise la sommabilité de la suite $(p_{j,k})_{(j,k) \in \mathbb{N}^2}$. D'après le théorème de sommation par paquets,

$$\begin{aligned}\sum_{(j,k) \in \mathbb{N}^2} p_{j,k} &= \sum_{n=0}^{+\infty} \left(\sum_{j+k=n} p_{j,k} \right) = \sum_{n=0}^{+\infty} \left(\sum_{j+k=n} \frac{(j+k) \left(\frac{1}{2}\right)^{j+k}}{e j! k!} \right) \\ &= \frac{1}{e} \sum_{n=1}^{+\infty} \left(\frac{n \left(\frac{1}{2}\right)^n}{n!} \sum_{j+k=n} \frac{n!}{j! k!} \right) = \frac{1}{e} \sum_{n=1}^{+\infty} \left(\frac{1}{(n-1)!} \left(\frac{1}{2}\right)^n (1+1)^n \right) = \frac{1}{e} \sum_{n=1}^{+\infty} \frac{1}{(n-1)!} = \frac{1}{e} \times e \\ &= 1 < +\infty.\end{aligned}$$

Puisque les $p_{j,k}$ sont des réels positifs, on sait que ce calcul montre en particulier que la famille est sommable.

2) Soit $j \in \mathbb{N}$. D'après la question précédente,

$$\mathbb{P}(X = j) = \sum_{k=0}^{+\infty} p_{j,k} = \frac{2j+1}{e^{\frac{1}{2}} 2^{j+1} j!}.$$

Par symétrie des rôles, pour tout $k \in \mathbb{N}$,

$$\mathbb{P}(Y = k) = \sum_{j=0}^{+\infty} p_{j,k} = \frac{2k+1}{e^{\frac{1}{2}} 2^{k+1} k!}.$$

2.1.3 Lois conditionnelles

DÉFINITION 6. Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles E et E' respectivement.

Soit $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$. La **loi conditionnelle de X sachant $\{Y = y\}$** est la loi de la variable aléatoire X dans l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P}_{\{Y=y\}})$ c'est-à-dire :

$$\forall x \in X(\Omega), \mathbb{P}_{Y=y}(X = x) = \frac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(Y = y)}.$$

Soit $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$. La loi conditionnelle de Y sachant $\{X = x\}$ est la loi de la variable aléatoire Y dans l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P}_{\{X=x\}})$ c'est-à-dire :

$$\forall y \in Y(\Omega), \mathbb{P}_{X=x}(Y = y) = \frac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(X = x)}.$$

Le théorème suivant est immédiat.

Théorème 7. Pour tout y de $Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$,

$$\forall x \in X(\Omega), \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(Y = y) \times \mathbb{P}_{Y=y}(X = x).$$

Pour tout x de $X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$,

$$\forall y \in Y(\Omega), \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \mathbb{P}(X = x) \times \mathbb{P}_{X=x}(Y = y).$$

Si pour tout y de $Y(\Omega)$, $\mathbb{P}(Y = y) \neq 0$, alors

$$\forall x \in X(\Omega), \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \sum_{y \in Y(\Omega)} \mathbb{P}(Y = y) \times \mathbb{P}_{Y=y}(X = x).$$

Si pour tout x de $X(\Omega)$, $\mathbb{P}(X = x) \neq 0$, alors

$$\forall y \in Y(\Omega), \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x) \times \mathbb{P}_{X=x}(Y = y).$$

2.2 Généralisation à un n -uplet de variables aléatoires discrètes

On généralise brièvement ce qui précède à n variables aléatoires X_1, \dots, X_n , ($n \geq 2$).

DÉFINITION 7. Soit (Ω, \mathcal{A}) un espace probabilisable. Soient X_1, \dots, X_n , des variables aléatoires discrètes sur cet espace ($n \geq 2$), à valeurs dans des ensembles E_1, \dots, E_n respectivement.

L'application $\Omega \rightarrow E_1 \times \dots \times E_n$ est appelée **n -uplet de variables aléatoires** sur l'espace (Ω, \mathcal{A}) et se note (X_1, \dots, X_n) . C'est une variable aléatoire sur l'espace (Ω, \mathcal{A}) à valeurs dans $E_1 \times \dots \times E_n$.

Dans le cas particulier où toutes les variables aléatoires sont réelles, le vecteur $\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ est appelé **vecteur aléatoire discret**.

DÉFINITION 8. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soient X_1, \dots, X_n , des variables aléatoires discrètes sur cet espace ($n \geq 2$), à valeurs dans des ensembles E_1, \dots, E_n respectivement.

La **loi conjointe** des variables aléatoires X_1, \dots, X_n est la loi du n -uplet (X_1, \dots, X_n) .

Les **lois marginales** du n -uplet (X_1, \dots, X_n) sont les lois des variables X_1, \dots, X_n .

Quand on dispose de la loi du n -uplet (X_1, \dots, X_n) , les lois marginales s'obtiennent de la façon suivante : pour $k \in \llbracket 1, n \rrbracket$ puis $a_k \in X_k(\Omega)$ donnés,

$$\mathbb{P}(\{X_k = a_k\}) = \sum_{\substack{(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \\ \in (X_1, \dots, X_{k-1}, X_{k+1}, \dots, X_n)(\Omega)}} \mathbb{P}(X = x_1, \dots, X_{k-1} = x_{k-1}, X_k = a_k, X_{k+1} = x_{k+1}, \dots, X_n = x_n).$$

3 Variables aléatoires indépendantes

3.1 Cas de deux variables

3.1.1 Définition de l'indépendance de deux variables aléatoires

DÉFINITION 9. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soient X et Y deux variables aléatoires sur cet espace à valeurs dans des ensembles non vides E_1 et E_2 respectivement.

Les variables X et Y sont **indépendantes** si et seulement si

$$\forall (x, y) \in (X(\Omega) \times Y(\Omega)), \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y).$$

Il revient au même de dire que pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $\{X = x\}$ et $\{Y = y\}$ sont indépendants.

Notation. Quand X et Y sont indépendantes, on écrit $X \perp\!\!\!\perp Y$.

Exercice 5. Les variables de l'exercice n° 4, page 6, sont-elles indépendantes ?

Solution 5. Pour tout $(j, k) \in \mathbb{N}^2$, $\mathbb{P}(X = j, Y = k) = \frac{(j+k) \left(\frac{1}{2}\right)^{j+k}}{e^j j! k!}$.

D'autre part, on a vu que pour tout $(j, k) \in \mathbb{N}^2$, $\mathbb{P}(X = j) = \frac{2^j + 1}{e^{\frac{1}{2}} 2^{j+1} j!}$ et $\mathbb{P}(Y = k) = \frac{2k + 1}{e^{\frac{1}{2}} 2^{k+1} k!}$.

Par suite, $\mathbb{P}(\{X = 0\} \cap \{Y = 0\}) = \mathbb{P}(X = 0, Y = 0) = 0$ et d'autre part, $\mathbb{P}(\{X = 0\}) \times \mathbb{P}(\{Y = 0\}) = \left(\frac{1}{2e^{\frac{1}{2}}}\right)^2 \neq 0$. Ainsi, $\mathbb{P}(\{X = 0\} \cap \{Y = 0\}) \neq \mathbb{P}(\{X = 0\}) \times \mathbb{P}(\{Y = 0\})$.
Donc, il existe $(j, k) \in \mathbb{N}^2$ tel que $\mathbb{P}(\{X = j\} \cap \{Y = k\}) \neq \mathbb{P}(\{X = j\}) \times \mathbb{P}(\{Y = k\})$. Les variables X et Y ne sont pas indépendantes.

3.1.2 Fonctions de deux variables indépendantes

Théorème 8. Soient X et Y deux variables aléatoires discrètes indépendantes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles E et E' respectivement.

Pour toute fonction f définie sur E à valeurs dans un certain ensemble F et toute fonction g définie sur E' à valeurs dans un certain ensemble F' , les variables $f(X)$ et $g(Y)$ sont indépendantes.

DÉMONSTRATION. Soit $(z, z') \in f(X)(\Omega) \times g(Y)(\Omega)$. L'événement $(f(X) = z, g(Y) = z')$ est la réunion disjointe des événements $(X = x, Y = y)$ où $(x, y) \in X(\Omega) \times Y(\Omega)$ est tel que $f(x) = z$ et $g(y) = z'$. Donc,

$$\begin{aligned} \mathbb{P}(f(X) = z, g(Y) = z') &= \sum_{(x, y) \in f^{-1}(\{z\}) \times g^{-1}(\{z'\})} \mathbb{P}(X = x, Y = y) \\ &= \sum_{(x, y) \in f^{-1}(\{z\}) \times g^{-1}(\{z'\})} \mathbb{P}(X = x) \times \mathbb{P}(Y = y) \text{ (car } X \text{ et } Y \text{ sont indépendantes)} \\ &= \left(\sum_{x \in f^{-1}(\{z\})} \mathbb{P}(X = x) \right) \left(\sum_{y \in g^{-1}(\{z'\})} \mathbb{P}(Y = y) \right) \\ &= \mathbb{P}(f(X) = z) \times \mathbb{P}(g(Y) = z'). \end{aligned}$$

Donc, les variables $f(X)$ et $g(Y)$ sont indépendantes. □

Ainsi, si par exemple X et Y sont deux variables aléatoires réelles indépendantes, alors X^2 et e^Y sont deux variables aléatoires indépendantes.

3.1.3 Somme de deux variables indépendantes à valeurs entières

De manière générale, si X et Y sont deux variables réelles discrètes telles que $X(\Omega) = Y(\Omega) = \mathbb{N}$, $X + Y$ est une variable à valeurs dans \mathbb{N} dont la loi est donnée par :

$$\forall n \in \mathbb{N}, \mathbb{P}(X + Y = n) = \sum_{i+j=n} \mathbb{P}(X = i, Y = j).$$

Si de plus, les variables X et Y sont indépendantes, on obtient la loi de $X + Y$ à partir des lois de X et de Y :

$$\forall n \in \mathbb{N}, \mathbb{P}(X + Y = n) = \sum_{i+j=n} \mathbb{P}(X = i) \times \mathbb{P}(Y = j).$$

Rappelons maintenant un calcul fait en maths sup : on montre que X et Y sont deux variables indépendantes suivant la loi binomiale $\mathcal{B}(n, p)$ et la loi binomiale $\mathcal{B}(m, p)$ respectivement ($(n, m) \in (\mathbb{N}^*)^2$ et $p \in]0, 1[$), alors la variable $X + Y$ suit la loi binomiale $\mathcal{B}(n + m, p)$.

On peut poser $X(\Omega) = \mathbb{N}$ puis pour $i \in \mathbb{N}$, $\mathbb{P}(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}$ (on rappelle que si $i > n$, $\binom{n}{i} = 0$ puis que $\mathbb{P}(X = i) = 0$). On pose aussi $Y(\Omega) = \mathbb{N}$ puis pour $j \in \mathbb{N}$, $\mathbb{P}(Y = j) = \binom{m}{j} p^j (1 - p)^{m-j}$. On a alors $(X + Y)(\Omega) = \mathbb{N}$ puis, pour $k \in \mathbb{N}$.

$$\begin{aligned} \mathbb{P}(X + Y = k) &= \sum_{i+j=k} \mathbb{P}(\{X = i\} \cap \{Y = j\}) = \sum_{i+j=k} \mathbb{P}(X = i) \times \mathbb{P}(Y = j) \text{ (car } X \text{ et } Y \text{ sont indépendantes)} \\ &= \sum_{i+j=k} \binom{n}{i} p^i (1 - p)^{n-i} \binom{m}{j} p^j (1 - p)^{m-j} = p^k (1 - p)^{(n+m)-k} \sum_{i+j=k} \binom{n}{i} \binom{m}{j}. \end{aligned}$$

Maintenant, $\sum_{i+j=k} \binom{n}{i} \binom{m}{j}$ est le coefficient de X^k dans le développement de

$$\begin{aligned} \left(\sum_{i=0}^{+\infty} \binom{n}{i} X^i \right) \left(\sum_{j=0}^{+\infty} \binom{m}{j} X^j \right) &= \left(\sum_{i=0}^n \binom{n}{i} X^i \right) \left(\sum_{j=0}^m \binom{m}{j} X^j \right) \\ &= (1 + X)^n (1 + X)^m = (1 + X)^{n+m} = \sum_{k=0}^{n+m} \binom{n+m}{k} X^k \\ &= \sum_{k=0}^{+\infty} \binom{n+m}{k} X^k \end{aligned}$$

En identifiant les coefficients, on obtient l'identité de VANDERMONDE : $\sum_{i+j=k} \binom{n}{i} \binom{m}{j} = \binom{n+m}{k}$ (y compris si $k > n + m$). Finalement, $\mathbb{P}(X + Y = k) = \binom{n+m}{k} p^k (1 - p)^{(n+m)-k}$ et donc $X + Y \sim \mathcal{B}(n + m, p)$.

3.2 Généralisation à n variables

DÉFINITION 10. Soient X_1, \dots, X_n , $n \geq 2$, des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles non vides E_1, \dots, E_n respectivement.

Les variables X_1, \dots, X_n , sont **deux à deux indépendantes** si et seulement si pour tout couple $(i, j) \in \llbracket 1, n \rrbracket^2$ tel que $i \neq j$, les variables X_i et X_j sont indépendantes.

DÉFINITION 11. Soient X_1, \dots, X_n , $n \geq 2$, des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles non vides E_1, \dots, E_n respectivement.

Les variables X_1, \dots, X_n , sont **mutuellement indépendantes** (ou plus simplement **indépendantes**) si et seulement si pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,

$$\mathbb{P}(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}) = \mathbb{P}(X_1 = x_1) \times \dots \times \mathbb{P}(X_n = x_n).$$

Théorème 9. Soient X_1, \dots, X_n , $n \geq 2$, des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles non vides E_1, \dots, E_n respectivement.

Les variables X_1, \dots, X_n , sont mutuellement indépendantes si et seulement si pour tout $(A_1, \dots, A_n) \in \mathcal{P}(E_1) \times \dots \times \mathcal{P}(E_n)$,

$$\mathbb{P}(\{X_1 \in A_1\} \cap \dots \cap \{X_n \in A_n\}) = \mathbb{P}(X_1 \in A_1) \times \dots \times \mathbb{P}(X_n \in A_n).$$

DÉMONSTRATION. Supposons que pour tout $(A_1, \dots, A_n) \in \mathcal{P}(E_1) \times \dots \times \mathcal{P}(E_n)$, on a

$$\mathbb{P}(\{X_1 \in A_1\} \cap \dots \cap \{X_n \in A_n\}) = \mathbb{P}(X_1 \in A_1) \times \dots \times \mathbb{P}(X_n \in A_n).$$

Alors pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$

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

Les variables X_1, \dots, X_n sont donc mutuellement indépendantes.

Réciproquement, supposons que les variables X_1, \dots, X_n soient mutuellement indépendantes. Soit $(A_1, \dots, A_n) \in E_1 \times \dots \times E_n$.

S'il existe $i_0 \in \llbracket 1, n \rrbracket$ tel que $A_{i_0} \cap X_{i_0}(\Omega) = \emptyset$ alors $\mathbb{P}(\{X_{i_0} \in A_{i_0}\}) = 0$ puis $\prod_{i=1}^n \mathbb{P}(\{X_i \in A_i\}) = 0$. D'autre part, $\bigcap_{i=1}^n \{X_i \in A_i\} = \emptyset$

$$\text{et donc } \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) = 0 = \prod_{i=1}^n \mathbb{P}(\{X_i \in A_i\}).$$

On suppose dorénavant que pour tout $i \in \llbracket 1, n \rrbracket$, $A_i \cap X_i(\Omega) \neq \emptyset$.

$$\begin{aligned} \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) &= \mathbb{P}\left(\bigcup_{(x_i)_{1 \leq i \leq n} \in \prod_{i=1}^n (A_i \cap X_i(\Omega))} \left(\bigcap_{i=1}^n \{X_i = x_i\}\right)\right) \\ &= \sum_{(x_i)_{1 \leq i \leq n} \in \prod_{i=1}^n (A_i \cap X_i(\Omega))} \mathbb{P}\left(\bigcap_{i=1}^n \{X_i = x_i\}\right) \quad (\text{événements deux à deux disjoints}) \\ &= \sum_{(x_i)_{1 \leq i \leq n} \in \prod_{i=1}^n (A_i \cap X_i(\Omega))} \left(\prod_{i=1}^n \mathbb{P}(X_i = x_i)\right) \\ &= \prod_{i=1}^n \left(\sum_{x_i \in A_i \cap X_i(\Omega)} \mathbb{P}(\{X_i = x_i\})\right) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i). \end{aligned}$$

□

Théorème 10. Soient X_1, \dots, X_n , $n \geq 2$, des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles non vides E_1, \dots, E_n respectivement.

Si les variables X_1, \dots, X_n , sont mutuellement indépendantes, alors pour toute $\{i_1, \dots, i_k\}$ de $\llbracket 1, n \rrbracket$ (avec $2 \leq k \leq n$ et $1 \leq i_1 < \dots < i_k \leq n$), les variables X_{i_1}, \dots, X_{i_k} sont mutuellement indépendantes.

En particulier, si les variables X_1, \dots, X_n , sont mutuellement indépendantes, alors les variables X_1, \dots, X_n , sont deux à deux indépendantes.

DÉMONSTRATION. Soit $I = \{i_1, \dots, i_k\} \subset \llbracket 1, n \rrbracket$ avec $k \geq 2$ et $i_1 < i_2 < \dots < i_k$. Montrons que les variables X_{i_1}, \dots, X_{i_k} sont mutuellement indépendantes.

Soit $(x_{i_1}, \dots, x_{i_k}) \in X_{i_1}(\Omega) \times \dots \times X_{i_k}(\Omega)$. Pour $i \in \llbracket 1, n \rrbracket$, on pose $A_i = \{X_i = x_i\}$ si $i \in I$ et $A_i = X_i(\Omega)$ si $i \notin I$. Alors,

$$\begin{aligned}
\mathbb{P}(\{X_{i_1} = x_{i_1}\} \cap \dots \cap \{X_{i_k} = x_{i_k}\}) &= \mathbb{P}(\{X_{i_1} \in A_{i_1}\} \cap \dots \cap \{X_{i_k} \in A_{i_k}\}) \\
&= \mathbb{P}(\{X_{i_1} \in A_{i_1}\} \cap \dots \cap \{X_{i_k} \in A_{i_k}\} \cap \Omega \cap \dots \cap \Omega) \\
&= \mathbb{P}\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i) \\
&= \mathbb{P}(X_{i_1} = x_{i_1}) \times \dots \times \mathbb{P}(X_{i_k} = x_{i_k}) \quad (\text{car si } i \notin I, \mathbb{P}(X_i \in A_i) = \mathbb{P}(\Omega) = 1).
\end{aligned}$$

Ceci montre que les variables X_{i_1}, \dots, X_{i_k} sont mutuellement indépendantes. □

Théorème 11. Soient X_1, \dots, X_n , $n \geq 2$, des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles non vides E_1, \dots, E_n respectivement.

Si les variables X_1, \dots, X_n , sont mutuellement indépendantes, alors pour tout $m \in \llbracket 1, n-1 \rrbracket$ et toutes fonctions f et g , les variables $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

DÉMONSTRATION. (X_1, \dots, X_m) et (X_{m+1}, \dots, X_n) sont des variables aléatoires (à valeurs dans $E_1 \times \dots \times E_m$ et $E_{m+1} \times \dots \times E_n$ respectivement). Vérifions que ces variables sont indépendantes.

Soient $(x_1, \dots, x_m) \in X_1(\Omega) \times \dots \times X_m(\Omega)$ et $(x_{m+1}, \dots, x_n) \in X_{m+1}(\Omega) \times \dots \times X_n(\Omega)$.

$$\begin{aligned}
\mathbb{P}((X_1, \dots, X_m) = (x_1, \dots, x_m), (X_{m+1}, \dots, X_n) = (x_{m+1}, \dots, x_n)) &= \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) \\
&= \mathbb{P}(X_1 = x_1) \dots \mathbb{P}(X_m = x_m) \mathbb{P}(X_{m+1} = x_{m+1}) \dots \mathbb{P}(X_n = x_n) \\
&= \mathbb{P}((X_1, \dots, X_m) = (x_1, \dots, x_m)) \times \mathbb{P}((X_{m+1}, \dots, X_n) = (x_{m+1}, \dots, x_n)) \\
&\quad (\text{car les variables } X_1, \dots, X_m \text{ d'une part et } X_{m+1}, \dots, X_n \text{ d'autre part sont indépendantes} \\
&\quad \text{d'après le théorème précédent}).
\end{aligned}$$

Ainsi, les deux variables (X_1, \dots, X_m) et (X_{m+1}, \dots, X_n) sont indépendantes. D'après le théorème 8, les variables $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes. □

Plus généralement, on donne sans démonstration le **lemme des coalitions** :

Théorème 12. Soient X_1, \dots, X_n , $n \geq 2$, des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans des ensembles non vides E_1, \dots, E_n respectivement. On suppose que les variables X_1, \dots, X_n sont mutuellement indépendantes.

On considère une partition de $\llbracket 1, n \rrbracket$ du type $\llbracket 1, n \rrbracket = I_1 \cup \dots \cup I_k$ (où les I_k sont non vides et deux à deux disjoints). Pour $j \in \llbracket 1, k \rrbracket$, on note Y_j une variable aléatoire qui est une fonction des variables X_i , $i \in I_j$. Alors, les variables Y_1, \dots, Y_k sont mutuellement indépendantes.

Par exemple, si X_1, \dots, X_6 , sont indépendantes, alors les variables $Y_1 = X_1 + e^{X_3 X_4}$, $Y_2 = X_2 + X_6$ et $Y_3 = X_5$ sont indépendantes.

3.3 Suites de variables aléatoires indépendantes

Intéressons nous au modèle du jeu de pile ou face infini : on lance une pièce de monnaie, éventuellement truquée, on note si on obtient pile ou face et on recommence ... sans prévoir de temps d'arrêt (par exemple, en s'intéressant à la première fois où on obtient pile). A chaque étape, on modélise le lancer par une variable aléatoire X_n suivant une loi de BERNOULLI $\mathcal{B}(p)$, $p \in]0, 1[$, l'événement $X_n = 1$ modélisant l'obtention d'un pile et l'événement $X_n = 0$ modélisant l'obtention d'un face. Ainsi, pour tout $n \in \mathbb{N}^*$, la probabilité de l'événement $X_n = 1$ est p et la probabilité de l'événement $X_n = 0$ est $1 - p$ (et on peut même compliquer en p_n et $1 - p_n$ en changeant de pièce de monnaie à chaque lancer). Chaque variable aléatoire X_n est définie sur un certain espace probabilisé.

On suppose que pour tout $n \in \mathbb{N}^*$, les variables aléatoires X_1, \dots, X_n , sont mutuellement indépendantes et on dit alors que la suite $(X_n)_{n \in \mathbb{N}^*}$ est une **suite de variables mutuellement indépendantes**. Le problème est alors l'existence d'un espace probabilisé commun à toutes ces variables. Le programme officiel demande d'admettre le théorème suivant qui règle le problème :

Théorème 13. Pour toute suite $(X_n)_{n \in \mathbb{N}^*}$ de variables aléatoires discrètes mutuellement indépendantes sur des espaces probabilisés $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$, de lois discrètes \mathcal{L}_n données, il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tel que chaque X_n soit une variable aléatoire sur cet espace, de loi \mathcal{L}_n .

Une suite de variables aléatoires mutuellement indépendantes est appelée **suite i.d.d** (suite de variables aléatoires **indépendantes** et **identiquement distribuées**). Le jeu de pile ou face infini se modélise donc par une suite i.d.d de variables de BERNOULLI.

4 Lois usuelles

On rappelle qu'en maths sup, on a analysé la loi uniforme sur $\llbracket 1, n \rrbracket$ et plus généralement sur $\llbracket a, b \rrbracket$ ($(a, b) \in \mathbb{Z}^2$) notée $\mathcal{U}(\llbracket a, b \rrbracket)$, la loi de BERNOULLI de paramètre p notée $\mathcal{B}(p)$ et plus généralement la loi binomiale de paramètres n et p notée $\mathcal{B}(n, p)$. On redonne d'abord brièvement les lois de maths sup puis on rajoute deux nouvelles lois discrètes.

4.1 Rappels de sup

4.1.1 La loi uniforme

Soit $n \in \mathbb{N}^*$. Une variable aléatoire X suit la loi uniforme sur $\llbracket 1, n \rrbracket$ (ou encore $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$) si et seulement si $X(\Omega) = \llbracket 1, n \rrbracket$ et $\forall k \in \llbracket 1, n \rrbracket, \mathbb{P}(X = k) = \frac{1}{n}$.

Plus généralement, soit $(a, b) \in \mathbb{Z}^2$ tel que $a \leq b$. Une variable aléatoire X suit la loi uniforme sur $\llbracket a, b \rrbracket$ (ou encore $X \sim \mathcal{U}(\llbracket a, b \rrbracket)$) si et seulement si $X(\Omega) = \llbracket a, b \rrbracket$ et $\forall k \in \llbracket a, b \rrbracket, \mathbb{P}(X = k) = \frac{1}{b - a + 1}$.

4.1.2 La loi de BERNOULLI. La loi binomiale

Soit $p \in]0, 1[$. Une variable aléatoire X suit la loi de BERNOULLI de paramètre p (ou encore $X \sim \mathcal{B}(p)$) si et seulement si $X(\Omega) = \{0, 1\}$ et $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$.

Soient $p \in]0, 1[$ et $n \in \mathbb{N}^*$. Une variable aléatoire X suit la loi binomiale de paramètres n et p (ou encore $X \sim \mathcal{B}(n, p)$) si et seulement si $X(\Omega) = \llbracket 0, n \rrbracket$ et $\forall k \in \llbracket 0, n \rrbracket, \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$. On note que $\mathcal{B}(1, p) \sim \mathcal{B}(p)$.

On rappelle que $X \sim X_1 + \dots + X_n$ où les X_i , $1 \leq n$, sont des variables mutuellement indépendantes suivant une loi de BERNOULLI de paramètre p . Ce résultat peut se démontrer par récurrence à partir du résultat réénoncé et redémontré page 10.

4.2 La loi géométrique de paramètre p

On revient sur le modèle du jeu de pile ou face infini. On s'intéresse à une suite $(X_n)_{n \in \mathbb{N}^*}$ de variables mutuellement indépendantes, toutes définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, toutes suivant une loi de BERNOULLI ayant le même paramètre $p \in]0, 1[$. On note X la variable aléatoire égale au rang du premier succès (c'est-à-dire à $\min\{n \in \mathbb{N}^* / X_n = 1\}$).

A priori, $X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$. Pour $k \in \mathbb{N}^*$, $X = k$ si et seulement si on obtient 1 à la k -ème étape et pas avant. Puisque les variables sont mutuellement indépendantes,

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

$$\text{Ensuite, } \mathbb{P}(X = +\infty) = 1 - \mathbb{P}(X \in \mathbb{N}^*) = 1 - \sum_{k=1}^{+\infty} p(1 - p)^{k-1} = 1 - p \times \frac{1}{1 - (1 - p)} = 1 - 1 = 0.$$

On pose donc la définition suivante :

DÉFINITION 12. Soit $p \in]0, 1[$.

On dit qu'une variable aléatoire X , définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, suit la loi géométrique de paramètre p et on écrit $X \sim \mathcal{G}(p)$ si et seulement si $X(\Omega) = \mathbb{N}^*$ et $\forall k \in \mathbb{N}^*, \mathbb{P}(X = k) = p(1 - p)^{k-1}$.

On note que $\sum_{k=1}^{+\infty} p(1 - p)^{k-1} = 1$. La loi géométrique est une loi discrète caractérisée par le fait que c'est une loi **sans mémoire** (ou sans vieillissement) :

Théorème 14.

1) Soit X une variable aléatoire suivant une loi géométrique de paramètre $p \in]0, 1[$. Alors,

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

2) Réciproquement, soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}^*$, vérifiant $(*)$, telle que $0 < \mathbb{P}(X = 1) < 1$. Alors, X suit la loi géométrique de paramètre $p = \mathbb{P}(X = 1)$.

DÉMONSTRATION .

1) Soit $(n, k) \in \mathbb{N}^2$.

$$\mathbb{P}(X > n) = \sum_{i=n+1}^{+\infty} \mathbb{P}(X = i) = \sum_{i=n+1}^{+\infty} p(1-p)^{i-1} = p(1-p)^n \frac{1}{1-(1-p)} = (1-p)^n$$

puis

$$\mathbb{P}_{X>n}(X > n + k) = \frac{\mathbb{P}((X > n) \cap (X > n + k))}{\mathbb{P}(X > n)} = \frac{\mathbb{P}(X > n + k)}{\mathbb{P}(X > n)} = \frac{(1-p)^{n+k}}{(1-p)^n} = (1-p)^k = \mathbb{P}(X > k).$$

2) Posons $p = \mathbb{P}(X = 1) \in]0, 1[$. Soit $k \geq 1$.

$$\begin{aligned} \mathbb{P}(X > k) &= \mathbb{P}((X > k-1) \cap (X > k)) = \mathbb{P}(X > k-1) \times \mathbb{P}_{X>k-1}(X > k) = \mathbb{P}(X > k-1) \times \mathbb{P}_{X>k-1}(X > (k-1) + 1) \\ &= \mathbb{P}(X > k-1) \times \mathbb{P}(X > 1) \\ &= \mathbb{P}(X > k-1) \times \mathbb{P}(X > 1) = \mathbb{P}(X > k-1) \times (1 - \mathbb{P}(X = 1)) \\ &= (1-p)\mathbb{P}(X > k-1). \end{aligned}$$

En tenant compte de $\mathbb{P}(X > 0) = 1$, on obtient pour tout $k \in \mathbb{N}$, $\mathbb{P}(X > k) = (1-p)^k$. Ensuite, pour $k \in \mathbb{N}^*$,

$$\begin{aligned} \mathbb{P}(X = k) &= \mathbb{P}((X > k-1) \setminus (X > k)) = \mathbb{P}(X > k-1) - \mathbb{P}(X > k) \\ &= (1-p)^{k-1} - (1-p)^k = (1-(1-p))(1-p)^{k-1} = p(1-p)^{k-1}. \end{aligned}$$

Ainsi, X suit la loi géométrique de paramètre $p = \mathbb{P}(X = 1)$. □

4.3 Loi de POISSON de paramètre λ **4.3.1 Approximation de la loi binomiale par la loi de POISSON**

On considère une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ suivant chacune une loi binomiale : $\forall n \in \mathbb{N}^*$, $X_n \sim \mathcal{B}(n, p_n)$ où p_n est un réel de $]0, 1[$. Ainsi,

$$\forall k \in \llbracket 0, n \rrbracket, \mathbb{P}(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}.$$

On suppose que la suite $(np_n)_{n \in \mathbb{N}^*}$ converge vers un certain réel strictement positif λ . Soit $k \in \mathbb{N}$. Pour $n > k$,

$$\mathbb{P}(X_n = k) = \frac{n(n-1)\dots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k}.$$

Déjà, puisque k est constant quand n varie, $\frac{n(n-1)\dots(n-k+1)}{k!} p_n^k \underset{n \rightarrow +\infty}{\sim} \frac{n^k}{k!} p_n^k = \frac{(np_n)^k}{k!} \underset{n \rightarrow +\infty}{\sim} \frac{\lambda^k}{k!}$. D'autre part,

$$\begin{aligned} (1-p_n)^{n-k} &= e^{(n-k)\ln(1-\frac{np_n}{n})} \underset{n \rightarrow +\infty}{=} e^{(n+o(n))\ln(1-\frac{\lambda+o(1)}{n})} \underset{n \rightarrow +\infty}{=} e^{(n+o(n))\ln(-\frac{\lambda}{n}+o(\frac{1}{n}))} \\ &\underset{n \rightarrow +\infty}{=} e^{-\lambda+o(1)} \underset{n \rightarrow +\infty}{\sim} e^{-\lambda} \end{aligned}$$

Finalement, pour $k \in \mathbb{N}$ fixé,

$$\mathbb{P}(X_n = k) \underset{n \rightarrow +\infty}{\sim} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Cette constatation nous conduit à la définition du paragraphe suivant.

4.3.2 Définition de la loi de POISSON

DÉFINITION 13. Soit $\lambda \in]0, +\infty[$.

On dit qu'une variable aléatoire X , définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, suit la loi de POISSON de paramètre λ et on écrit $X \sim \mathcal{P}(\lambda)$ si et seulement si $X(\Omega) = \mathbb{N}$ et $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

⇒ **Commentaire.**

◇ On note que $\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$ et en particulier, pour tout $k \in \mathbb{N}, 0 \leq \frac{\lambda^k}{k!} e^{-\lambda} \leq 1$.

◇ Pour k fixé dans \mathbb{N} , si np_n tend vers λ ou encore si $p_n \sim \frac{\lambda}{n}, \binom{n}{k} p_n^k (1-p_n)^{n-k}$ tend vers $\frac{\lambda^k}{k!} e^{-\lambda}$.

Dit autrement, quand n est grand et p est proche de 0 de sorte que l'espérance de $\mathcal{B}(n, p)$ soit proche de $\lambda > 0$, la loi de POISSON de paramètre λ est une « bonne approximation » de la loi binomiale de paramètres n et p .

Une variable aléatoire suivant une loi de POISSON peut donc s'interpréter comme le nombre de succès à la suite d'un grand nombre d'épreuves de BERNOULLI, la probabilité de succès à chaque épreuve étant faible. Pour cette raison, la loi de POISSON est souvent appelée **lois des événements rares**.

La loi de POISSON est par exemple utilisée pour décrire le nombre d'appels à un standard téléphonique durant un intervalle de temps donné, le nombre de taxis passant place de la Concorde à Paris dans un intervalle de temps donné, le nombre de clients rentrant dans un magasin dans un intervalle de temps donné ... □

Exercice 6. Soient X_1 et X_2 deux variables aléatoires indépendantes suivant des lois de POISSON de paramètres respectifs $\lambda_1 > 0$ et $\lambda_2 > 0$.

Déterminer la loi de $X_1 + X_2$.

Solution 6. $(X_1 + X_2)(\Omega) = \mathbb{N}$. Soit alors $n \in \mathbb{N}$.

$$\begin{aligned} \mathbb{P}(X_1 + X_2 = n) &= \sum_{k=0}^n \mathbb{P}(X_1 = k, X_2 = n - k) \\ &= \sum_{k=0}^n \mathbb{P}(X_1 = k) \times \mathbb{P}(X_2 = n - k) \quad (\text{car } X_1 \text{ et } X_2 \text{ sont indépendantes}) \\ &= \sum_{k=0}^n \frac{\lambda_1^k}{k!} e^{-\lambda_1} \frac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2} = \frac{e^{-(\lambda_1 + \lambda_2)}}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda_1^k \lambda_2^{n-k} \\ &= \frac{(\lambda_1 + \lambda_2)^n}{n!} e^{-(\lambda_1 + \lambda_2)}. \end{aligned}$$

Par suite, $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.

5 Espérance, variance

5.1 Espérance d'une variable aléatoire

5.1.1 Définition de l'espérance

DÉFINITION 14. Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^+ \cup \{+\infty\}$. L'**espérance** de X , notée $E(X)$, est l'élément de $[0, +\infty]$ égal à la somme de la famille $(x \mathbb{P}(X = x))_{x \in X(\Omega)}$:

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

Dans le cas où $E(X) \in [0, +\infty[$, on dit que la variable X est d'**espérance finie**. Ceci équivaut à la sommabilité de la famille $(x \mathbb{P}(X = x))_{x \in X(\Omega)}$.

Dans la définition précédente, on a adopté la convention suivante : si $x = +\infty$ et $\mathbb{P}(X = +\infty) = 0$, alors $x \mathbb{P}(X = x) = 0$ et si $x = +\infty$ et $\mathbb{P}(X = +\infty) > 0$, alors $x \mathbb{P}(X = x) = +\infty$.

Pour les variables aléatoires entières (c'est à dire à valeurs dans \mathbb{N}), on a dispose d'une autre expression de l'espérance :

Théorème 15. Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Alors,

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

DÉMONSTRATION . Si $\mathbb{P}(X = +\infty) > 0$, alors $E(X) \geq +\infty \times \mathbb{P}(X = +\infty) = +\infty$. D'autre part, pour tout $n \in \mathbb{N}$, $\mathbb{P}(X \geq n) \geq \mathbb{P}(X = +\infty) > 0$. Donc, $\mathbb{P}(X \geq n)$ ne tend pas vers 0 quand n tend vers $+\infty$ puis la série de terme général $\mathbb{P}(X \geq n)$ est grossièrement divergente. Dans ce cas (les séries étant à termes positifs), on a

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

Dorénavant, on suppose que $\mathbb{P}(X = +\infty) = 0$. On note que

$$\sum_{k=0}^{+\infty} k\mathbb{P}(X = k) = \sum_{k=1}^{+\infty} k\mathbb{P}(X = k) = \sum_{k=1}^{+\infty} \left(\sum_{n=1}^k \mathbb{P}(X = k) \right).$$

On pose donc $A = \{(k, n) \in (\mathbb{N}^*)^2 / 1 \leq n \leq k\}$ puis pour $(k, n) \in A$, $u_{k,n} = \mathbb{P}(X = k)$.

Pour $k \in \mathbb{N}^*$, posons $I_k = \{(k, n) \in (\mathbb{N}^*)^2 / n \leq k\}$ (k est fixé) et pour $n \in \mathbb{N}^*$, posons $J_n = \{(k, n) \in (\mathbb{N}^*)^2 / k \geq n\}$ (n est fixé). $(I_k)_{k \in \mathbb{N}^*}$ et $(J_n)_{n \in \mathbb{N}^*}$ sont des partitions de A .

1er cas. Supposons que la suite $(u_{k,l})_{(k,l) \in A}$ est sommable. D'après le théorème de sommation par paquets,

$$\begin{aligned} E(X) &= \sum_{k=1}^{+\infty} k\mathbb{P}(X = k) = \sum_{k=1}^{+\infty} \left(\sum_{(k,n) \in I_k} u_{k,n} \right) = \sum_{(k,n) \in A} u_{k,n} = \sum_{n=1}^{+\infty} \left(\sum_{(k,n) \in J_n} u_{k,n} \right) \\ &= \sum_{n=1}^{+\infty} \left(\sum_{k=n}^{+\infty} \mathbb{P}(X = k) \right) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n). \end{aligned}$$

2ème cas. Supposons que la famille $(u_{k,l})_{(k,l) \in A}$ n'est pas sommable. Puisque cette famille est réelle positive,

$$E(X) = \sum_{k=1}^{+\infty} k\mathbb{P}(X = k) = \sum_{(k,n) \in A} u_{k,n} = +\infty = \sum_{n=1}^{+\infty} \mathbb{P}(X \geq n).$$

□

DÉFINITION 15. Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles ou complexe. Si la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable, on dit que X est d'**espérance finie**. On écrit dans ce cas $X \in L^1$. L'**espérance** de X , notée $E(X)$, est alors égale à la somme de la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$:

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

5.1.2 Propriétés de l'espérance

Théorème 16. (théorème de transfert)

Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans un ensemble non vide E et f une fonction de E dans \mathbb{R} ou \mathbb{C} . La variable $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable et dans ce cas

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

DÉMONSTRATION . Pour $y \in f(X(\Omega))$, l'événement $\{f(X) = y\}$ est la réunion disjointe des événements $\{X = x\}$ où $x \in f^{-1}(\{y\}) \subset X(\Omega) \subset E$. Pour $y \in f(X(\Omega))$, $f^{-1}(\{y\})$ est une partie de $X(\Omega)$ qui est au plus dénombrable et est donc au plus dénombrable puis

$$\mathbb{P}(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x).$$

De plus, Ω est la réunion disjointe des événements $\{f(X) = y\}$, $y \in f(X)(\Omega)$.

• Supposons la famille $(f(x) \mathbb{P}(X = x))_{x \in X(\Omega)}$ sommable. Pour $y \in f(X(\Omega))$ donné, la sous-famille $(f(x) \mathbb{P}(X = x))_{x \in f^{-1}(\{y\})}$ est sommable et

$$\sum_{x \in f^{-1}(\{y\})} f(x) \mathbb{P}(X = x) = y \sum_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x) = y \times \mathbb{P}(f(X) = y)$$

puis d'après le théorème de sommation par paquets, la famille $(y \mathbb{P}(f(X) = y))_{y \in f(X)(\Omega)}$ est sommable ou encore la variable $f(X)$ admet est d'espérance finie et de plus,

$$\begin{aligned} E(f(X)) &= \sum_{y \in f(X)(\Omega)} y \mathbb{P}(f(X) = y) = \sum_{y \in f(X)(\Omega)} \left(\sum_{x \in f^{-1}(\{y\})} f(x) \mathbb{P}(X = x) \right) \\ &= \sum_{x \in X(\Omega)} f(x) \mathbb{P}(X = x). \end{aligned}$$

• Supposons que $f(X)$ admette soit d'espérance finie ou encore supposons la famille $(y \mathbb{P}(f(X) = y))_{y \in f(X)(\Omega)}$ sommable. Il est équivalent de dire $\sum_{y \in f(X)(\Omega)} |y| \mathbb{P}(f(X) = y) < +\infty$. Montrons alors que la famille $(f(x) \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable. Puisque la famille $(|f(x)| \mathbb{P}(X = x))_{x \in X(\Omega)}$ est réelle positive, on peut regrouper par paquets et on obtient

$$\begin{aligned} \sum_{x \in X(\Omega)} |f(x)| \mathbb{P}(X = x) &= \sum_{y \in f(X)(\Omega)} \left(\sum_{x \in f^{-1}(\{y\})} |f(x)| \mathbb{P}(X = x) \right) = \sum_{y \in f(X)(\Omega)} \left(|y| \sum_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x) \right) \\ &= \sum_{y \in f(X)(\Omega)} |y| \mathbb{P}(f(X) = y) < +\infty \end{aligned}$$

Mais alors, d'après le paragraphe précédent, la variable $f(X)$ admet est d'espérance finie et de plus

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

□

Théorème 17. (linéarité de l'espérance)

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles ou complexes. Si X et Y sont d'espérance finie, alors pour tout $(\lambda, \mu) \in \mathbb{K}^2$, la variable $\lambda X + \mu Y$ est d'espérance finie et de plus $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$.

Dit autrement, l'ensemble des variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$ qui sont d'espérance finie est un \mathbb{K} -espace vectoriel et l'application $X \mapsto E(X)$ est une forme linéaire sur cet espace.

DÉMONSTRATION . Supposons que X soit d'espérance finie ou encore supposons sommable la famille $(x \mathbb{P}(X = x))_{x \in X(\Omega)}$. Soit $\lambda \in \mathbb{K}$. La famille $(\lambda x \mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable. D'après le théorème de transfert appliqué avec la fonction $f : x \mapsto \lambda x$, la variable λX admet une espérance et

$$E(\lambda X) = \sum_{x \in X(\Omega)} \lambda x \mathbb{P}(X = x) = \lambda \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) = \lambda E(X).$$

Soient X et Y deux variables discrètes d'espérance finie. Vérifions que $X + Y$ est d'espérance finie et que $E(X + Y) = E(X) + E(Y)$. (X, Y) est une variable aléatoire à valeurs dans $X(\Omega) \times Y(\Omega) \subset \mathbb{K}^2$. Soit f l'application de \mathbb{R}^2 dans \mathbb{R} définie par : $\forall (x, y) \in \mathbb{R}^2$, $f((x, y)) = x + y$. Vérifions que la famille $((x + y) \mathbb{P}((X, Y) = (x, y)))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est sommable.

$$\begin{aligned} \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} |x + y| \mathbb{P}((X, Y) = (x, y)) &\leq \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} (|x| + |y|) \mathbb{P}((X, Y) = (x, y)) \\ &= \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} |x| \mathbb{P}((X, Y) = (x, y)) + \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} |y| \mathbb{P}((X, Y) = (x, y)) \\ &= \sum_{x \in X(\Omega)} |x| \left(\sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y)) \right) + \sum_{y \in Y(\Omega)} |y| \left(\sum_{x \in X(\Omega)} \mathbb{P}((X, Y) = (x, y)) \right) \\ &= \sum_{x \in X(\Omega)} |x| \mathbb{P}(X = x) + \sum_{y \in Y(\Omega)} |y| \mathbb{P}(Y = y) = E(|X|) + E(|Y|) < +\infty \end{aligned}$$

car par hypothèse, les familles $(|x| \mathbb{P}(X = x))_{x \in X(\Omega)}$ et $(|y| \mathbb{P}(Y = y))_{y \in Y(\Omega)}$ sont sommables. Puisque la famille $((x + y) \mathbb{P}((X, Y) = (x, y)))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est sommable, le théorème de transfert permet d'affirmer que

$$E(X + Y) = E(f(X, Y)) = \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} (x + y) \mathbb{P}((X, Y) = (x, y)).$$

On a vu au passage que les familles $(x \mathbb{P}((X, Y) = (x, y)))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ et $(y \mathbb{P}((X, Y) = (x, y)))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ sont sommables. On peut donc séparer en deux sommes et on obtient

$$\begin{aligned} E(X + Y) &= \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} x \mathbb{P}((X, Y) = (x, y)) + \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} y \mathbb{P}((X, Y) = (x, y)) \\ &= \sum_{x \in X(\Omega)} x \left(\sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y)) \right) + \sum_{y \in Y(\Omega)} y \left(\sum_{x \in X(\Omega)} \mathbb{P}((X, Y) = (x, y)) \right) \\ &= \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) + \sum_{y \in Y(\Omega)} y \mathbb{P}(Y = y) = E(X) + E(Y). \end{aligned}$$

□

Théorème 18. (positivité de l'espérance, croissance de l'espérance)

- 1) a) Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles. Si X est positive, alors $E(X) \geq 0$.
b) Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles positives. $E(X) = 0$ si et seulement si $\forall x \in X(\Omega) \setminus \{0\}, \mathbb{P}(X = x) = 0$.
- 2) Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles, toutes deux d'espérance finie. Si $X \leq Y$ (c'est-à-dire si $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$), alors $E(X) \leq E(Y)$.

DÉMONSTRATION .

1) a) est immédiat (y compris dans le cas $E(X) = +\infty$).

b) S'il existe $x_0 \in X(\Omega) \setminus \{0\}$ tel que $\mathbb{P}(X = x_0) > 0$, alors $E(X) \geq x_0 \mathbb{P}(X = x_0) > 0$. Par contraposition, si $E(X) = 0$, alors $\forall x \in X(\Omega) \setminus \{0\}, \mathbb{P}(X = x) = 0$. La réciproque est claire.

2) Si $X \leq Y$, alors $Y - X \geq 0$ puis $E(Y) - E(X) = E(Y - X) \geq 0$.

□

Les deux théorèmes suivants sont immédiats à partir des propriétés des familles sommables :

Théorème 19. Soient X une variable aléatoire discrète à valeurs complexes et Y une variable aléatoire discrète à valeurs réelles sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Si $|X| \leq Y$ et si Y est d'espérance finie, alors X admet une espérance.

Théorème 20. Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles. Si X est d'espérance finie (ou encore si $E(|X|) < +\infty$), alors X admet une espérance et de plus $|E(X)| \leq E(|X|)$.

Théorème 21. (espérance d'un produit de deux variables indépendantes)

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles ou complexes, toutes deux d'espérance finie. Si X et Y sont indépendantes, alors XY est d'espérance finie et de plus $E(XY) = E(X)E(Y)$.

DÉMONSTRATION . Vérifions que la famille $(xy \mathbb{P}((X, Y) = (x, y)))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est sommable. Puisque X et Y sont indépendantes,

$$\begin{aligned} \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} |xy| \mathbb{P}((X, Y) = (x, y)) &= \sum_{(x, y) \in X(\Omega) \times Y(\Omega)} |x| \mathbb{P}(X = x) \times |y| \mathbb{P}(Y = y) \\ &= \left(\sum_{x \in X(\Omega)} |x| \mathbb{P}(X = x) \right) \left(\sum_{y \in Y(\Omega)} |y| \mathbb{P}(Y = y) \right) = E(|X|)E(|Y|) < +\infty. \end{aligned}$$

Le théorème de transfert appliqué au couple (X, Y) et à la fonction $f : (x, y) \mapsto x \times y$ montre que XY admet une espérance et que

$$\begin{aligned} E(XY) &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy \mathbb{P}((X,Y) = (x,y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} x \mathbb{P}(X=x) \times y \mathbb{P}(Y=y) \\ &= \left(\sum_{x \in X(\Omega)} x \mathbb{P}(X=x) \right) \left(\sum_{y \in Y(\Omega)} y \mathbb{P}(Y=y) \right) = E(X)E(Y). \end{aligned}$$

□

Théorème 22. Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit $A \in \mathcal{A}$ (ou encore soit A un événement). Alors $\mathbb{P}(A) = E(1_A)$.

DÉMONSTRATION. $E(1_A) = 0 \times \mathbb{P}(1_A = 0) + 1 \times \mathbb{P}(1_A = 1) = \mathbb{P}(1_A = 1) = \mathbb{P}(\{\omega \in \Omega / 1_A(\omega) = 1\}) = \mathbb{P}(A)$.

□

5.1.3 Inégalité de MARKOV

Théorème 23. Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs réelles positives, d'espérance finie.

Pour tout réel strictement positif a , $\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}$.

DÉMONSTRATION.

1ère démonstration. Soit $a > 0$.

$$\begin{aligned} E(X) &= \sum_{x \in X(\Omega)} x \mathbb{P}(X=x) = \sum_{x \in X(\Omega), x \geq a} x \mathbb{P}(X=x) + \sum_{x \in X(\Omega), x < a} x \mathbb{P}(X=x) \\ &\geq \sum_{x \in X(\Omega), x \geq a} x \mathbb{P}(X=x) \text{ (car } X \text{ est à valeurs dans } \mathbb{R}^+) \\ &\geq a \sum_{x \in X(\Omega), x \geq a} \mathbb{P}(X=x) = a \mathbb{P}(X \geq a) \end{aligned}$$

et donc, puisque $a > 0$, $\mathbb{P}(X \geq a) \leq \frac{E(X)}{a}$.

2ème démonstration. Soit $a > 0$. Soit A l'événement $\{X \geq a\}$. D'après le théorème 22,

$$\mathbb{P}(\{X \geq a\}) = \mathbb{P}(A) = E(1_A).$$

Maintenant, pour $\omega \in \Omega$, si $\omega \in A$, $a 1_A(\omega) = a \leq X(\omega)$ et si $\omega \notin A$, $a 1_A(\omega) = 0 \leq X(\omega)$. Donc, $a 1_A \leq X$ puis, puisque $a > 0$, $1_A \leq \frac{X}{a}$. Par croissance et linéarité de l'espérance,

$$\mathbb{P}(\{X \geq a\}) = E(1_A) \leq E\left(\frac{X}{a}\right) = \frac{E(X)}{a}.$$

□

5.2 Moments d'ordre r d'une variable aléatoire

DÉFINITION 16. Soient X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et r un entier naturel.

On dit que X **admet un moment d'ordre r** si et seulement si X^r est d'espérance finie. Dans ce cas, le moment d'ordre r de X est $E(X^r)$.

Quand la famille $(x^r \mathbb{P}(X=x))_{x \in X(\Omega)}$ est sommable, le théorème de transfert fournit le moment d'ordre r de X :

$$E(X^r) = \sum_{x \in X(\Omega)} x^r \mathbb{P}(X=x).$$

5.3 Variance et écart-type d'une variable aléatoire

5.3.1 Définition de la variance et de l'écart-type

On prépare la définition de la variance par quelques résultats préliminaires.

Théorème 24. Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Si X admet un moment d'ordre 2, alors X est d'espérance finie.

DÉMONSTRATION . Soit X une variable admettant un moment d'ordre 2.

Puisque $0 \leq (|X| - 1)^2 = X^2 - 2|X| + 1$, on en déduit que $|X| \leq \frac{1}{2}(X^2 + 1)$. La variable constante 1 est d'espérance finie (égale à 1) et par hypothèse, la variable X^2 est d'espérance finie puis la variable $\frac{1}{2}(X^2 + 1)$ est d'espérance finie d'après le théorème 17. Mais alors, la variable $|X|$ est d'espérance finie puis X est d'espérance finie d'après le théorème 20. □

Si X est une variable aléatoire réelles discrète admettant un moment d'ordre 2, on écrit $X \in L^2$. Donc $X \in L^2 \Leftrightarrow E(X^2) < +\infty$.

Théorème 25. (Inégalité de CAUCHY-SCHWARZ).

Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Si X et Y admettent un moment d'ordre 2, alors XY est d'espérance finie (dit autrement, si X et Y sont dans L^2 , alors XY est dans L^1) et de plus

$$(E(XY))^2 \leq E(X^2) E(Y^2).$$

DÉMONSTRATION . L'inégalité $(|X| - |Y|)^2 \geq 0$ fournit $|XY| \leq \frac{1}{2}(X^2 + Y^2)$. Puisque X^2 et Y^2 sont d'espérance finies, il en est de même de la variable $\frac{1}{2}(X^2 + Y^2)$ puis de la variable $|XY|$ et enfin de la variable XY .

Pour $\lambda \in \mathbb{R}$, on pose $P(\lambda) = E((\lambda X + Y)^2) = \lambda^2 E(X^2) + 2\lambda E(XY) + E(Y^2)$. La variable $(\lambda X + Y)^2$ est positive et donc, par positivité de l'espérance, le polynôme P est positif sur \mathbb{R} .

1er cas. Si $E(X^2) > 0$, P est un trinôme du second degré de signe constant sur \mathbb{R} . Son discriminant réduit est négatif ou nul ce qui fournit :

$$0 \geq \Delta' = (E(XY))^2 - E(X^2) E(Y^2)$$

puis $(E(XY))^2 \leq E(X^2) E(Y^2)$.

2ème cas. Si $E(X^2) = 0$, P est une fonction affine de signe constant sur \mathbb{R} . Le coefficient de λ est donc nul ce qui fournit $2E(XY) = 0$ puis $(E(XY))^2 = 0 = E(X^2) E(Y^2)$.

L'inégalité de CAUCHY-SCHWARZ est démontrée dans tous les cas. □

Théorème 26. Les variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ admettant un moment d'ordre 2 (c'est-à-dire L^2) constituent un \mathbb{R} -espace vectoriel.

DÉMONSTRATION . La variable 0 admet un moment d'ordre 2 et si X et Y sont deux variables admettant un moment d'ordre 2 et λ et μ sont deux réels, $(\lambda X + \mu Y)^2 = \lambda^2 X^2 + 2\lambda\mu XY + \mu^2 Y^2$ admet une espérance d'après le théorème 25 entre autre ou encore $\lambda X + \mu Y$ admet un moment d'ordre 2.

Par suite, l'ensemble des variables aléatoires admettant un moment d'ordre 2 est un sous-espace de l'espace des variables aléatoires et en particulier, est un espace vectoriel. □

DÉFINITION 17. Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X admet une espérance $E(X)$ et que $X - E(X)$ admet un moment d'ordre 2.

La **variance** de X est : $V(X) = E((X - E(X))^2)$ et l'**écart-type** de X est $\sigma(X) = \sqrt{V(X)}$.

Théorème 27. (formule de KOENIG-HUYGENS)

Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Si X admet un moment d'ordre 2, alors X admet une variance et de plus

$$V(X) = E(X^2) - (E(X))^2.$$

DÉMONSTRATION . D'après le théorème 24, si X admet un moment d'ordre 2, X est d'espérance finie. Mais alors, $(X - E(X))^2 = X^2 - 2E(X)X + (E(X))^2$ est d'espérance finie ou encore X admet une variance et, par linéarité de l'espérance,

$$V(X) = E(X^2) - 2E(X)E(X) + (E(X))^2 = E(X^2) - (E(X))^2.$$

□

Théorème 28. Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Si X admet un moment d'ordre 2, alors pour tout $(\lambda, \mu) \in \mathbb{R}^2$, $\lambda X + \mu$ admet une variance et $V(\lambda X + \mu) = \lambda^2 V(X)$.

DÉMONSTRATION . Puisque X admet un moment d'ordre 2, X admet une espérance. Il en est de même de $\lambda X + \mu$ et $(\lambda X + \mu) - E(\lambda X + \mu) = \lambda(X - E(X))$.

Puisque X admet un moment d'ordre 2, X admet une variance ou encore $(X - E(X))^2$ admet une espérance. Il en est de même de $\lambda^2(X - E(X))^2$ et de plus

$$V(\lambda X + \mu) = E(\lambda^2(X - E(X))^2) = \lambda^2 E((X - E(X))^2) = \lambda^2 V(X).$$

□

5.3.2 Variables centrées réduites

DÉFINITION 18. Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

X est **centrée** si et seulement si $E(X) = 0$. X est **centrée réduite** si et seulement si $E(X) = 0$ et $\sigma(X) = 1$.

Théorème 29. Soit X une variable aléatoire réelle sur un espace probabilisé au plus dénombrable $(\Omega, \mathcal{A}, \mathbb{P})$.

Si X est d'espérance finie, la variable $Y = X - E(X)$ est centrée.

Si X admet un moment d'ordre 2 et si $\sigma(X) > 0$, la variable $Y = \frac{X - E(X)}{\sigma(X)}$ est centrée réduite.

DÉMONSTRATION . $E(X - E(X)) = E(X) - E(X) = 0$. Si de plus, $\sigma(X) > 0$, $E\left(\frac{X - E(X)}{\sigma(X)}\right) = \frac{E(X) - E(X)}{\sigma(X)} = 0$ et

$$V\left(\frac{X - E(X)}{\sigma(X)}\right) = \frac{V(X)}{(\sigma(X))^2} = 1.$$

□

5.3.3 Inégalité de BIENAYMÉ-TCHEBYCHEV

Théorème 30. Soit X une variable aléatoire réelle sur un espace probabilisé au plus dénombrable $(\Omega, \mathcal{A}, \mathbb{P})$. Si X admet un moment d'ordre 2, alors

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

DÉMONSTRATION . Puisque X admet un moment d'ordre 2, X admet une espérance et une variance. D'après l'inégalité de MARKOV (théorème 23, page 19), puisque pour tout $\omega \in \Omega$, $|X(\omega) - E(X)| \geq \varepsilon \Leftrightarrow (X(\omega) - E(X))^2 \geq \varepsilon^2$,

$$\mathbb{P}(|X - E(X)| \geq \varepsilon) = \mathbb{P}((X - E(X))^2 \geq \varepsilon^2) \leq \frac{E((X - E(X))^2)}{\varepsilon^2} = \frac{V(X)}{\varepsilon^2}.$$

□

5.4 Covariance d'un couple de variables aléatoires

5.4.1 Définition et propriétés de la covariance

DÉFINITION 19. Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que X et Y admettent une espérance puis que $(X - E(X))(Y - E(Y))$ admet une espérance.

La **covariance** du couple (X, Y) est $\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$.

On note que $\text{cov}(X, X) = E((X - E(X))^2) = V(X)$.

Théorème 31. (formule de KOENIG-HUYGENS)

Soient X et Y deux variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Si X et Y admettent un moment d'ordre 2, alors le couple (X, Y) admet une covariance et de plus

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

DÉMONSTRATION . Si X et Y admettent un moment d'ordre 2, X et Y et XY admettent une espérance d'après les théorèmes 24 et 25, page 20. Mais alors, $(X - E(X))(Y - E(Y)) = XY - E(X)Y - E(Y)X + E(X)E(Y)$ admet une espérance d'après le théorème 15, page 16, et de plus

$$\begin{aligned}\text{cov}(X, Y) &= E(XY - E(X)Y - E(Y)X + E(X)E(Y)) = E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y).\end{aligned}$$

□

A partir du théorème 21, page 18, on en déduit immédiatement :

Théorème 32. Soient X et Y deux variables aléatoires réelles sur un espace probabilisé au plus dénombrable $(\Omega, \mathcal{A}, \mathbb{P})$. Si X et Y admettent un moment d'ordre 2 et si X et Y sont indépendantes, alors $\text{cov}(X, Y) = 0$.

De manière générale, quand deux variables X et Y vérifient $\text{cov}(X, Y) = 0$, on dit que les deux variables X et Y sont **non corrélées** ou **décorélées**. Le théorème précédent affirme que deux variables indépendantes sont non corrélées. On peut démontrer que la réciproque est fausse c'est-à-dire qu'il existe des couples de variables non indépendantes vérifiant $\text{cov}(X, Y) = 0$.

Théorème 33. La covariance est une forme bilinéaire, symétrique, positive sur l'espace des variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, admettant un moment d'ordre 2.

DÉMONSTRATION . On rappelle (théorème 26, page 20) que l'ensemble des variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, admettant un moment d'ordre 2, est un espace vectoriel, l'espace L^2 .

cov est bien une application de $(L^2)^2$ dans \mathbb{R} d'après le théorème 31.

Si X et Y sont deux éléments de L^2 , $\text{cov}(X, Y) = E(XY) - E(X)E(Y) = E(YX) - E(Y)E(X) = \text{cov}(Y, X)$. Donc, cov est symétrique.

Si X_1, X_2 et Y sont trois éléments de L^2 et $(\lambda_1, \lambda_2) \in \mathbb{R}^2$, par linéarité de l'espérance,

$$\begin{aligned}\text{cov}(\lambda_1 X_1 + \lambda_2 X_2, Y) &= E((\lambda_1 X_1 + \lambda_2 X_2)Y) - E(\lambda_1 X_1 + \lambda_2 X_2)E(Y) \\ &= \lambda_1 (E(X_1 Y) - E(X_1)E(Y)) + \lambda_2 (E(X_2 Y) - E(X_2)E(Y)) \\ &= \lambda_1 \text{cov}(X_1, Y) + \lambda_2 \text{cov}(X_2, Y).\end{aligned}$$

cov est une forme linéaire par rapport à sa première variable sur L^2 et donc bilinéaire par symétrie.

Enfin, pour tout X de L^2 , $\text{cov}(X, X) = V(X) \geq 0$. Donc, cov est une forme bilinéaire, symétrique, positive sur L^2 .

□

5.4.2 Variance d'une somme finie de variables

Théorème 34. Soient X_1, \dots, X_n des variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, admettant chacune un moment d'ordre 2. Alors, $X_1 + \dots + X_n$ admet une variance et de plus

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

DÉMONSTRATION . Puisque les X_i , $1 \leq i \leq n$, admettent chacune un moment d'ordre 2, chaque couple (X_i, X_j) , $(i, j) \in \llbracket 1, n \rrbracket^2$, admet une covariance. Par bilinéarité,

$$\begin{aligned} V\left(\sum_{i=1}^n X_i\right) &= \text{cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) = \sum_{1 \leq i, j \leq n} \text{cov}(X_i, X_j) \\ &= \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j). \end{aligned}$$

□

Une conséquence immédiate des théorèmes 31 et 29 est

Théorème 35. Soient X_1, \dots, X_n des variables aléatoires réelles sur un espace probabilisé au plus dénombrable $(\Omega, \mathcal{A}, \mathbb{P})$, admettant chacune un moment d'ordre 2 et deux à deux non corrélées (ce qui est le cas si les variables X_1, \dots, X_n sont deux à deux indépendantes). Alors, $X_1 + \dots + X_n$ admet une variance et de plus

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i).$$

5.5 Espérances et variances des lois usuelles

Théorème 36. (espérances et variances des lois usuelles)

- 1) Soit $n \in \mathbb{N}^*$. Si $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$, $E(X) = \frac{n+1}{2}$ et $V(X) = \frac{n^2-1}{12}$.
- 2) a) Soit $p \in]0, 1[$, Si $X \sim \mathcal{B}(p)$, alors $E(X) = p$ et $V(X) = p(1-p)$.
b) Soit $(n, p) \in \mathbb{N}^* \times]0, 1[$. Si $X \sim \mathcal{B}(n, p)$, $E(X) = np$ et $V(X) = np(1-p)$.
- 3) Soit $p \in]0, 1[$. Si $X \sim \mathcal{G}(p)$, $E(X) = \frac{1}{p}$ et $V(X) = \frac{1-p}{p^2}$.
- 4) Soit $\lambda \in]0, +\infty[$. Si $X \sim \mathcal{P}(\lambda)$, $E(X) = \lambda$ et $V(X) = \lambda$.

DÉMONSTRATION . 1) et 2) ont été démontrés en maths sup. On redonne ces démonstrations.

1) a) Soit X une variable aléatoire telle que $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$.

$$E(X) = \sum_{k=1}^n k \mathbb{P}(X=k) = \frac{1}{n} \sum_{k=1}^n k = \frac{1}{n} \times \frac{n(n+1)}{2} = \frac{n+1}{2},$$

et

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 = \sum_{k=1}^n k^2 \times \frac{1}{n} - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} \\ &= \frac{n+1}{12}(2(2n+1) - 3(n+1)) = \frac{n^2-1}{12} \end{aligned}$$

2) a) Soit X une variable aléatoire telle que $X \sim \mathcal{B}(p)$. $E(X) = 0 \times (1-p) + 1 \times p = p$ et $V(X) = (0^2 \times (1-p) + 1^2 \times p) - p^2 = p - p^2 = p(1-p)$.

b) Soit X une variable aléatoire telle que $X \sim \mathcal{B}(n, p)$. Soient X_1, \dots, X_n , n variables de BERNOULLI mutuellement indépendantes (et en particulier deux à deux indépendantes) de paramètre p . On sait que $X_1 + \dots + X_n \sim X$. On en déduit que $E(\mathcal{B}(n, p)) = \sum_{k=1}^n E(X_k) = \sum_{k=1}^n p = np$ et que $V(X) = \sum_{i=1}^n V(X_i) = \sum_{i=1}^n p(1-p) = np(1-p)$.

On peut aussi rappeler le calcul direct de $E(X)$ et $V(X)$:

$$\begin{aligned} E(X) &= \sum_{k=0}^n k \times \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} = np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} = np(p + (1-p))^{n-1} = np, \end{aligned}$$

et

$$\begin{aligned}
V(X) &= \sum_{k=0}^n k^2 \times \binom{n}{k} p^k (1-p)^{n-k} - (np)^2 = \sum_{k=2}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} + \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} - n^2 p^2 \\
&= \sum_{k=2}^n n(n-1) \binom{n-2}{k-2} p^k (1-p)^{n-k} + np - n^2 p^2 \\
&= n(n-1)p^2 \sum_{k=2}^n \binom{n-2}{k-2} p^{k-2} (1-p)^{(n-2)-(k-2)} + np - n^2 p^2 \\
&= n(n-1)p^2 + np - n^2 p^2 = np - np^2 = np(1-p).
\end{aligned}$$

3) On rappelle que pour tout réel $t \in]0, 1[$, $\frac{1}{1-t} = \sum_{k=0}^{+\infty} t^k$ puis par dérivation, pour tout réel $t \in]0, 1[$, $\frac{1}{(1-t)^2} = \sum_{k=1}^{+\infty} k t^{k-1}$ et $\frac{2}{(1-t)^3} = \sum_{k=2}^{+\infty} k(k-1) t^{k-2}$.

Soit X une variable aléatoire telle que $X \sim \mathcal{G}(p)$. Puisque $p \in]0, 1[$, on a encore $1-p \in]0, 1[$ et donc

$$E(X) = \sum_{k=1}^{+\infty} k p (1-p)^{k-1} = p \sum_{k=1}^{+\infty} k (1-p)^{k-1} = p \frac{1}{(1-(1-p))^2} = \frac{1}{p}.$$

Ensuite,

$$\begin{aligned}
V(X) &= \sum_{k=1}^{+\infty} k^2 p (1-p)^{k-1} - \frac{1}{p^2} = \sum_{k=1}^{+\infty} k(k-1) p (1-p)^{k-1} + \sum_{k=1}^{+\infty} k p (1-p)^{k-1} - \frac{1}{p^2} \\
&= p(1-p) \sum_{k=2}^{+\infty} k(k-1) (1-p)^{k-2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2p(1-p)}{(1-(1-p))^3} + \frac{1}{p} - \frac{1}{p^2} \\
&= \frac{2(1-p) + p - 1}{p^2} = \frac{1-p}{p^2}.
\end{aligned}$$

4) Soit X une variable aléatoire telle que $X \sim \mathcal{P}(\lambda)$.

$$E(X) = \sum_{k=0}^{+\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{+\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$

et

$$\begin{aligned}
V(X) &= \sum_{k=0}^{+\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 = \sum_{k=0}^{+\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \sum_{k=0}^{+\infty} k \frac{\lambda^k}{k!} e^{-\lambda} - \lambda^2 \\
&= \lambda^2 e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda - \lambda^2 = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda - \lambda^2 = \lambda.
\end{aligned}$$

□

5.6 Loi faible des grands nombres

Théorème 37. (loi faible des grands nombres)

Soient X_1, \dots, X_n , des variables aléatoires réelles discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, deux à deux indépendantes et de mêmes lois, admettant toutes un moment d'ordre 2. En posant pour $n \in \mathbb{N}^*$, $S_n = \sum_{i=1}^n X_i$ et en notant m l'espérance de X_1 (qui est encore l'espérance de chaque X_i), on a

$$\forall \varepsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) = 0.$$

DÉMONSTRATION. Soit $\varepsilon > 0$. Soit $n \in \mathbb{N}^*$. Puisque chaque X_i admet un moment d'ordre 2, chaque X_i admet une espérance et une variance. Il en est de même de $\frac{S_n}{n}$ et de plus, par linéarité de l'espérance,

$$E \left(\frac{S_n}{n} \right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n m = m$$

et d'autre part, les variables X_i étant deux à deux indépendantes,

$$V\left(\frac{S_n}{n}\right) = \frac{1}{n^2} V\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{V(X_1)}{n}.$$

D'après l'inégalité de BIENAYMÉ-TCHEBYCHEV,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geq \varepsilon\right) \leq \frac{V(S_n/n)}{\varepsilon^2} = \frac{V(X_1)}{n\varepsilon^2}.$$

Quand n tend vers $+\infty$, on obtient

$$\lim_{n \rightarrow +\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) = 0.$$

□

Interprétons ce résultat. Considérons un événement donné A de probabilité p (comme par exemple obtenir un 1 après le lancer d'un dé bien équilibré et donc $p = \frac{1}{6}$). Notons X_n , $n \in \mathbb{N}^*$, la variable égale à 1 si l'événement A est réalisé et à 0 sinon. X_n suit la loi de BERNOULLI de paramètre p .

La variable $F_n = \frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n}$ est alors la fréquence d'apparition de l'événement A au cours de n épreuves. En supposant ces épreuves indépendantes, la loi faible des grands nombres affirme que pour n grand, la probabilité que la fréquence (« probabilité empirique ») s'écarte de la probabilité théorique d'au moins ε est quasiment nulle (au cours d'un grand nombre de lancers de dés, la fréquence d'apparition du 1 doit être très proche de $\frac{1}{6}$).

6 Fonctions génératrices

6.1 Fonction génératrice d'une variable aléatoire

6.1.1 Définition

DÉFINITION 20. Soit X une variable aléatoire sur un espace probabilisé au plus dénombrable $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N} .

La **fonction génératrice** de la variable aléatoire X , notée G_X , est la série entière associée à la suite $(\mathbb{P}(X = n))_{n \in \mathbb{N}}$:

$$G_X(t) = \sum_{n=0}^{+\infty} t^n \mathbb{P}(X = n).$$

Théorème 38. G_X est définie et continue sur $[-1, 1]$ au moins.

DÉMONSTRATION . Pour $n \in \mathbb{N}$, on pose $a_n = \mathbb{P}(X = n)$. Pour tout $n \in \mathbb{N}$, $0 \leq a_n \leq 1$ et donc $R_a \geq 1$. Par suite, G_X est définie et continue sur $] -1, 1[$ au moins.

Pour $n \in \mathbb{N}$, soit $f_n : t \mapsto a_n t^n$. Pour $t \in [-1, 1]$, $|a_n t^n| = a_n |t|^n \leq a_n$ avec égalité effectivement obtenue pour $t = 1$ puis $\|f_n\|_{\infty, [-1, 1]} = a_n$. La suite (a_n) est positive et la série de terme général a_n converge (et a pour somme 1). Ceci montre que la série de fonctions de terme général f_n , $n \in \mathbb{N}$, converge normalement et en particulier uniformément sur $[-1, 1]$ vers la fonction G_X . Puisque chaque fonction f_n est définie et continue sur $[-1, 1]$, la fonction G_X est définie et continue sur $[-1, 1]$ au moins.

□

6.1.2 Dérivées et applications aux calculs des moments

Tout d'abord puisque le rayon de la série entière de somme G_X est supérieur ou égal à 1 un résultat classique sur les séries entières montre que l'on retrouve la loi de X à partir de la connaissance de sa fonction génératrice :

Théorème 39. G_X est de classe C^∞ sur $] -1, 1[$ et

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

Ensuite, on peut utiliser la fonction génératrice pour des calculs d'espérance et de variance (ou plus généralement des calculs de moments) comme le montrent les théorèmes 41 et 42. On commence par établir un résultat général sur les séries entières.

Théorème 40. Soit $(a_n)_{n \in \mathbb{N}}$ une suite réelle positive telle que la série de terme général a_n , $n \in \mathbb{N}$, converge. La fonction $f : t \mapsto \sum_{n=0}^{+\infty} a_n t^n$ est dérivable en 1 si et seulement si la série de terme général na_n , $n \in \mathbb{N}$, converge et dans ce cas,

$$f'(1) = \sum_{n=1}^{+\infty} na_n.$$

DÉMONSTRATION . Puisque la série de terme général a_n , $n \in \mathbb{N}$, converge, f est définie sur $[-1, 1]$ au moins. On note que le résultat du théorème est immédiat si $R_a > 1$ auquel cas f est de classe C^∞ sur $] -R_a, R_a[$ et pour tout t de $] -R_a, R_a[$, $f'(t) = \sum_{n=1}^{+\infty} na_n t^{n-1}$.

En particulier, f est dérivable en 1, la série de terme général na_n , $n \in \mathbb{N}$, converge et $f'(1) = \sum_{n=1}^{+\infty} na_n$. On peut donc supposer dorénavant que $R_a = 1$ et analyser la dérivabilité de f en 1 à gauche.

Pour $t \in [0, 1[$,

$$\frac{f(t) - f(1)}{t - 1} = \sum_{n=0}^{+\infty} a_n \frac{t^n - 1}{t - 1} = \sum_{n=1}^{+\infty} a_n (t^{n-1} + t^{n-2} + \dots + t + 1) \quad (*).$$

En particulier, la fonction $t \mapsto \frac{f(t) - f(1)}{t - 1}$ est positive et croissante sur $[0, 1[$ en tant que somme de fonctions croissantes sur $[0, 1[$.

On en déduit que la fonction $t \mapsto \frac{f(t) - f(1)}{t - 1}$ admet une limite en 1 à gauche dans $[0, +\infty]$.

• Supposons f dérivable en 1 à gauche. Soit $N \in \mathbb{N}^*$. Pour $t \in [0, 1[$,

$$\sum_{n=1}^N a_n (t^{n-1} + t^{n-2} + \dots + t + 1) \leq \sum_{n=1}^{+\infty} a_n (t^{n-1} + t^{n-2} + \dots + t + 1) = \frac{f(t) - f(1)}{t - 1}.$$

Quand t tend vers 1, on obtient $\sum_{n=1}^N na_n \leq f'(1)$. Mais alors, la suite des sommes partielles $\left(\sum_{n=1}^N na_n \right)_{N \in \mathbb{N}}$ est majorée par $f'(1)$.

Puisque la suite $(na_n)_{n \in \mathbb{N}}$ est positive, on en déduit que la série de terme général na_n converge et que $\sum_{n=1}^{+\infty} na_n \leq f'(1)$.

• Supposons que la série de terme général na_n converge. Pour $t \in [0, 1[$,

$$\frac{f(t) - f(1)}{t - 1} = \sum_{n=1}^{+\infty} a_n (t^{n-1} + t^{n-2} + \dots + t + 1) \leq \sum_{n=1}^{+\infty} a_n (1 + 1 + \dots + 1 + 1) = \sum_{n=1}^{+\infty} na_n < +\infty.$$

Donc, la fonction $t \mapsto \frac{f(t) - f(1)}{t - 1}$ est majorée sur $[0, 1[$. On en déduit que $\frac{f(t) - f(1)}{t - 1}$ a une limite réelle quand t tend vers 1 ou encore que f est dérivable en 1 et de plus, quand t tend vers 1, on obtient $f'(1) \leq \sum_{n=1}^{+\infty} na_n$.

En résumé, f est dérivable en 1 si et seulement si la série de terme général na_n converge et dans ce cas, $f'(1) \leq \sum_{n=1}^{+\infty} na_n$ et

$$f'(1) \geq \sum_{n=1}^{+\infty} na_n \text{ et finalement } f'(1) = \sum_{n=1}^{+\infty} na_n.$$

□

⇒ **Commentaire .** Pour l'implication « $\sum na_n$ converge $\Rightarrow G_X$ dérivable en 1 », on peut aussi utiliser le théorème de dérivation terme à terme en montrant d'abord la convergence normale de la série des dérivées $n \mapsto na_n t^{n-1}$ sur $[0, 1]$. Mais le théorème de dérivation terme à terme ne permet pas d'établir la réciproque : G_X dérivable en 1 $\Rightarrow \sum na_n$ converge.

On peut maintenant énoncer les théorèmes fournissant espérance et variance d'une variable aléatoire quand on connaît sa fonction génératrice (X désignant toujours une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{N}) :

Théorème 41. X admet une espérance si et seulement si G_X est dérivable en 1 et dans ce cas,

$$E(X) = G'_X(1).$$

DÉMONSTRATION . X est d'espérance finie si et seulement si la série numérique de terme général $(n\mathbb{P}(X = n))_{n \in \mathbb{N}}$ converge. D'après le théorème 40, appliqué à $a_n = \mathbb{P}(X = n)$, ceci équivaut à la dérivabilité de G_X en 1 et en cas de dérivabilité en 1,

$$G'_X(1) = \sum_{n=1}^{+\infty} n\mathbb{P}(X = n) = E(X).$$

□

Théorème 42. X admet un moment d'ordre 2 (et donc une espérance et une variance) si et seulement si G_X est deux fois dérivable en 1 et dans ce cas,

$$E(X^2) = G''_X(1) + G'_X(1),$$

puis

$$V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2.$$

DÉMONSTRATION . X admet un moment d'ordre 2 si et seulement si la série de terme général $n^2\mathbb{P}(X = n)$. Or, $n(n-1)\mathbb{P}(X = n) = n^2\mathbb{P}(X = n) - n\mathbb{P}(X = n)$. Donc, $n(n-1)\mathbb{P}(X = n) \underset{n \rightarrow +\infty}{=} O(n^2\mathbb{P}(X = n))$ et $n^2\mathbb{P}(X = n) \underset{n \rightarrow +\infty}{=} O(n(n-1)\mathbb{P}(X = n))$.

Ainsi, X admet un moment d'ordre 2 si et seulement si la série de terme général $n(n-1)\mathbb{P}(X = n)$.

Si la série de terme général $n(n-1)\mathbb{P}(X = n)$ converge, alors en particulier, la série de terme général $n\mathbb{P}(X = n)$ converge (car $n\mathbb{P}(X = n) \underset{n \rightarrow +\infty}{=} O(n(n-1)\mathbb{P}(X = n))$) et donc G_X est dérivable en 1 puis sur $[0, 1]$. Le théorème 40 appliqué à $a_n = n\mathbb{P}(X = n)$ et $f = G'_X$ montre alors que G'_X est dérivable en 1 ou encore G_X est deux fois dérivable en 1.

Réciproquement, si G_X est deux fois dérivable en 1, G_X est en particulier dérivable en 1 et donc la série de terme général $n\mathbb{P}(X = n)$ converge. Le théorème 40 appliqué à $a_n = n\mathbb{P}(X = n)$ et $f = G'_X$ montre que la série de terme général $n(n-1)\mathbb{P}(X = n)$ converge.

De plus, toujours d'après le théorème 40, dans le cas où G_X est deux fois dérivable en 1, on a

$$\begin{aligned} G''_X(1) &= \sum_{n=2}^{+\infty} n(n-1)\mathbb{P}(X = n) = \sum_{n=1}^{+\infty} n(n-1)\mathbb{P}(X = n) \\ &= \sum_{n=1}^{+\infty} n^2\mathbb{P}(X = n) - \sum_{n=1}^{+\infty} n\mathbb{P}(X = n) \text{ (les deux séries convergent)} \\ &= E(X^2) - E(X) \text{ (d'après le théorème de transfert)} \\ &= E(X^2) - G'_X(1), \end{aligned}$$

et donc

$$E(X^2) = G''_X(1) + G'_X(1)$$

puis

$$V(X) = E(X^2) - (E(X))^2 = G''_X(1) + G'_X(1) - (G'_X(1))^2.$$

□

6.1.3 Fonction génératrice d'une somme finie de variables indépendantes

Théorème 43.

Soient X et Y deux variables aléatoires réelles discrètes sur une espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Si X et Y sont indépendantes, alors pour tout $t \in [-1, 1]$ (au moins)

$$G_{X+Y}(t) = G_X(t) \times G_Y(t).$$

Soient X_1, \dots, X_n des variables aléatoires réelles discrètes sur une espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Si X_1, \dots, X_n sont mutuellement indépendantes, alors pour tout $t \in [-1, 1]$ (au moins)

$$G_{X_1+\dots+X_n}(t) = G_{X_1}(t) \times \dots \times G_{X_n}(t).$$

DÉMONSTRATION . Si X et Y sont indépendantes, pour $t \in [-1, 1]$,

$$\begin{aligned}
G_{X+Y}(t) &= \sum_{n=0}^{+\infty} t^n \mathbb{P}(X+Y=n) = \sum_{n=0}^{+\infty} \left(\sum_{k=0}^n \mathbb{P}(\{X=k\} \cap \{Y=n-k\}) \right) t^n \\
&= \sum_{n=0}^{+\infty} \left(\sum_{k=0}^n \mathbb{P}(X=k) \times \mathbb{P}(Y=n-k) \right) t^n \text{ (car } X \text{ et } Y \text{ sont indépendantes)} \\
&= \left(\sum_{n=0}^{+\infty} t^n \mathbb{P}(X=n) \right) \left(\sum_{n=0}^{+\infty} t^n \mathbb{P}(Y=n) \right) \text{ (produit de CAUCHY de deux séries entières)} \\
&= G_X(t) \times G_Y(t).
\end{aligned}$$

Le résultat général s'en déduit alors par récurrence à partir du fait que si X_1, \dots, X_{n+1} , sont mutuellement indépendantes, alors $X_1 + \dots + X_n$ et X_{n+1} sont indépendantes d'après le lemme des coalitions. \square

6.2 Fonctions génératrices des variables usuelles

6.2.1 Fonction génératrice d'un variable aléatoire de BERNOULLI

Soit $p \in]0, 1[$. Soit X une variable aléatoire telle que $X \sim \mathcal{B}(p)$.

$$\text{Pour } n \in \mathbb{N}, a_n = \mathbb{P}(X=n) = \begin{cases} 1-p & \text{si } n=0 \\ p & \text{si } n=1 \\ 0 & \text{si } n \geq 2 \end{cases}. \text{ Donc, } R_a = +\infty \text{ et}$$

$$\forall t \in \mathbb{R}, G_X(t) = \mathbb{P}(X=0) + t\mathbb{P}(X=1) = pt + 1 - p.$$

On retrouve alors $E(X) = G'_X(1) = p$ et $V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = 0 + p - p^2 = p(1-p)$.

6.2.2 Fonction génératrice d'un variable aléatoire suivant une loi binomiale

Soit $(n, p) \in \mathbb{N}^* \times]0, 1[$. Soit X une variable aléatoire telle que $X \sim \mathcal{B}(n, p)$.

$$\text{Pour } k \in \mathbb{N}, a_k = \mathbb{P}(X=k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } 0 \leq k \leq n \\ 0 & \text{si } k > n \end{cases}. \text{ Donc, } R_a = +\infty \text{ et pour } t \in \mathbb{R},$$

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

On peut aussi appliquer le théorème 43 en écrivant $X \sim X_1 + \dots + X_n$ où X_1, \dots, X_n , sont n variables de BERNOULLI mutuellement indépendantes de même paramètre p :

$$G_X(t) = \prod_{k=1}^n G_{X_i}(t) = \prod_{k=1}^n (pt + 1 - p) = (pt + 1 - p)^n.$$

On retrouve alors $E(X) = G'_X(1) = np$ et $V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = np(1-p)$.

6.2.3 Fonction génératrice d'un variable aléatoire suivant une loi géométrique

Soit $p \in]0, 1[$. Soit X une variable aléatoire telle que $X \sim \mathcal{G}(p)$. Pour $n \in \mathbb{N}$, $a_n = \mathbb{P}(X=n) = \begin{cases} 0 & \text{si } n=0 \\ p(1-p)^{n-1} & \text{si } n \geq 1 \end{cases}$.

Donc, $R_a = \frac{1}{1-p}$ (avec $\frac{1}{1-p} > 1$) et pour $t \in \left] -\frac{1}{1-p}, \frac{1}{1-p} \right[$,

$$G_X(t) = \sum_{n=1}^{+\infty} t^n p(1-p)^{n-1} = pt \sum_{k=1}^{+\infty} (t(1-p))^{k-1} = \frac{pt}{1 - (1-p)t}.$$

Retrouvons alors les expressions de l'espérance et de la variance du théorème 36, page 23. Pour $t \in \left] -\frac{1}{1-p}, \frac{1}{1-p} \right[$,

$$G'_X(t) = p \frac{(1 - (1-p)t) - t(-(1-p))}{(1 - (1-p)t)^2} = \frac{p}{(1 - (1-p)t)^2}$$

et donc

$$E(X) = G'_X(1) = \frac{p}{(1 - (1 - p))^2} = \frac{1}{p}.$$

Ensuite, pour $t \in \left] -\frac{1}{1-p}, \frac{1}{1-p} \right[$, $G''_X(t) = \frac{p \times (-2)(-(1-p))}{(1 - (1-p)t)^3}$ et donc $G''_X(1) = \frac{2(1-p)}{p^2}$ puis

$$E(X^2) = G''_X(1) + G'_X(1) = \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2-p}{p^2}$$

et enfin

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

6.2.4 Fonction génératrice d'un variable aléatoire suivant une loi de POISSON

Soit $\lambda \in \mathbb{R}$. Soit X une variable aléatoire telle que $X \sim \mathcal{P}(\lambda)$. Pour $n \in \mathbb{N}$, $a_n = \mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}$. Donc, $R_a = +\infty$ et pour $t \in \mathbb{R}$,

$$G_X(t) = \sum_{k=0}^{+\infty} t^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda t)^k}{k!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}.$$

Retrouvons alors les expressions de l'espérance et de la variance du théorème 36. Pour $t \in \mathbb{R}$, $G'_X(t) = \lambda e^{\lambda(t-1)}$ et $G''_X(t) = \lambda^2 e^{\lambda(t-1)}$. Donc $E(X) = G'_X(1) = \lambda$ et $V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

On résume les différentes fonctions génératrices dans un tableau :

Si $X \sim \mathcal{B}(p)$, $p \in]0, 1[$, pour tout $t \in \mathbb{R}$, $G_X(t) = pt + 1 - p$.

Si $X \sim \mathcal{B}(n, p)$, $(n, p) \in \mathbb{N}^* \times]0, 1[$, pour tout $t \in \mathbb{R}$, $G_X(t) = (pt + 1 - p)^n$.

Si $X \sim \mathcal{G}(p)$, $p \in]0, 1[$, pour tout $t \in \left] -\frac{1}{1-p}, \frac{1}{1-p} \right[$, $G_X(t) = \frac{pt}{1 - (1-p)t}$.

Si $X \sim \mathcal{P}(\lambda)$, $\lambda \in]0, +\infty[$, pour tout $t \in \mathbb{R}$, $G_X(t) = e^{\lambda(t-1)}$.

Ainsi, si X_1, \dots, X_n , sont n variables mutuellement indépendantes suivant des lois de POISSON de paramètres respectifs $\lambda_1, \dots, \lambda_n$, alors la fonction génératrice de $X = X_1 + \dots + X_n$ est $G_X : t \mapsto \prod_{k=1}^n G_{X_k}(t) = e^{(\lambda_1 + \dots + \lambda_n)t}$ et donc X suit la loi de POISSON de paramètre $\lambda_1 + \dots + \lambda_n$.