Mathématiques

PT

Lycée Couffignal

CHAPITRE 15

Variables aléatoires discrètes

Plan du chapitre

Introduction

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

Dans ce chapitre, nous allons étendre la notion de variable aléatoire réelle au cadre des univers infinis. Nous verrons notamment qu'une variable aléatoire n'admet plus nécessairement une espérance et une variance dans ce nouveau contexte.

Partie I Variables aléatoires discrètes

I.A - Généralités

Rappelons que l'on dit qu'un ensemble est au plus dénombrable s'il est fini ou dénombrable.

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

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

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

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

Remarque 1 : L'ensemble $X(\Omega)$ est l'ensemble des valeurs prises par la variable aléatoire X.

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

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

De plus, X est une variable aléatoire réelle sur $(\Omega, \mathscr{A})$, alors pour tout $x \in \mathbb{R}$, on note

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

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

Proposition 1 : Si $X : \Omega \to E$ est une variable aléatoire discrète sur un espace probabilisable $(\Omega, \mathscr{A})$ et si $U \subset E$, alors l'ensemble $(X \in U)$ est un évènement, i.e. $(X \in U)$ est un élément de $\mathscr{A}$.

Remarque 2 : On en déduit que les six ensembles ci-dessus sont des évènements de $(\Omega, \mathscr{A})$.

Exemple 1 : On lance un dé cubique rouge et un dé cubique bleu. Un univers permettant de modéliser cette expérience aléatoire est $\Omega = [\![1, 6]\!]^2$. Comme l'univers est fini, on peut considérer la tribu $\mathscr{A} = \mathscr{P}(\Omega)$.

Les applications définies sur Ω par $X : (k, \ell) \mapsto k$ et $Y : (k, \ell) \mapsto \ell$ sont des variables aléatoires sur $(\Omega, \mathscr{A})$. Elles correspondent respectivement au résultat du dé rouge et au résultat du dé bleu. On a $X(\Omega) = Y(\Omega) = [\![1, 6]\!]$.

Proposition 2 : Soit X une variable aléatoire discrète sur un espace probabilisable $(\Omega, \mathscr{A})$. Si f est une fonction définie sur $X(\Omega)$, alors l'application $f \circ X$ est une variable aléatoire discrète sur $(\Omega, \mathscr{A})$.

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

Remarque 3 : Plus généralement, toute fonction d'un nombre fini de variables aléatoires discrètes sur $(\Omega, \mathscr{A})$ est encore une variable aléatoire discrète sur $(\Omega, \mathscr{A})$.

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

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

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

I.B - Loi d'une variable aléatoire discrète

Pour toute variable aléatoire discrète $X : \Omega \to E$ définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, on peut définir l'application $P_X : \mathscr{P}(X(\Omega)) \to \mathbb{R}$ par $P_X : A \mapsto P_X(A) = P(X \in A)$.

Définition (Loi d'une variable aléatoire discrète) : Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. L'application P_X s'appelle la loi de la variable aléatoire X.

Remarque 4 : On peut vérifier que P_X est une probabilité sur l'espace probabilisable $\big(X(\Omega), \mathscr{P}(X(\Omega))\big)$.

Proposition 3 : Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. La loi P_X est déterminée par la famille des nombres $P(X = x)$ pour $x \in X(\Omega)$.

Remarques 5 :

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

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

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

Exemple 2 : Si X désigne le résultat du lancer d'un dé cubique équilibré, alors la loi de X est donnée par

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

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

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

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

Exemples 3 :

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

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

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

Proposition 4 : Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. Si f est une application définie sur $X(\Omega)$ et si X et Y suivent la même loi, alors $f(X)$ et $f(Y)$ suivent la même loi.

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

Remarque 6 : D'après la proposition 3, on en déduit que pour donner la loi conditionnelle d'une variable aléatoire discrète X sachant un évènement B, il suffit de spécifier l'ensemble $X(\Omega)$ et de calculer le nombre $P_B(X = x)$ pour tout $x \in X(\Omega)$.

Exemple 4 : On lance un dé cubique équilibrée et on note X le résultat.

- La variables aléatoire X suit une loi uniforme sur l'ensemble $[\![1, 6]\!]$.
- On note B l'évènement « le résultat du dé est pair ». Déterminons la loi conditionnelle de X sachant l'évènement B. Pour tout $k \in \{2, 4, 6\}$, on a

$$P_B(X = k) = \frac{P\big((X = k) \cap B\big)}{P(B)} = \frac{P(X = k)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}$$

tandis que pour tout $k \in \{1, 3, 5\}$, on a

$$P_B(X = k) = \frac{P\big((X = k) \cap B\big)}{P(B)} = \frac{P(\varnothing)}{P(B)} = 0.$$

I.C - Couples de variables aléatoires discrètes

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

Remarque 7 : En reprenant les notations ci-dessus, l'application $Z = (X, Y) : \Omega \to E \times F$ est une variable aléatoire.

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

Remarque 8 : D'après la proposition 3, on en déduit que pour donner la loi d'un couple de variables aléatoires discrète (X, Y), il suffit de spécifier les ensembles $X(\Omega)$ et $Y(\Omega)$, puis de calculer le nombre

$$P\big((X, Y) = (x, y)\big) = P\big((X = x) \cap (Y = y)\big)$$

pour tout $x \in X(\Omega)$ et tout $y \in Y(\Omega)$.

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

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

Remarque 9 : La loi conjointe de deux variables aléatoires discrètes détermine entièrement les lois marginales. En effet, par la formule des probabilités totales, on a les relations

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

Exemple 5 : On considère le couple de variables aléatoires (X, Y) dont la loi est donnée par $X(\Omega) = Y(\Omega) = \mathbb{N}$ et

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

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

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

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

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

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

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

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

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

Remarque 10 : Toutes les définitions précédentes s'étendent naturellement aux n-uplets de variables aléatoires.

I.D - Indépendance de variables aléatoires discrètes

Dans cette partie, on étend les différentes notions d'indépendance de variables aléatoires vues en première année dans le cadre des univers finis.

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

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

Remarque 11 : Dans le cas où X et Y sont indépendantes, la loi du couple (X, Y) est déterminée par ses lois marginales.

Exemple 6 : On lance un dé cubique rouge et un dé cubique bleu. Les variables aléatoires X et Y donnant respectivement le résultat du dé rouge et le résultat du dé bleu sont indépendantes. La probabilité que les deux dés donnent un résultat d'au moins 5 est

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

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

Proposition 5 : Deux variables aléatoires discrètes X et Y sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ sont indépendantes si et seulement si

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

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

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

Proposition 6 : Soit $n \in \mathbb{N}^*$. Des variables aléatoires discrètes $X_1, \ldots, X_n$ sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ sont indépendantes si et seulement si

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

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

Remarque 12 : Le résultat ci-dessus se généralise si l'on considère $k \in \mathbb{N}^*$ fonctions $f_1, \ldots, f_k$ à la place de f et g.

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

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

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

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

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

Exemple 8 : Pour modéliser une succession infinie de lancers d'une pièce, on considère une suite i.i.d. de variables aléatoires de Bernoulli.

Pour terminer, rappelons le résultat suivant de première année.

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

Partie II Moments d'une variable aléatoire discrète réelle

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

II.A - Espérance d'une variable aléatoire discrète réelle

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

Soit $X : \Omega \to E$ est une variable aléatoire discrète sur $(\Omega, \mathscr{A}, P)$. Par définition d'une variable aléatoire discrète, l'ensemble $X(\Omega)$ est une partie de E au plus dénombrable : il est fini ou dénombrable.

(i) Si l'ensemble $X(\Omega)$ est fini de cardinal $n \in \mathbb{N}^*$, on peut écrire $X(\Omega) = \{x_1, \dots, x_n\}$ où les éléments $x_1, \dots, x_n \in E$ sont deux à deux distincts.

(ii) Si l'ensemble $X(\Omega)$ est dénombrable, on peut écrire $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ où $(x_n)_{n\in\mathbb{N}}$ est une suite injective d'éléments de E.

Dans les deux cas, on dit que l'écriture précédente est une énumération de l'ensemble $X(\Omega)$.

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

(i) Si l'ensemble $X(\Omega)$ est fini, il admet une énumération de la forme $X(\Omega) = \{x_1, \dots, x_n\}$.
On dit que la variable aléatoire X est d'espérance finie et on appelle espérance de X le nombre réel

$$E(X) = \sum_{k=1}^{n} x_k P(X = x_k).$$

(ii) Si $X(\Omega)$ est dénombrable, il admet une énumération de la forme $X(\Omega) = \{x_n \in \mathbb{R} \mid n \in \mathbb{N}\}$.
On dit que la variable aléatoire X est d'espérance finie si la série

$$\sum_{n\geqslant 0} x_n P(X = x_n)$$

converge absolument. Dans ce cas, on appelle espérance de X et on note $\mathrm{E}(X)$ le nombre réel

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

Remarques 13 :

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

b) Si $X(\Omega)$ est fini, alors l'espérance ne dépend pas du choix de l'énumération de l'ensemble $X(\Omega)$, car toute énumération de $X(\Omega)$ est une permutation des éléments de $x_1, \dots, x_n$ et la somme est commutative.

c) Si $X(\Omega)$ est dénombrable et si la série de la définition ne converge pas absolument, on dit que la variable aléatoire X n'est pas d'espérance finie.

d) Si $X(\Omega)$ est dénombrable, on impose dans la définition que la série converge absolument, car dans ce cas, on peut vérifier que la somme de la série ne dépend pas du choix de l'énumération de $X(\Omega)$.

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

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

Exemples 9 :

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

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

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

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

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

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

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

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

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

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

alors X est d'espérance finie, car la série

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

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

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

Proposition 7 : Soit X une variable aléatoire réelle sur $(\Omega, \mathscr{A}, P)$. Si X est à valeurs dans $\mathbb{N}$ et si X est d'espérance finie, alors la série $\sum\limits_{n\geqslant 1} P(X \geqslant n)$ converge et on a

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

Exemple 10 : En reprenant l'exemple 9.c, on a avec la somme de la série géométrique que

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

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

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

Théorème du transfert : Soient X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$ et f une fonction définie sur $X(\Omega)$ à valeurs réelles.

(i) Si l'ensemble $X(\Omega)$ est fini, il admet une énumération de la forme $X(\Omega) = \{x_1, \dots, x_n\}$.
La variable aléatoire $f(X)$ est d'espérance finie et on a

$$E\big(f(X)\big) = \sum_{k=1}^{n} f(x_k) P(X = x_k).$$

(ii) Si $X(\Omega)$ est dénombrable, il admet une énumération de la forme $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$.
La variable aléatoire $f(X)$ est d'espérance finie si et seulement si la série

$$\sum_{n \geqslant 0} f(x_n) P(X = x_n)$$

converge absolument. Dans ce cas, on a

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

DÉMONSTRATION ADMISE

Remarque 14 : En particulier, le théorème s'applique au cas où X est un n-uplet de variables aléatoires.
Par exemple, en considérant la fonction $f : (x, y) \longmapsto xy$, si X et Y sont des variables aléatoires discrètes réelles et si on a des énumérations $X(\Omega) = \{x_m \in \mathbb{R} \mid m \in \mathbb{N}\}$ et $Y(\Omega) = \{y_n \in \mathbb{R} \mid n \in \mathbb{N}\}$, alors la variable aléatoire XY est d'espérance finie si et seulement si la série double

$$\sum_{m=0}^{+\infty} \left(\sum_{n=0}^{+\infty} |x_m y_n| P\big((X = x_m) \cap (Y = y_n)\big) \right)$$

converge et dans ce cas, on a

$$\mathrm{E}(XY) = \sum_{m=0}^{+\infty} \left(\sum_{n=0}^{+\infty} x_m y_n P\big((X = x_m) \cap (Y = y_n)\big) \right).$$

Exemples 11 :

a) En reprenant l'exemple 9.a, comme $X(\Omega)$ est fini, la variable aléatoire X^2 est d'espérance finie par le théorème du transfert et on a

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

b) En reprenant l'exemple 9.c, on trouve que X^2 est d'espérance finie par le théorème du transfert, car la série

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

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

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

Proposition 8 : Soient X et Y deux variables aléatoires réelles d'espérance finie sur $(\Omega, \mathscr{A}, P)$.

(i) Linéarité de l'espérance : pour tout $(\lambda, \mu) \in \mathbb{R}^2$, la variable aléatoire $\lambda X + \mu Y$ est d'espérance finie et on a

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

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

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

DÉMONSTRATION ADMISE

Proposition 9 : Soient X et Y deux variables aléatoires réelles sur $(\Omega, \mathscr{A}, P)$.

(i) Si $|X| \leqslant Y$ et si Y est d'espérance finie, alors X est d'espérance finie.

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

DÉMONSTRATION ADMISE

Proposition 10 : Soient X et Y deux variables aléatoires réelles d'espérance finie sur $(\Omega, \mathscr{A}, P)$.
Si X et Y sont indépendantes, alors la variable aléatoire XY est d'espérance finie et on a $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$.

DÉMONSTRATION ADMISE

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

Proposition 11 (Inégalité de Cauchy-Schwarz) : Soient X et Y deux variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$. Si X^2 et Y^2 sont d'espérance finie, alors XY est d'espérance finie et on a

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

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

Remarque 15 : En particulier, en prenant $Y = 1$, on en déduit que si X^2 est d'espérance finie, alors X est d'espérance finie.

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

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

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

Remarques 16 :

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

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

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

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

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

Théorème de Koenig-Huyghens : Soit X une variable aléatoire réelle sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.
Si X est de variance finie, alors on a
$$\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2.$$

Exemple 12 : En reprenant l'exemple 9.c, on obtient que $\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 = 3 - 1^2 = 2$.

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

Remarque 17 : Si X est une variable aléatoire de variance finie, alors la variable aléatoire $\dfrac{X - \mathrm{E}(X)}{\sigma(X)}$ est centrée et réduite.

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

II.C - Covariance de deux variables aléatoires réelles

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

Remarques 18 :

a) La covariance existe d'après l'inégalité de Cauchy-Schwarz.
b) Deux variables aléatoires dont la covariance est nulle sont dites décorrélées.

Remarque 19 : On peut interpréter la covariance de la façon suivante.

a) Si $\mathrm{Cov}(X, Y) > 0$, alors Y a tendance à augmenter lorsque X augmente.
b) Si $\mathrm{Cov}(X, Y) < 0$, alors Y a tendance à diminuer lorsque X augmente.
c) Si $\mathrm{Cov}(X, Y) = 0$, alors il n'y a pas de lien entre les variations des variables aléatoires X et Y.

Théorème 2 : Soient X et Y deux variables aléatoires réelles sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$.
Si X et Y sont de variance finie, alors on a
$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y).$$

Proposition 13 : Soient X et Y deux variables aléatoires réelles sur $(\Omega, \mathscr{A}, P)$. Si X et Y sont de variance finie, alors pour tout $(a, b) \in \mathbb{R}^2$, la variable aléatoire $aX + bY$ est de variance finie finie et on a
$$\mathrm{V}(aX + bY) = a^2\,\mathrm{V}(X) + 2ab\,\mathrm{Cov}(X, Y) + b^2\,\mathrm{V}(Y).$$

Théorème 3 : Soient X et Y deux variables aléatoires réelles sur $(\Omega, \mathscr{A}, P)$. Si X et Y sont des variables aléatoires indépendantes de variance finie, alors

(i) $\mathrm{Cov}(X, Y) = 0$,
(ii) $\mathrm{V}(X + Y) = \mathrm{V}(X) + \mathrm{V}(Y)$.

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

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

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

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

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

II.D - Coefficient de corrélation linéaire ($\star$)

Définition (Coefficient de corrélation linéaire) : Soient X et Y deux variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, P)$. Si X et Y sont de variance finie non nulle, on définit le coefficient de corrélation linéaire de X et Y comme le nombre réel

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}.$$

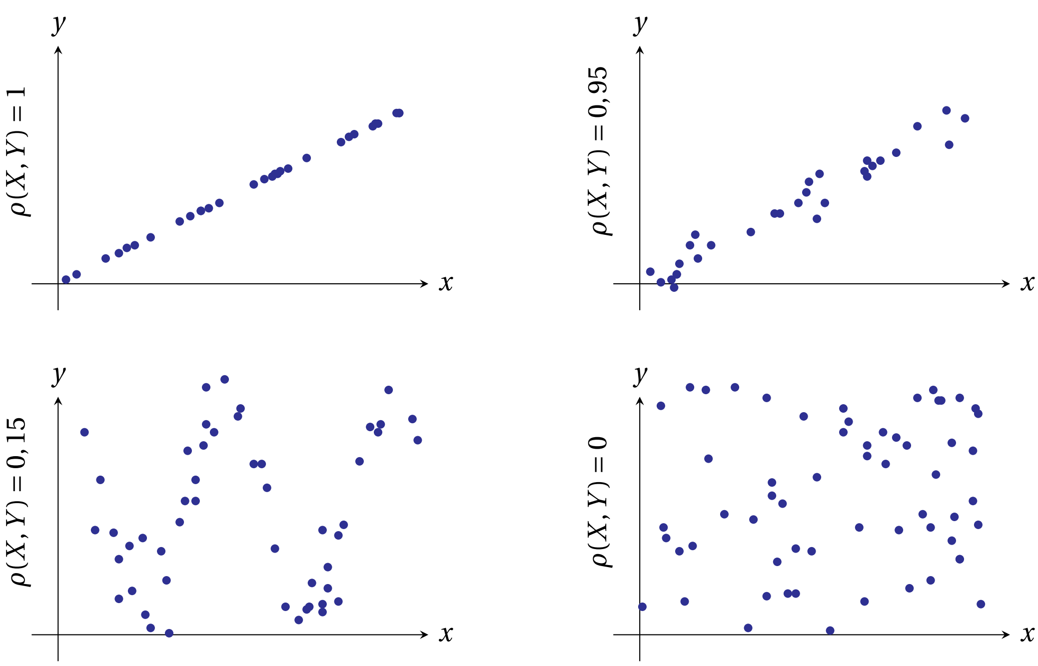
Proposition 14 : En reprenant les hypothèses de la définition précédente, on a les propriétés suivantes.

(i) Si X et Y sont indépendantes, alors $\rho(X, Y) = 0$.

(ii) On a $|\rho(X, Y)| \leqslant 1$.

(iii) On a $|\rho(X, Y)| = 1$ si et seulement si il existe $(a, b) \in \mathbb{R}^2$ tel que $Y = aX + b$ presque surement.

Illustration : Avec un ordinateur, on simule de nombreuses fois un couple de variables aléatoires (X, Y) et on place les points obtenus dans le plan. Voici les résultats des simulations de quatre expériences différentes.



Remarque 20 : On peut interpréter le coefficient de corrélation linéaire de la façon suivante.

a) S'il vaut 1 ou -1, les variables aléatoires sont linéairement liées.

b) Plus il est proche de 1 ou de -1, plus les variables aléatoires sont linéairement corrélées.

c) Lorsqu'il est nul, il n'y a pas de corrélation linéaire entre les deux variables aléatoires.

II.E - Fonctions génératrices d'une variable aléatoire à valeurs entières

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

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

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

Remarques 21 :

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

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

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

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

Exemples 13 :

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

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

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

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

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

Remarque 22 : Cette propriété est une conséquence directe de la relation

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

Théorème 4 : Soit X une variable aléatoire définie sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$.
La variable aléatoire X est d'espérance finie si et seulement si la fonction G_X est dérivable en 1.
Dans ce cas, on a $\mathrm{E}(X) = G_X'(1)$.

Remarque 23 : Admettons que si la fonction G_X est deux fois dérivables en 1, alors on a les égalités

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

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

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

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

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

Exemple 14 : Si X suit une loi binomiale de paramètre $(n,p) \in \mathbb{N} \times [0,1]$, alors sa fonction génératrice G_X est polynomiale, donc elle est de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$. De plus, on a

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

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

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

Théorème 5 : Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$.
Si les variables aléatoires X et Y sont indépendantes, alors pour tout $t \in]-1,1[$, on a $G_{X+Y}(t) = G_X(t)G_Y(t)$.

Exemple 15 : Soient X et Y deux variables aléatoires indépendantes suivant des lois binomiales de paramètres respectifs (n_1, p) et (n_2, p) avec $(n_1, n_2, p) \in \mathbb{N}^2 \times [0,1]$. Déterminons la loi de $Z = X+Y$.
Comme les variables X et Y sont indépendantes, on a d'après le théorème ci-dessus que

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

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

Remarque 24 : Le théorème se généralise directement à plusieurs variables aléatoires. Si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes définies sur $(\Omega, \mathscr{A}, P)$ à valeurs dans $\mathbb{N}$, alors on a

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

Partie III Lois usuelles

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

III.A - Loi géométrique

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

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

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

Remarques 25 :

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

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

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

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

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

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

Exemple 16 : On lance un dé équilibré autant de fois que nécessaire jusqu'à obtenir un 6. Si on note X le nombre de lancers effectués, alors X suit une loi géométrique de paramètre $1/6$.

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

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

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

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

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

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

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

Remarque 26 : On dit que la loi géométrique est sans mémoire.

III.B - Loi de Poisson

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

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

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

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

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

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

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

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

Corollaire 2 ($\star$) : Si X et Y sont deux variables indépendantes définies sur $(\Omega, \mathscr{A}, P)$ suivant des lois de Poisson de paramètres respectifs $\lambda \in \mathbb{R}_+^*$ et $\mu \in \mathbb{R}_+^*$, alors $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.

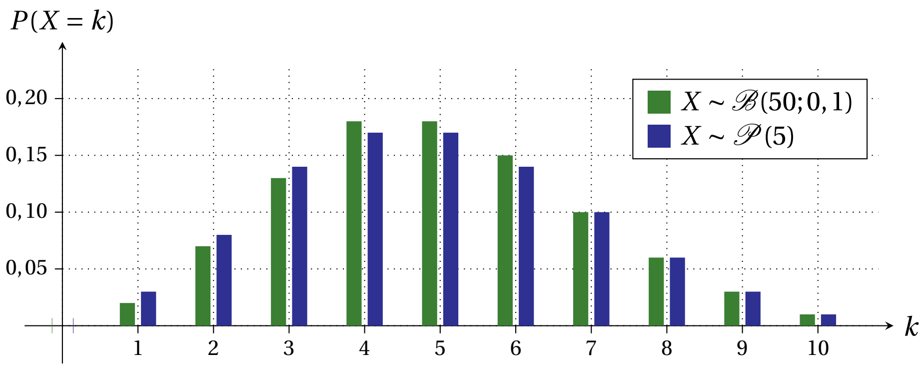
Remarque 27 : Le résultat précédent se généralise aisément : la somme de $n \in \mathbb{N}^*$ variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs $\lambda_1, \ldots, \lambda_n$ est une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda_1 + \cdots + \lambda_n$.

Théorème d'approximation de la loi binomiale par la loi de Poisson ($\star$) : Si $(S_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires telle que $S_n \sim \mathscr{B}(n, p_n)$ pour tout $n \in \mathbb{N}$ et que $\lim\limits_{n\to+\infty} np_n = \lambda$, alors on a

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P(S_n = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}.$$

Remarques 28 :

a) On en déduit que l'on peut approcher numériquement toute loi binomiale de paramètre $(n, p) \in \mathbb{N} \times]0,1[$ par une loi de Poisson de paramètre $\lambda = np$. En pratique, cette approximation n'est satisfaisante que si l'entier naturel n est suffisamment grand et si le nombre $np(1-p)$ est suffisamment petit.



b) Si des événements indépendants ont une très faible probabilité de réalisation, alors le nombre des événements se réalisant suit approximativement une loi de Poisson. Ainsi, on qualifie souvent la loi de Poisson de loi des événements rares.

Partie IV Inégalités probabilistes

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

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

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

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

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

Remarque 29 : On retrouve que la variance est un indicateur de dispersion : plus la variance de X est faible, plus les valeurs de X sont concentrées avec une probabilité élevée autour de son espérance $\mathrm{E}(X)$.

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

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

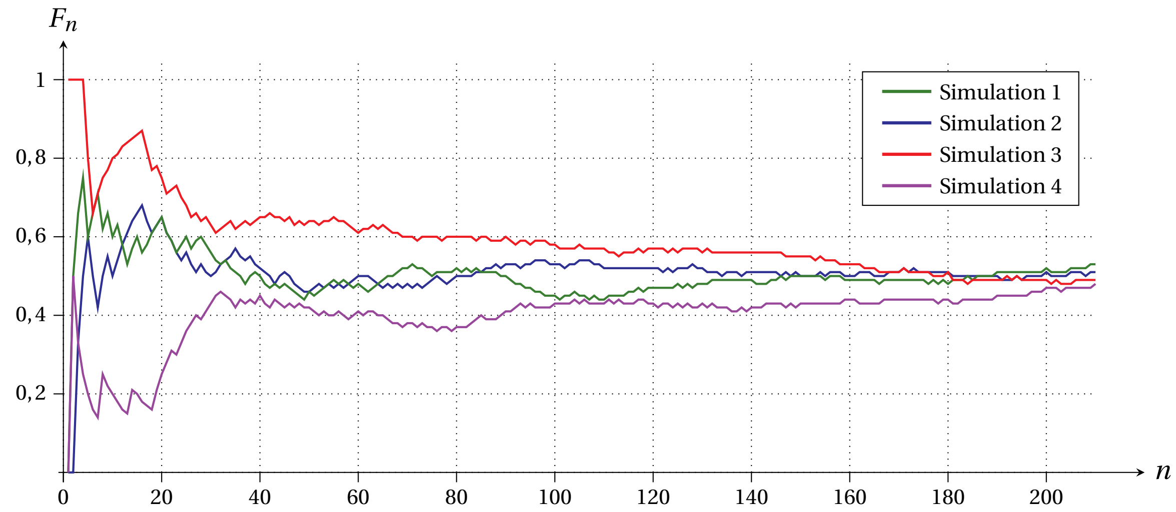
Reformulons cette définition avec notre modèle : on considère une suite de variable aléatoire $(X_n)_{n\in\mathbb{N}^*}$ définie par

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

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

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

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



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

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

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

Le théorème suivant formalise les observations et intuitions précédentes.

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

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

Remarques 30 :

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

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

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

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

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

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

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

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

Partie V Synthèse des lois usuelles

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