12 | Moments d'une variable aléatoire

« Ce calcul délicat s'étend aux questions les plus importantes de la vie, qui ne sont en effet, pour la plupart, que des problèmes de probabilité. »
Pierre-Simon, marquis de Laplace (1812)

Plan de cours

Dans tout ce chapitre, $(\Omega, \mathscr{A}, \mathbf{P})$ désignera un espace probabilisé quelconque et X une variable aléatoire – sauf spécification contraire – à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$.

I | Espérance d'une variable aléatoire complexe

En MPSI, on définit l'espérance d'une variable aléatoire réelle X à support fini comme la moyenne des valeurs prises par X, pondérées par leur probabilité d'apparition :

$$\mathbf{E}(X) = \sum_{x\in X(\Omega)} x \cdot \mathbf{P}(X=x) = \sum_{k=1}^{n} x_k \cdot \mathbf{P}(X=x_k) \quad \text{où} \quad X(\Omega) = \{x_k\}_{1\leqslant k\leqslant n}$$

L'espérance s'interprète naturellement dans les jeux de hasard comme le gain espéré, d'où la dénomination.

Il n'est guère difficile d'étendre la définition précédente aux variables aléatoires discrètes à valeurs réelles ou complexes. Encore faut-il pouvoir donner un sens à $\displaystyle\sum_{x\in X(\Omega)} x \cdot \mathbf{P}(X=x)$ et s'assurer que l'espérance ne dépend pas de l'ordre d'indexation choisi pour $X(\Omega)$. D'où la définition suivante.

Définition 12.1 : Espérance mathématique

Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{C}$.
On dit que X est d'espérance finie, ou bien qu'elle admet une espérance, lorsque la famille $(x\mathbf{P}(X=x))_{x\in X(\Omega)}$ est sommable. Dans ce cas, on appelle espérance de X le nombre complexe défini par :

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

On note $L^1(\Omega, \mathbf{P})$, ou plus simplement L^1, l'ensemble des variables aléatoires d'espérance finie.

Cette définition appelle plusieurs remarques :

- L'espérance de X, lorsqu'elle existe, ne dépend que de la loi de X.
- Toute variable aléatoire à support fini admet bien entendu une espérance.
- Lorsque $X(\Omega) = \{x_n\}_{n\in\mathbb{N}}$, X est d'espérance finie ssi la série $\displaystyle\sum_{n\in\mathbb{N}} x_n \cdot \mathbf{P}(X=x_n)$ converge absolument.
  Lorsque X est à valeurs réelles positives, on se contentera en pratique de montrer que $\mathbf{E}(X) < +\infty$.
- X est d'espérance finie si et seulement $|X|$ est d'espérance finie.

Lorsque $\mathbf{E}(X) = 0$, on dira que la variable X est *centrée*.

Exemples

- Toute variable aléatoire bornée admet une espérance. En effet, si $|X| \leqslant a$ avec $a \in \mathbb{R}_+$, alors, pour tout $x \in X(\Omega)$, $|x\mathbf{P}(X=x)| \leqslant a\mathbf{P}(X=x)$. Donc $|\mathbf{E}(X)| \leqslant \mathbf{E}(|X|) \leqslant a \sum_{x \in X(\Omega)} \mathbf{P}(X=x) = a$.
- Si A est un événement, $\mathbb{1}_A$ est d'espérance finie et $\mathbf{E}(\mathbb{1}_A) = \sum_{x \in \{0,1\}} x\mathbf{P}(\mathbb{1}_A = x) = \sum_{\omega \in \Omega} \mathbf{P}(\omega \in A) = \mathbf{P}(A)$.

Exercice 1

Dans un concours de sauts, la probabilité qu'un sauteur passe la n^{e} barre est $1/n$ et est indépendante des sauts précédents. On note X la dernière barre que le sauteur a franchi avant d'échouer et A_n l'événement « le sauteur a franchi n^{e} barre ».

1. Déterminer la probabilité de l'événement « le sauteur a franchi une infinité de barres ».
2. Calculer $\mathbf{P}(X \geqslant n)$ pour tout $n \in \mathbb{N}$ puis déterminer la loi de X.
3. En déduire l'espérance de X.

Toute variable aléatoire suivant une des lois usuelles figurant au programme est d'espérance finie.

Théorème 12.2 : Espérances usuelles (♡)

Soient X une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbf{P})$, $n \in \mathbb{N}^*$, $p \in]0,1]$ et $\lambda \in \mathbb{R}_+^*$.

(i) Si $X \sim \mathscr{U}(\{x_1, \dots, x_n\})$, alors $\mathbf{E}(X) = \dfrac{1}{n}\sum_{k=1}^{n} x_k$. En particulier, si $E = [\![1,n]\!]$, $\mathbf{E}(X) = \dfrac{n+1}{2}$.

(ii) Si $X \sim \mathscr{B}(n,p)$, alors $\mathbf{E}(X) = np$. En particulier, si $X \sim \mathscr{B}(p)$, $\mathbf{E}(X) = p$.

(iii) Si $X \sim \mathscr{G}(p)$, alors $X \in L^1$ et $\mathbf{E}(X) = \dfrac{1}{p}$.

(iv) Si $X \sim \mathscr{P}(\lambda)$, alors $X \in L^1$ et $\mathbf{E}(X) = \lambda$.

Démonstration

Pour les deux dernières lois, on s'assure de la sommabilité d'une famille de réels positifs pour justifier l'existence de l'espérance.

(ii) Si $X \sim \mathscr{B}(p)$, $\mathbf{E}(X) = 0 \times \mathbf{P}(X=0) + 1 \times \mathbf{P}(X=1) = p$. Comme $\mathbb{1}_A \sim \mathscr{B}(\mathbf{P}(A))$, on retrouve $\mathbf{E}(\mathbb{1}_A) = \mathbf{P}(A)$. Supposons maintenant que $X \sim \mathscr{B}(n,p)$. Grâce aux formules du binôme de Newton et du capitaine,

$$\mathbf{E}(X) = \sum_{k=1}^{n} k\binom{n}{k}p^k q^{n-k} = n\sum_{k=1}^{n}\binom{n-1}{k-1}p^k q^{n-k} = np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k q^{n-1-k} = np(p+q)^{n-1} = np$$

Nous retrouverons plus rapidement ce résultat au moyen de la linéarité de l'espérance.

(iii) Supposons que $X \sim \mathscr{G}(p)$. $X(\Omega) = \mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$, $\mathbf{P}(X=n) = pq^{n-1}$.
La série entière $\sum_{n \in \mathbb{N}} q^n$ et la série dérivée $\sum_{n \in \mathbb{N}^*} nq^{n-1}$ convergent absolument sur $]-1,1[$ d'après le théorème de dérivation terme à terme. En outre, $\sum_{n=1}^{+\infty} nq^{n-1} = \dfrac{1}{(1-q)^2}$. Ainsi,

$$\mathbf{E}(X) = \sum_{n=1}^{+\infty} n\mathbf{P}(X=n) = \sum_{n=1}^{+\infty} npq^{n-1} = \frac{p}{(1-q)^2} = \frac{1}{p} < +\infty$$

En interprétant X comme le temps d'attente du premier succès, ce résultat ne surprend pas : il nous faut en moyenne $1/p$ tirages pour espérer l'apparition du premier succès.

(iv) Supposons que $X \sim \mathscr{P}(\lambda)$. $X(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$, $\mathbf{P}(X=n) = \mathrm{e}^{-\lambda}\dfrac{\lambda^n}{n!}$.

$$\mathbf{E}(X) = \sum_{n=1}^{+\infty} n\mathbf{P}(X=n) = \lambda\mathrm{e}^{-\lambda}\sum_{n=1}^{+\infty}\frac{\lambda^{n-1}}{(n-1)!} = \lambda < +\infty$$

■

Lorsque la variable X est à valeurs positives entières, on dispose d'une autre expression de l'espérance qui peut dans certains cas simplifier les calculs.

Proposition 12.3

Soit X une variable aléatoire à valeurs dans $\mathbb{N}\cup\{+\infty\}$. Alors :

$$\mathbf{E}(X)=\sum_{n=0}^{+\infty}\mathbf{P}(X>n)=\sum_{n=1}^{+\infty}\mathbf{P}(X\geqslant n)\quad(\text{égalité dans }[0,+\infty])$$

Démonstration

La variable X étant à valeurs dans $\mathbb{N}\cup\{+\infty\}$, si $k\in\mathbb{N}$, $\mathbf{P}(X\geqslant k)=\sum_{n=k}^{+\infty}\mathbf{P}(X=n)$. D'où l'égalité dans $\overline{\mathbb{R}_+}$:

$$\mathbf{E}(X)=\sum_{n=0}^{+\infty}n\mathbf{P}(X=n)=\sum_{n=0}^{+\infty}\sum_{k=1}^{n}\mathbf{P}(X=n)=\sum_{k=1}^{+\infty}\sum_{n=k}^{+\infty}\mathbf{P}(X=n)=\sum_{k=1}^{+\infty}\mathbf{P}(X\geqslant k)$$

Par sommation par paquets, si la série $\sum_{k\in\mathbb{N}^*}\mathbf{P}(X\geqslant k)$ converge, $X\in L^1$ et $\mathbf{E}(X)=\sum_{n=1}^{+\infty}\mathbf{P}(X\geqslant n)$. ∎

Exemple

Appliquons ce résultat à la loi géométrique. Si $X\sim\mathcal{G}(p)$ avec $p\in]0,1]$, $\mathbf{P}(X\geqslant n)=q^{n-1}$. Or $q\in[0,1[$, donc $\sum q^{n-1}$ converge et $X\in L^1$. De plus, $\mathbf{E}(X)=\sum_{n=1}^{+\infty}\mathbf{P}(X\geqslant n)=\sum_{n=1}^{+\infty}q^{n-1}=\dfrac{1}{1-q}=\dfrac{1}{p}$.

Le calcul de l'espérance d'une variable aléatoire repose sur la connaissance *a priori* de sa loi, loi qui peut être très délicate à déterminer. Le résultat suivant permet de calculer l'espérance de $f(X)$ en s'appuyant seulement sur la loi de X.

Théorème 12.4 : Formule de transfert

Soient X une variable aléatoire discrète sur l'espace probabilisé $(\Omega,\mathscr{A},\mathbf{P})$ et $f:X(\Omega)\to\mathbb{C}$.
$f(X)$ est d'espérance finie si et seulement si la famille $\big(f(x)\mathbf{P}(X=x)\big)_{x\in X(\Omega)}$ est sommable. Dans ce cas,

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

Ce théorème précise les conditions d'existence d'une espérance (finie) pour $f(X)$ et constitue un outil formidable pour la calculer.

Démonstration

Justifions la sommabilité de la famille à la condition (nécessaire et suffisante) que $f(X)\in L^1$.
Posons $Y=|f(X)|$. Pour $y\in Y(\Omega)$, on introduit $I_y=\{x\in X(\Omega)\mid y=|f(x)|\}$. $X(\Omega)=\bigsqcup_{y\in Y(\Omega)}I_y$ et,

$$\begin{aligned}\sum_{x\in X(\Omega)}|f(x)|\mathbf{P}(X=x)&=\sum_{y\in Y(\Omega)}\sum_{x\in I_y}|f(x)|\mathbf{P}(X=x)=\sum_{y\in Y(\Omega)}y\sum_{x\in I_y}\mathbf{P}(X=x)\\&=\sum_{y\in Y(\Omega)}y\mathbf{P}(X\in I_y)=\sum_{y\in Y(\Omega)}y\mathbf{P}(Y=y)\end{aligned}$$

D'après le théorème de sommation par paquets, $f(X)$ est d'espérance finie si et seulement si la famille $\big(f(x)\mathbf{P}(X=x)\big)_{x\in X(\Omega)}$ est sommable. En cas de sommabilité, $\mathbf{E}(f(X))=\sum_{x\in X(\Omega)}f(x)\mathbf{P}(X=x)$. ∎

Exercice 2

Soient X une variable aléatoire et A un événement. Calculer de deux manières $\mathbf{E}(\mathbb{1}_A(X))$.

On observera dans les hypothèses de la formule de transfert, qu'aucune spécification n'est faite sur la nature de $X(\Omega)$ (autre que sa dénombrabilité). Souvent, $X(\Omega)\subset\mathbb{C}$. Mais on peut aussi choisir de considérer deux variables discrètes complexes X_1, X_2 et poser $X=(X_1,X_2)$. $X(\Omega)\subset\mathbb{C}^2$, si bien que la formule de transfert devient :

$$\mathbf{E}(f(X_1,X_2))=\sum_{(x_1,x_2)\in(X_1,X_2)(\Omega)} f(x_1,x_2)\cdot\mathbf{P}(X_1=x_1,X_2=x_2)\quad\text{(sous réserve de sommabilité)}$$

Exemple

On considère une urne contenant n boules numérotées de 1 à n. On tire successivement deux boules avec remise. On note X le numéro de la $1^{\text{ère}}$ boule, Y le numéro de la 2^{nde}. Déterminons $\mathbf{E}(Z)$, où $Z=\max(X,Y)$. $X(\Omega)=Y(\Omega)=Z(\Omega)=[\![1,n]\!]$; $X, Y\sim\mathscr{U}([\![1,n]\!])$ et sont indépendantes. De plus,

$$\begin{aligned}\mathbf{E}(Z)&=\sum_{1\leqslant i,j\leqslant n}\max(i,j)\mathbf{P}\big(X=i,Y=j\big)=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\max(i,j)=\frac{1}{n^2}\sum_{i=1}^{n}\left[\sum_{j=1}^{i} i+\sum_{j=i+1}^{n} j\right]\\&=\frac{1}{n^2}\sum_{i=1}^{n}\left[i^2+\frac{n(n+1)}{2}-\frac{i(i+1)}{2}\right]=\frac{1}{2n^2}\sum_{i=1}^{n}\left[n(n+1)+i^2-i\right]=\frac{(n+1)(4n-1)}{6n}\end{aligned}$$

Que se passe-t-il pour un tirage sans remise ?

Théorème 12.5 : Propriétés de l'espérance (1)

Soient X et Y deux variables aléatoires discrètes d'espérances finies sur $(\Omega,\mathscr{A},\mathbf{P})$.
Les propriétés (ii) et (iii) sont énoncées pour des variables supposées à valeurs réelles.

(i) *Linéarité de l'espérance*
Pour tout $\lambda,\mu\in\mathbb{C}$, $\lambda X+\mu Y$ est d'espérance finie et $\mathbf{E}(\lambda X+\mu Y)=\lambda\mathbf{E}(X)+\mu\mathbf{E}(Y)$.

(ii) *Positivité de l'espérance*
Si $X\geqslant 0$, c'est-à-dire si $X(\omega)\geqslant 0$ pour tout $\omega\in\Omega$, alors $\mathbf{E}(X)\geqslant 0$.

(iii) *Croissance de l'espérance*
Si $X\leqslant Y$, c'est-à-dire si $X(\omega)\leqslant Y(\omega)$ pour tout $\omega\in\Omega$, alors $\mathbf{E}(X)\leqslant\mathbf{E}(Y)$.

Démonstration

(i) La linéarité découle du théorème de transfert en considérant $f:(x,y)\mapsto\lambda x+\mu y$.

$$\begin{aligned}\mathbf{E}(\lambda X+\mu Y)&\underset{\text{transfert}}{=}\sum_{(x,y)\in(X,Y)(\Omega)}(\lambda x+\mu y)\cdot\mathbf{P}\big(X=x,Y=y\big)\\&\underset{\text{sommabilité}}{=}\lambda\sum_{(x,y)\in(X,Y)(\Omega)} x\cdot\mathbf{P}\big(X=x,Y=y\big)+\mu\sum_{(x,y)\in(X,Y)(\Omega)} y\cdot\mathbf{P}\big(X=x,Y=y\big)\\&\underset{\text{probas totales}}{=}\lambda\sum_{x\in X(\Omega)} x\cdot\mathbf{P}(X=x)+\mu\sum_{y\in Y(\Omega)} y\cdot\mathbf{P}\big(Y=y\big)=\lambda\mathbf{E}(X)+\mu\mathbf{E}(Y)\end{aligned}$$

d'après la formule des probabilités totales et par sommabilité des familles $(x\mathbf{P}(X=x))$ et $(y\mathbf{P}(Y=y))$.

(ii) $\mathbf{E}(X)=\sum_{x\in X(\Omega)} x\cdot\mathbf{P}(X=x)\geqslant 0$.

(iii) $Y-X$ est une variable positive, donc par linéarité, $\mathbf{E}(Y-X)=\mathbf{E}(Y)-\mathbf{E}(X)\geqslant 0$. ■

Quelques remarques supplémentaires :

- Nous venons en fait de prouver que L^1 est un $\mathbb{C}$-espace vectoriel et que $\mathbf{E}$ est une forme linéaire sur L^1.
- Si la variable positive X est d'espérance nulle, la positivité des termes de la somme $\sum_{x\in X(\Omega)} x\cdot\mathbf{P}(X=x)=0$ conduit à $\mathbf{P}(X=x)=0$ pour tout $x\neq 0$. Les variables positives d'espérance nulle sont donc les variables presque sûrement nulles.
- La variable $X-\mathbf{E}(X)$ est *centrée* par simple linéarité de l'espérance.

La linéarité est un principe-clé pour déterminer l'espérance d'une variable aléatoire.

Exemple (loi binomiale)

Soient $X_1, \dots, X_n \sim \mathscr{B}(p)$ supposées indépendantes. On sait que $X = X_1 + \cdots + X_n \sim \mathscr{B}(n,p)$.

Mais surtout, $\mathbf{E}(X) = \sum_{i=1}^{n} \mathbf{E}(X_i) = np$ par linéarité de l'espérance.

Exercice 3

Que vaut en moyenne la somme de deux nombres pris au hasard entre 1 et n ?

Exercice 4 (formule du crible)

Soient $A_1, \dots, A_n$ des événements. Montrer que :

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{\substack{I \subset [\![1,n]\!] \\ \operatorname{card}(I)=k}} \mathbf{P}\left(\bigcap_{i\in I} A_i\right)$$

Rajoutons deux propriétés dont les preuves sont immédiates.

Théorème 12.6 : Propriétés de l'espérance (2)

Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{C}$.

(i) *Inégalité triangulaire*
Si $|X|$ est d'espérance finie, X également et $|\mathbf{E}(X)| \leqslant \mathbf{E}(|X|)$.

(ii) *Règle de comparaison*
Si Y est d'espérance finie et $|X| \leqslant Y$, alors X est d'espérance finie.

Étendons un dernier résultat classique de première année aux variables aléatoires discrètes.

Théorème 12.7 : Espérance et indépendance

Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{C}$, d'espérance finie et indépendantes.
Alors, XY est d'espérance finie et $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

Démonstration

La preuve repose une nouvelle fois sur la formule de transfert, appliquée à la fonction $f : (x,y) \mapsto xy$.
Sous réserve de sommabilité, par indépendance,

$$\mathbf{E}(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy \cdot \mathbf{P}\big(X = x, Y = y\big) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy \cdot \mathbf{P}(X = x)\mathbf{P}(Y = y)$$
$$= \left(\sum_{x\in X(\Omega)} x \cdot \mathbf{P}(X = x)\right) \cdot \left(\sum_{y\in Y(\Omega)} y \cdot \mathbf{P}(Y = y)\right) = \mathbf{E}(X)\mathbf{E}(Y)$$

La sommabilité n'offre aucune difficulté. ■

Attention, l'indépendance de deux variables $X, Y \in L^1$ n'est pas pour autant caractérisée par la relation $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$, comme le montre le contre-exemple suivant.

Exemple

Considérons une variable aléatoire X, de loi donnée par :

x_i	-2	-1	0	1	2
$\mathbf{P}(X = x_i)$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{6}$

Par symétrie, $\mathbf{E}(X) = \mathbf{E}(X^3) = 0$ donc $\mathbf{E}(X^3) = \mathbf{E}(X)\mathbf{E}(X^2)$. Pourtant, il est clair que X et X^2 ne sont pas indépendantes.

II | Variance, covariance et écart-type d'une variable aléatoire réelle

A – Quelques considérations sur les moments d'ordre quelconque (HP)

Définition 12.8 : Moment d'ordre p

Soient X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, à valeurs dans $\mathbb{C}$ et $p \in \mathbb{N}^*$.
On dit que X admet un moment d'ordre p si X^p est d'espérance finie, c'est-à-dire si $\mathbf{E}(|X|^p) < +\infty$. On appelle alors moment d'ordre p le nombre complexe $\mathbf{E}(X^p)$.
On note $L^p(\Omega, \mathbf{P})$, ou L^p, l'ensemble des variables aléatoires admettant un moment d'ordre p.

Proposition 12.9

Pour tout $p \in \mathbb{N}^*$, $L^p(\Omega, \mathbf{P})$ est un $\mathbb{C}$-espace vectoriel. De plus, si $q \in \mathbb{N}^*$ vérifie $p \leqslant q$, $L^p(\Omega, \mathbf{P}) \subset L^q(\Omega, \mathbf{P})$.

Démonstration

Soient $p \in \mathbb{N}^*$ et X, Y deux variables admettant un moment d'ordre p. Remarquons déjà que par convexité,

$$\forall x, y \geqslant 0, \quad \left(\frac{x}{2} + \frac{y}{2}\right)^p \leqslant \frac{x^p + y^p}{2} \quad \text{donc} \quad (x+y)^p \leqslant 2^{p-1}(x^p + y^p)$$

$|X+Y|^p \leqslant 2^{p-1}(|X|^p + |Y|^p) \in L^1$ donc $X+Y$ admet un moment d'ordre p. C'est à peu près tout ce qu'il nous fallait pour prouver que $L^p(\Omega, \mathbf{P})$ possède une structure d'espace vectoriel.
Soient maintenant un entier q supérieur ou égal à p et $X \in L^q(\Omega, \mathbf{P})$. Si $x \in \mathbb{R}_+$, de deux choses l'une : soit $x \leqslant 1$ et alors $x^p \leqslant 1$, soit $x > 1$ et alors $x^p \leqslant x^q$. Ainsi, pour tout $x \geqslant 0$, $x^p \leqslant \max(1, x^q) \leqslant 1 + x^q$. $|X|^p \leqslant 1 + |X|^q \in L^1$ donc $X \in L^p(\Omega, \mathbf{P})$. ■

Corollaire 12.10

Toute variable admettant un moment d'ordre 2 est d'espérance finie.

On retrouve ce résultat en exploitant directement l'inégalité $|X| \leqslant \dfrac{X^2+1}{2}$ et la condition $X^2 \in L^1$.

B – Variance

Dans toute la suite, on ne considérera plus que des variables aléatoires réelles.

Théorème / Définition 12.11 : Variance

Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors, $(X - \mathbf{E}(X))^2$ est d'espérance finie. On appelle variance de X et on note $\mathbf{V}(X)$ le réel positif :

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

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

Démonstration

Soit $X \in L^2$. Puisque $X \in L^1$, $(X - \mathbf{E}(X))^2 = X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2 \in L^1$.
Enfin, l'écart type est bien défini par positivité de l'espérance : $\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2) \geqslant 0$. ■

La variance, qui n'est rien d'autre que la moyenne du carrée de la distance entre les valeurs de X et la moyenne de X, mesure la dispersion de X par rapport à son espérance. Il en va de même pour l'écart type qui a l'avantage d'être homogène à X. Les variables aléatoires de variance nulle sont les variables presque sûrement constantes :

$$\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2) = 0 \quad \Longrightarrow \quad X = \mathbf{E}(X) \ \text{p.s.}$$

Proposition 12.12 : Formule de Kœnig-Huygens

Si la variable aléatoire discrète X admet un moment d'ordre 2, $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.

Démonstration

Par linéarité, $\mathbf{V}(X) = \mathbf{E}((X - \mathbf{E}(X))^2) = \mathbf{E}(X^2 - 2\mathbf{E}(X)X + \mathbf{E}(X)^2) = \mathbf{E}(X^2) - 2\mathbf{E}(X)^2 + \mathbf{E}(X)^2 = \mathbf{E}(X^2) - \mathbf{E}(X)^2$. ■

Théorème 12.13 : Variances usuelles (♡)

Soient X une variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbf{P})$, $n \in \mathbb{N}^*$, $p \in]0,1]$ et $\lambda \in \mathbb{R}_+^*$.

(i) Si $X \sim \mathscr{U}([\![1,n]\!])$, alors $\mathbf{V}(X) = \dfrac{n^2-1}{12}$.

(ii) Si $X \sim \mathscr{B}(n,p)$, alors $\mathbf{V}(X) = npq$. En particulier, si $X \sim \mathscr{B}(p)$, $\mathbf{V}(X) = pq$.

(iii) Si $X \sim \mathscr{G}(p)$, alors $\mathbf{V}(X) = \dfrac{q}{p^2}$.

(iv) Si $X \sim \mathscr{P}(\lambda)$, alors $\mathbf{V}(X) = \lambda$.

Démonstration

Pour les deux dernières lois, on s'assure de la sommabilité pour justifier l'existence d'un moment d'ordre 2.

(i) Supposons que $X \sim \mathscr{U}([\![1,n]\!])$.

$$\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \sum_{k=1}^{n} k^2 \mathbf{P}(X=k) - \mathbf{E}(X)^2 = \sum_{k=1}^{n} \frac{k^2}{n} - \mathbf{E}(X)^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

(ii) Si $X \sim \mathscr{B}(p)$, $X^2 \sim \mathscr{B}(p)$ donc $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = p - p^2 = pq$.
Supposons maintenant que $X \sim \mathscr{B}(n,p)$. Nous allons calculer $\mathbf{E}(X(X-1))$ plutôt que $\mathbf{E}(X^2)$, en utilisant la formule de transfert :

$$\begin{aligned}\mathbf{E}(X(X-1)) &= \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k q^{n-k} = \sum_{k=2}^{n} k(k-1)\binom{n}{k} p^k q^{n-k} = n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2} p^k q^{n-k} \\ &= n(n-1)p^2 \sum_{k=0}^{n-2}\binom{n-2}{k} p^k q^{n-2-k} = n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2\end{aligned}$$

On a alors $\mathbf{E}(X(X-1)) = \mathbf{E}(X^2) - \mathbf{E}(X)$ par linéarité, et donc :

$$\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2 = n(n-1)p^2 + np - (np)^2 = npq$$

Nous retrouverons plus rapidement ce résultat dans le prochain paragraphe.

(iii) Supposons que $X \sim \mathscr{G}(p)$. $X(\Omega) = \mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$, $\mathbf{P}(X=n) = pq^{n-1}$.
Les séries entières $\sum_{n\in\mathbb{N}} q^n$, $\sum_{n\in\mathbb{N}^*} nq^{n-1}$ et $\sum_{n\geqslant 2} n(n-1)q^{n-2}$ convergent absolument sur $]-1,1[$ d'après le théorème de dérivation terme à terme. En outre, $\sum_{n=2}^{+\infty} n(n-1)q^{n-2} = \dfrac{2}{(1-q)^3}$. Ainsi, par transfert,

$$\mathbf{E}(X(X-1)) = \sum_{n=1}^{+\infty} n(n-1)\mathbf{P}(X=n) = \sum_{n=2}^{+\infty} n(n-1)pq^{n-1} = \frac{2pq}{(1-q)^3} = \frac{2q}{p^2} < +\infty$$

D'où $\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2 = \dfrac{2q}{p^2} + \dfrac{1}{p} - \dfrac{1}{p^2} = \dfrac{q}{p^2}$.

(iv) Supposons que $X \sim \mathscr{P}(\lambda)$. $X(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$, $\mathbf{P}(X=n) = e^{-\lambda}\dfrac{\lambda^n}{n!}$. Par transfert,

$$\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2 = \lambda^2 e^{-\lambda} \sum_{n=2}^{+\infty} \frac{\lambda^{n-2}}{(n-2)!} + \lambda - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

■

Proposition 12.14

Soient $\lambda, \mu \in \mathbb{R}$. Alors, $\mathbf{V}(\lambda X + \mu) = \lambda^2 \mathbf{V}(X)$.

Démonstration

$\mathbf{V}(\lambda X+\mu) = \mathbf{E}((\lambda X+\mu)^2) - \mathbf{E}(\lambda X+\mu)^2 = \mathbf{E}(\lambda^2 X^2 + 2\lambda\mu X + \mu^2) - (\lambda \mathbf{E}(X)+\mu)^2 = \lambda^2(\mathbf{E}(X^2) - \mathbf{E}(X)^2) = \lambda^2 \mathbf{V}(X)$

Exercice 5

Montrer que si $X \sim \mathscr{U}([\![a, b]\!])$, $\mathbf{V}(X) = \dfrac{(b-a)(b-a+2)}{12}$.

Lorsque $\sigma(X) = 1$, on dira que la variable X est *réduite*. Si $\mathbf{V}(X) \neq 0$, $\dfrac{X - \mathbf{E}(X)}{\sigma(X)}$ est *centrée réduite* :

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

C – Covariance

Remarquons que si $X, Y \in L^2$, alors $XY \in L^1$ grâce à l'inégalité $|XY| \leqslant \dfrac{X^2+Y^2}{2}$. On définit alors la covariance de X et Y comme l'espérance du produit des variables centrées.

Théorème / Définition 12.15

Soient X, Y deux variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{R}$. On suppose que X et Y admettent un moment d'ordre 2. Alors, la variable aléatoire $(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))$ admet une espérance. On appelle alors covariance de X et Y le réel noté $\operatorname{cov}(X, Y)$ défini par :

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

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

Théorème 12.16 : Formule de Kœnig-Huygens

Sous les hypothèses précédentes, $\operatorname{cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X) \cdot \mathbf{E}(Y)$.

Démonstration

Par linéarité de l'espérance,

$$\begin{aligned} \operatorname{cov}(X, Y) &= \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))] = \mathbf{E}[XY - \mathbf{E}(X)Y - \mathbf{E}(Y)X + \mathbf{E}(X)\mathbf{E}(Y)] \\ &= \mathbf{E}(XY) - 2\mathbf{E}(X)\mathbf{E}(Y) + \mathbf{E}(X)\mathbf{E}(Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) \end{aligned}$$

■

Si X et Y sont indépendantes, $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ donc $\operatorname{cov}(X, Y) = 0$. En revanche, deux variables peuvent être décorrélées sans être indépendantes ; nous avons déjà rencontré un contre-exemple.

Exercice 6

Exemple de l'urne avec 4 boules, $Z = (X_1, Y)$ avec $Y = \max(X_1, X_2)$. Que vaut $\operatorname{cov}(X_1, Y)$?

La covariance est une forme bilinéaire symétrique positive (mais non définie).

Proposition 12.17 : Propriétés de la covariance

Si X, Y et Z sont trois variables aléatoires admettant un moment d'ordre 2 et $a, b, c, d \in \mathbb{R}$, alors :

(i) $\operatorname{cov}(X, X) = \mathbf{V}(X)$;

(ii) $\operatorname{cov}(X, Y) = \operatorname{cov}(Y, X)$;

(iii) $\operatorname{cov}(aX + bY, Z) = a\operatorname{cov}(X, Z) + b\operatorname{cov}(Y, Z)$;

(iv) $\operatorname{cov}(aX + b, cY + d) = ac\operatorname{cov}(X, Y)$.

Proposition 12.18 : Variance d'une somme

Soient $X_1, \dots, X_n$ des variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, à valeurs dans $\mathbb{R}$, admettant un moment d'ordre 2. Alors,

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

En particulier, si les variables sont décorrélées, $\mathbf{V}(X_1 + \cdots + X_n) = \displaystyle\sum_{i=1}^{n} \mathbf{V}(X_i)$.

Le résultat s'établit par récurrence et repose sur la forme simplifiée suivante :

Proposition 12.19

Si X et Y sont deux variables aléatoires admettant un moment d'ordre 2 et $a, b \in \mathbb{R}$, alors :

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

Pour $a = b = 1$, $\mathbf{V}(X + Y) = \mathbf{V}(X) + \mathbf{V}(Y) + 2\operatorname{cov}(X, Y)$.

Démonstration

$$\begin{aligned}\mathbf{V}(aX + bY) &= \operatorname{cov}(aX + bY, aX + bY) = a\operatorname{cov}(X, aX + bY) + b\operatorname{cov}(Y, aX + bY) \\ &= a^2 \operatorname{cov}(X, X) + b^2 \operatorname{cov}(Y, Y) + 2ab\operatorname{cov}(X, Y) = a^2 \mathbf{V}(X) + b^2 \mathbf{V}(Y) + 2ab\operatorname{cov}(X, Y)\end{aligned}$$

■

Exemple (loi binomiale)

Soient $X_1, \dots, X_n \sim \mathscr{B}(p)$ supposées indépendantes. On sait que $X = X_1 + \cdots + X_n \sim \mathscr{B}(n, p)$.

Mais surtout, comme les variables sont décorrélées, $\mathbf{V}(X) = \displaystyle\sum_{i=1}^{n} \mathbf{V}(X_i) = npq$ par linéarité de l'espérance.

D – Coefficient de corrélation linéaire (HP)

Dans ce paragraphe, seule l'inégalité de Cauchy-Schwarz figure au programme. La définition du coefficient de corrélation linéaire en découle directement.

Proposition 12.20 : Inégalité de Cauchy-Schwarz

Soient X et Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et admettant un moment d'ordre 2. Alors, XY est d'espérance finie et :

$$|\mathbf{E}(XY)| \leqslant \sqrt{\mathbf{E}(X^2)}\sqrt{\mathbf{E}(Y^2)}$$

Il y a égalité si et seulement si X et Y sont proportionnelles presque sûrement.

Démonstration

Soient $X, Y \in L^2$ et $\lambda \in \mathbb{R}$. Alors, $\mathbf{E}((\lambda X + Y)^2) = \lambda^2 \mathbf{E}(X^2) + 2\lambda \mathbf{E}(XY) + \mathbf{E}(Y^2) \geqslant 0$. Ce trinôme en λ de signe constant a donc un discriminant $\Delta = 4(\mathbf{E}(XY)^2 - \mathbf{E}(X^2)\mathbf{E}(Y^2))$ strictement négatif. D'où l'inégalité.
De plus, en cas d'égalité, $\Delta = 0$ donc il existe $\lambda_0 \in \mathbb{R}$ tel que $\mathbf{E}((\lambda_0 X + Y)^2) = 0$. La variable $\lambda_0 X + Y$ est donc nulle presque sûrement. En d'autres termes, $Y = -\lambda_0 X$ presque sûrement. ■

En appliquant ce résultat aux deux variables centrées $X - \mathbf{E}(X)$ et $Y - \mathbf{E}(Y)$, pour peu qu'elles admettent un moment d'ordre 2, il vient :

$$|\operatorname{cov}(X, Y)| \leqslant \sqrt{\mathbf{V}(X)}\sqrt{\mathbf{V}(Y)} = \sigma(X) \cdot \sigma(Y)$$

Définition 12.21

Soient X et Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et admettant un moment d'ordre 2. On suppose de plus leurs écarts types non nuls.

On appelle coefficient de corrélation linéaire de X et Y le réel $\rho(X,Y) = \dfrac{\operatorname{cov}(X,Y)}{\sigma(X)\cdot\sigma(Y)}$.

La grandeur $\rho(X,Y)$ est sans unité. On déduit de l'inégalité de Cauchy-Schwarz probabiliste les propriétés :

Proposition 12.22

Si $X, Y \in L^2$ sont d'écart types non nuls, alors :

(i) $\rho(X,Y) \in [-1,1]$ (ii) $\rho(X,Y) = \pm 1 \iff \exists a, b \in \mathbb{R},\ Y = aX + b$ p. s.

III | Fonctions génératrices

Dans toute cette partie, on ne considérera que des variables aléatoires à valeurs dans $\mathbb{N}$.

Définition 12.23 : Fonction génératrice

La fonction génératrice de la variable X est définie par :

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

On appelle série génératrice la série entière associée. Cette série converge pour $t=1$ et $G_X(1) = \sum_{n=0}^{+\infty} \mathbf{P}(X=n) = 1$ puisque la famille $((X=n)_{n\in\mathbb{N}})$ est un système complet d'événements. D'où le résultat suivant.

Proposition 12.24

Le rayon de convergence de la série entière $\sum_{n\in\mathbb{N}} \mathbf{P}(X=n)t^n$ est supérieur ou égal à 1.

En fait, pour tout $t \in D_f(0,1)$, $|\mathbf{P}(X=n)t^n| \leqslant \mathbf{P}(X=n)$ donc la série converge normalement sur $D_f(0,1)$. Rajoutons que lorsque la variable aléatoire est finie, la fonction génératrice est polynomiale et $R = +\infty$.

Nom	**Notation**	$\boldsymbol{G(t)}$
Bernoulli	$\mathscr{B}(p)$	$q+pt$
Binomiale	$\mathscr{B}(n,p)$	$(q+pt)^n$
Uniforme	$\mathscr{U}([\![1;n]\!])$	$\dfrac{t}{n}\cdot\dfrac{1-t^n}{1-t}$
Géométrique	$\mathscr{G}(p)$	$\dfrac{pt}{1-qt}$
Poisson	$\mathscr{P}(\lambda)$	$\mathrm{e}^{\lambda(t-1)}$

FONCTIONS GÉNÉRATRICES DES LOIS USUELLES

L'application G_X est de classe $\mathscr{C}^\infty$ sur (au moins) l'intervalle ouvert $]-R,R[$ et on peut dériver terme à terme la somme de la série. On sait de plus que G_X est la somme de sa série de Taylor :

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

L'application G_X détermine donc entièrement la loi de X, et réciproquement. On dit que la loi de X est caractérisée par sa série génératrice.

Conséquence pratique : il sera dans certains cas intéressant de déterminer la fonction génératrice d'une variable aléatoire pour trouver sa loi de probabilité.

Le recours aux séries génératrices est particulièrement avisé lorsque l'on souhaite déterminer la loi de la somme de variables aléatoires indépendantes.

Théorème 12.25 : Fonction génératrice de la somme de deux variables indépendantes

Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$.
Notons G_X et G_Y leurs fonctions génératrices de rayon de convergence respectifs R_X et R_Y.
Si X et Y sont indépendantes, la fonction génératrice G_{X+Y} de $X+Y$ est définie sur au moins $]-R, R[$ avec $R \geqslant \min(R_X, R_Y)$ et :

$$\forall t \in D(0, \min(R_X, R_Y)) \quad G_{X+Y}(t) = \mathbf{E}\left(t^{X+Y}\right) = \mathbf{E}\left(t^X\right)\mathbf{E}\left(t^Y\right) = G_X(t)G_Y(t)$$

Démonstration

Par produit de Cauchy de séries absolument convergentes,

$$\begin{aligned} G_X(t)G_Y(t) &= \left(\sum_{n=0}^{+\infty} \mathbf{P}(X=n)t^n\right) \cdot \left(\sum_{n=0}^{+\infty} \mathbf{P}(Y=n)t^n\right) \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \mathbf{P}(X=k)\mathbf{P}(Y=n-k)t^n = \sum_{n=0}^{+\infty} \mathbf{P}(X+Y=n)t^n = G_{X+Y}(t) \end{aligned}$$

■

On généralise aisément par récurrence : $G_{X_1+\cdots+X_n} = G_{X_1} \times \cdots \times G_{X_n}$ en cas d'indépendance.

Exercice 7

Retrouver la loi d'une somme de variables indépendantes suivant des lois de Bernoulli / de Poisson.

Théorème 12.26 : Fonction génératrice et moments

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans $\mathbb{N}$.

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

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

Démonstration

Justifions seulement la première assertion.

$\boxed{\Longrightarrow}$ Supposons que $X \in L^1$. G_X est dérivable sur $]-1, 1[$ et :

$$\forall t \in]-1, 1[, \quad G_X'(t) = \sum_{n=1}^{+\infty} n\mathbf{P}(X=n)t^{n-1}$$

$\sum n\mathbf{P}(X=n)$ converge donc la série $\sum_{n\geqslant 1} n\mathbf{P}(X=n)t^{n-1}$ converge normalement sur $[-1, 1]$. Donc $G_X'(t) \xrightarrow[t\to 1^-]{} \mathbf{E}(X)$. D'après le théorème de la limite de la dérivée, G_X est dérivable en 1 et $G_X'(1) = \mathbf{E}(X)$.

$\boxed{\Longleftarrow}$ Raisonnons par contraposée en supposant que X n'est pas d'espérance finie. G_X' est croissante sur $[0, 1[$ donc admet une limite (finie ou infinie) en 1. Par positivité,

$$\forall p \in \mathbb{N}^*, \quad \forall t \in [0, 1[, \quad \sum_{n=1}^{p} n\mathbf{P}(X=n)t^{n-1} \leqslant \sum_{n=1}^{+\infty} n\mathbf{P}(X=n)t^{n-1}$$

En faisant tendre t vers 1^-, $\sum_{n=1}^{p} n\mathbf{P}(X=n) \leqslant \lim_{t\to 1^-} G_X'(t)$. En faisant tendre p vers $+\infty$, $\lim_{t\to 1^-} G_X'(t) = +\infty$.
Si G_X était dérivable en 1, G_X' serait continue en 1 par convergence normale, d'où la contradiction.

Rajoutons que sous réserve de dérivabilité, $G_X''(1) = \sum_{n=0}^{+\infty} n(n-1)\mathbf{P}(X=n) = \mathbf{E}(X(X-1))$.

Comme $\mathbf{V}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2 = \mathbf{E}(X(X-1)) + \mathbf{E}(X) - \mathbf{E}(X)^2$, $\mathbf{V}(X) = G_X''(1) + G_X'(1) - G_X'(1)^2$. ■

L'expression de la variance en fonction de $G_X'(1)$ et de $G_X''(1)$ n'est pas à retenir.

IV | Inégalités de concentration

Lemme 12.27 : Inégalité de Markov

Soit X une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, à valeurs positives et admettant une espérance.

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

Démonstration

Soit $a > 0$. $X \in L^1$ et $X(\Omega) \subset \mathbb{R}_+$ donc :

$$\mathbf{E}(X) = \sum_{x \in X(\Omega)} x\mathbf{P}(X = x) = \sum_{x \geqslant a} x\mathbf{P}(X = x) + \sum_{x < a} x\mathbf{P}(X = x) \geqslant \sum_{x \geqslant a} x\mathbf{P}(X = x) \geqslant a \sum_{x \geqslant a} \mathbf{P}(X = x) = a\mathbf{P}(X \geqslant a)$$

On pourra préférer la rédaction plus concise suivante :

$X \geqslant X \cdot \mathbb{1}_{(X \geqslant a)} \geqslant a\mathbb{1}_{(X \geqslant a)}$ donc par croissance de l'espérance, $\mathbf{E}(X) \geqslant a\mathbf{P}(X \geqslant a)$. ■

Exercice 8

Soit X une variable aléatoire d'espérance finie et à valeurs dans $\mathbb{R}_+$. Montrer que $\mathbf{P}(X \geqslant n) \underset{n \to +\infty}{=} o\left(\frac{1}{n}\right)$.

Proposition 12.28 : Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire définie sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ admettant un moment d'ordre 2.

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

Démonstration

On applique l'inégalité de Markov à la variable $(X - \mathbf{E}(X))^2$, positive et d'espérance finie, en prenant $a = \varepsilon^2$.

Cette inégalité, bien que grossière comme nous pourrons le constater dans l'exercice ci-dessous, estime la probabilité qu'une variable aléatoire s'écarte de sa moyenne. La variance est donc un indicateur de dispersion.

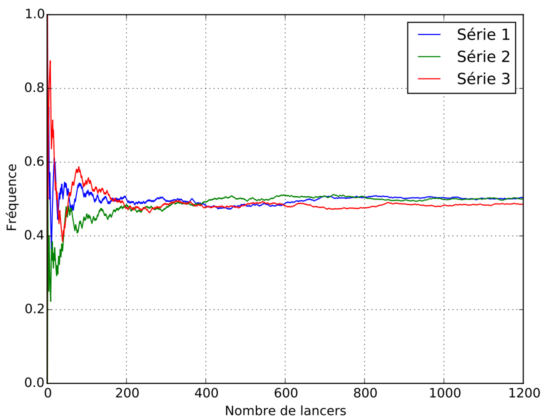
Exercice 9

(i) Comparer $\mathbf{P}(|X - \mathbf{E}(X)| \geqslant 1)$ et $\mathbf{V}(X)$ dans le cas où X représente la somme des chiffres obtenus lors du lancer de deux dés non pipés.

(ii) Trouver $n \in \mathbb{N}$ tel que $\mathbf{P}(X \leqslant n) \leqslant \frac{1}{2}$.

Cette inégalité a avant tout un intérêt théorique, nous allons en particulier l'utiliser pour prouver la loi faible des grands nombres.

Mais avant de présenter le résultat, intéressons-nous aux lancers successifs d'une pièce équilibrée. Intuitivement, la fréquence d'apparition de « Pile » au cours d'un très grand nombre de lancers devrait être proche de $\frac{1}{2}$. De manière plus générale, pour estimer la probabilité d'un événement A, on peut vouloir répéter l'expérience aléatoire sous-jacente un très grand nombre de fois dans l'espoir que la fréquence de réalisation soit proche de la probabilité recherchée. Fol espoir ? Non, même si rien ne nous garantit qu'au cours d'une série particulière de 10 000 lancers d'une pièce équilibrée, la fréquence d'apparition soit proche de $1/2$.



FRÉQUENCE D'APPARITION DE « PILE » AU COURS DE 1200 LANCERS

Considérons une suite de variables aléatoires $(X_n)_{n\geqslant 1}$ indépendantes et de même loi. On suppose qu'elles admettent toutes un moment d'ordre 2 (cadre commode pour la démonstration). Posons alors :

$$M_n = \frac{X_1 + \cdots + X_n}{n}$$

M_n s'interprète comme une moyenne empirique, celle des résultats observés au cours des n premiers « tirages ». Remarquons qu'il n'y a aucune raison que M_n suive une loi semblable aux variables aléatoires X_i. Nous allons cependant montrer que la probabilité que M_n s'écarte de l'espérance commune $\mathbf{E}(X_i)$ tend vers 0 quand n tend vers l'infini. C'est la loi faible des grands nombres.

Théorème 12.29 : Loi faible des grands nombres

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On suppose toutes les variables indépendantes et de même loi, admettant une espérance m et un écart type σ. Posons $S_n = \sum_{i=1}^{n} X_i$. Alors,

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

Démonstration

D'après l'inégalité de Bienaymé-Tchebychev et par indépendance des variables aléatoires,

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbf{V}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0$$

■

L'espérance d'une variable aléatoire s'interprète donc comme le résultat moyen obtenu lors d'une succession d'expériences identiques répétées une infinité de fois. Cette loi faible des grands nombres s'apparente à un théorème de réconciliation des points de vue : l'axiomatique mise en place par Kolmogorov valide l'approche fréquentiste que l'on peut avoir des probabilités.

V | Lois usuelles – synthèse

Nom	**Notation**	$X(\Omega)$	$\mathbf{P}(X=k)$	$\mathbf{E}(X)$	$\mathbf{V}(X)$	$G(t)$
Bernoulli	$\mathscr{B}(p)$	$\{0,1\}$	$\begin{cases} p \text{ si } k=1 \\ q \text{ si } k=0 \end{cases}$	p	pq	$q+pt$
Binomiale	$\mathscr{B}(n,p)$	$[\![0,n]\!]$	$\binom{n}{k} p^k q^{n-k}$	np	npq	$(q+pt)^n$
Uniforme	$\mathscr{U}([\![1,n]\!])$	$[\![1,n]\!]$	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{t-t^{n+1}}{n(1-t)}$
Géométrique	$\mathscr{G}(p)$	$\mathbb{N}^*$	$q^{k-1}p$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pt}{1-qt}$
Poisson	$\mathscr{P}(\lambda)$	$\mathbb{N}$	$\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$	λ	λ	$\mathrm{e}^{\lambda(t-1)}$