

Chapitre XXVI

Espérance, variance

On fixe dans ce chapitre un espace probabilisé (Ω, \mathbb{P}) **fini**.

1. Espérance

a) Notion d'espérance

Définition XXVI.1. Soit X une variable aléatoire réelle ou complexe ; on appelle **espérance** de X la quantité

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

▮ **Exemple XXVI.1.** Si $X \sim \mathcal{U}(n)$, on a :

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

✎ **Remarque XXVI.1.**

- De façon générale, l'espérance de X est la moyenne de ses valeurs pondérée par leur fréquence (cette dernière étant donnée par la loi de X).
- L'espérance de X ne dépend que de sa loi.

Vocabulaire. Si $\mathbb{E}(X) = 0$, on dit que la variable aléatoire X est **centrée**.

Proposition XXVI.1. Soit X une variable aléatoire réelle ou complexe ; alors :

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

Démonstration. On a :

$$\begin{aligned}
 \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) X(\omega) &= \sum_{x \in X(\Omega)} \sum_{\omega \in X^{-1}(\{x\})} X(\omega) \mathbb{P}(\{\omega\}) \\
 &= \sum_{x \in X(\Omega)} \sum_{\omega \in X^{-1}(\{x\})} x \mathbb{P}(\{\omega\}) \\
 &= \sum_{x \in X(\Omega)} x \sum_{\omega \in X^{-1}(\{x\})} \mathbb{P}(\{\omega\}) \\
 &= \sum_{x \in X(\Omega)} x \mathbb{P}(X^{-1}(\{x\})) \\
 &= \sum_{x \in X(\Omega)} x \mathbb{P}(X = x)
 \end{aligned}$$

d'où le résultat. □

✎ **Exercice XXVI.1.** Soit $A \subset \Omega$. Démontrer que $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

b) Zoologie

◇ Variable (presque) constante

Définition XXVI.2. Une variable aléatoire est dite **constante presque sûrement** si il existe $m \in \mathbb{C}$ tel que $\mathbb{P}(X = m) = 1$.

✂ **Remarque XXVI.2.** Il est aisé de démontrer que si X est constante presque sûrement égale à $m \in \mathbb{C}$ alors $\mathbb{E}(X) = m$. On en déduit que pour toute variable aléatoire réelle ou complexe Y on a :

$$\mathbb{E}(\mathbb{E}(Y)) = \mathbb{E}(Y).$$

◇ Loi uniforme

Proposition XXVI.2. Soient $x_1, \dots, x_n \in \mathbb{C}$ et soit $X \sim \mathcal{U}(x_1, \dots, x_n)$. Alors :

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

✂ **Remarque XXVI.3.** Il s'agit de la moyenne arithmétique des valeurs de X .

◇ Loi de Bernoulli

Proposition XXVI.3. Soit $p \in [0, 1]$ et soit $X \sim \mathcal{B}(p)$. Alors :

$$\mathbb{E}(X) = p.$$

Démonstration. Il suffit de remarquer que :

$$\mathbb{E}(X) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) = p.$$

□

◇ Loi binomiale

Proposition XXVI.4. Soit $p \in [0, 1]$ et $n \in \mathbb{N}^*$. Alors, si $X \sim \mathcal{B}(n, p)$ on a :

$$\mathbb{E}(X) = np.$$

Démonstration. On a :

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

d'où le résultat. □

2. Propriétés de l'espérance

a) Propriétés élémentaires

Définition XXVI.3. Un événement $A \subset \Omega$ sera dit **presque sûr** si $\mathbb{P}(A) = 1$.

▣► **Exemple XXVI.2.** La variable aléatoire comptant le nombre de fois où une pièce équilibrée tombe sur la tranche parmi 42 lancers est nulle presque sûrement dans le modèle précédemment établi pour cette expérience aléatoire.

☞ **Remarque XXVI.4.** Avec cette terminologie, on peut dire que connaître la loi d'une variable aléatoire revient à connaître cette dernière presque sûrement.

Proposition XXVI.5 (Linéarité de l'espérance). Soient X, Y deux variables aléatoires réelles ou complexes et soit $\lambda \in \mathbb{C}$. Alors :

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

Démonstration. On a :

$$\mathbb{E}(X + \lambda Y) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\})(X(\omega) + \lambda Y(\omega))$$

d'où le résultat. □

Proposition XXVI.6. Soit X une variable aléatoire réelle telle que $X \geq 0$ presque sûrement (*i.e* $\mathbb{P}(X \geq 0) = 1$); alors :

- (i) $\mathbb{E}(X) \geq 0$;
- (ii) $\mathbb{E}(X) = 0 \Leftrightarrow X$ est nulle presque sûrement (*i.e* $\mathbb{P}(X = 0) = 1$).

Démonstration. Posons $X(\Omega) = \{x_1, \dots, x_n\}$. Alors la quantité

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

est une somme de termes positifs, d'où le résultat. □

Corollaire XXVI.6.a (Croissance de l'espérance). Soient X, Y deux variables aléatoires réelles telles que $X \leq Y$ presque sûrement; alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

✂ **Remarque XXVI.5.** On retrouve ainsi l'espérance d'une variable aléatoire de loi binomiale à l'aide de la proposition [XXIV.13](#).

▣➡ **Exemple XXVI.3.** Soient $X, Y \sim \mathcal{U}(6)$; alors $\mathbb{E}(X + Y) = 7$. Nous venons de déterminer la somme moyenne des résultats de deux dés à 6 faces.

b) Lemme de transfert

Proposition XXVI.7 (Lemme de transfert). Soit $X : \Omega \rightarrow E$ une variable aléatoire quelconque et soit $\varphi : E \rightarrow \mathbb{R}$ ou \mathbb{C} une application. Alors :

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

✂ **Remarque XXVI.6.** La somme *supra* est bien une somme finie, car si $x \notin X(\Omega)$, $\mathbb{P}(X = x) = 0$.

Démonstration. Posons $X(\Omega) = \{x_1, \dots, x_n\}$. Alors, si l'on pose pour $k \in \llbracket 1, n \rrbracket$ $A_k = \{X = x_k\}$ on a :

$$X = \sum_{k=1}^n x_k \mathbb{1}_{A_k}$$

et donc :

$$\varphi(X) = \sum_{k=1}^n \varphi(x_k) \mathbb{1}_{A_k}.$$

Ainsi :

$$\begin{aligned} \mathbb{E}(\varphi(X)) &= \mathbb{E}\left(\sum_{k=1}^n \varphi(x_k) \mathbb{1}_{A_k}\right) \\ &= \sum_{k=1}^n \varphi(x_k) \mathbb{E}(\mathbb{1}_{A_k}) \\ &= \sum_{k=1}^n \varphi(x_k) \mathbb{P}(A_k) \end{aligned}$$

d'après le résultat de l'exercice **XXVI.1**. On en déduit le résultat. \square

☞ Remarque XXVI.7.

- Une conséquence importante du lemme de transfert est que l'espérance de $\varphi(X)$ est totalement déterminée par la loi de X et la fonction φ .
- Noter que le lemme de transfert s'applique au cas où X est un couple ou un n -uplet (se placer sur un espace produit à l'arrivée).

☛ Exemple XXVI.4.

- Soit $X \sim \mathcal{U}(n)$. Alors :

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_{k=1}^n \frac{k^2}{n} \\ &= \frac{(n+1)(2n+1)}{6}. \end{aligned}$$

- Soit X une variable aléatoire réelle ou complexe. Alors :

$$\begin{aligned} \mathbb{E}(|X|) &= \sum_{x \in X(\Omega)} |x| \mathbb{P}(X = x) \\ &\geq \left| \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) \right| \end{aligned}$$

et donc on obtient une "inégalité triangulaire" :

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$$

☛ Exercice XXVI.2. Soient $X, Y \sim \mathcal{U}(n)$ deux variables aléatoires indépendantes ; déterminer $\mathbb{E}(XY^2)$.

☛ Exercice XXVI.3. Soient $X_1, \dots, X_n \sim \mathcal{B}(p)$ mutuellement indépendantes, avec $p \in [0, 1]$. On pose :

$$Y_n = \frac{1}{n} \sum_{k=1}^n X_k.$$

Déterminer, pour $t \in \mathbb{R}$, la quantité $\mathbb{E}(e^{tY_n})$.

➔ **Correction :** On commence par remarquer que :

$$\mathbb{E}(e^{tY_n}) = \mathbb{E}\left(\prod_{k=1}^n e^{tX_k/n}\right).$$

Or, par lemme des coalitions, les variables $e^{tX_1/n}, \dots, e^{tX_n/n}$ sont mutuellement indépendantes, ergo :

$$\begin{aligned}\mathbb{E}(e^{tY_n}) &= \prod_{k=1}^n \mathbb{E}(e^{tX_k/n}) \\ &= (1 + p(e^{t/n} - 1))^n\end{aligned}$$

la dernière égalité étant obtenue par lemme de transfert (proposition **XXVI.7**).

c) Inégalité de Markov

La proposition *infra* est ainsi nommée en l'honneur du mathématicien russe Andreï Andreïevitch Markov (1856—1922), connu comme l'un des pionniers de la théorie moderne des probabilités (on lui doit notamment la notion de processus stochastique).

Théorème XXVI.8 (Inégalité de Markov).

Soit X une variable aléatoire réelle telle que $X \geq 0$ presque sûrement et soit $t \in \mathbb{R}_+^*$. Alors :

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Démonstration. Posons $A = \{x \in X(\Omega) \mid x \geq t\}$. Alors :

$$\begin{aligned}t \mathbb{P}(X \geq t) &= t \sum_{x \in A} \mathbb{P}(X = x) \\ &= \sum_{x \in A} t \mathbb{P}(X = x) \\ &\leq \sum_{x \in A} x \mathbb{P}(X = x) \\ &\leq \mathbb{E}(X)\end{aligned}$$

d'où le résultat. □

✂ **Remarque XXVI.8.** Soit X une variable aléatoire réelle. Alors on a, par exemple, que :

$$\mathbb{P}(X \geq 100\mathbb{E}(X)) \leq \frac{1}{100}.$$

Ceci nous permet d'estimer la répartition des valeurs de notre variable aléatoire X et de visualiser la façon dont la probabilité d'un écart à la moyenne donné évolue.

▮ **Exemple XXVI.5.** Avec cette inégalité, on peut démontrer que la probabilité qu'un dé à 6 faces fasse au moins 4 est inférieure à

$\frac{7}{8}$, ou que la probabilité qu'il dépasse 7 est de moins de $\frac{1}{2}$: on n'a pas ici affaire à l'estimation la plus fine qui soit !

d) Lien à l'indépendance

Proposition XXVI.9. Soient X et Y deux variables aléatoires réelles ou complexes indépendantes. Alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

✘ **ATTENTION :** l'indépendance est ici essentielle. Si $X \sim \mathcal{B}(p)$, on a $\mathbb{E}(X^2) = \mathbb{E}(X) = p$ car $X^2 = X$ presque sûrement. Pour peu que p ne soit pas égal à 0 ou 1, on a donc $\mathbb{E}(X^2) \neq \mathbb{E}(X)^2$.

Démonstration. Posons $X(\Omega) = \{x_1, \dots, x_n\}$ et $Y(\Omega) = \{y_1, \dots, y_m\}$. Alors, en appliquant le lemme de transfert au couple (X, Y) on obtient :

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{k=1}^n \sum_{\ell=1}^m x_k y_\ell \mathbb{P}(X = x_k, Y = y_\ell) \\ &= \sum_{k=1}^n \sum_{\ell=1}^m x_k y_\ell \mathbb{P}(X = x_k) \mathbb{P}(Y = y_\ell) \end{aligned}$$

par indépendance. *In fine*, on a donc :

$$\begin{aligned} \mathbb{E}(XY) &= \left(\sum_{k=1}^n x_k \mathbb{P}(X = x_k) \right) \left(\sum_{\ell=1}^m y_\ell \mathbb{P}(Y = y_\ell) \right) \\ &= \mathbb{E}(X)\mathbb{E}(Y). \end{aligned}$$

□

✘ **ATTENTION :** la réciproque est **FAUSSE** : si $X = m$ presque sûrement pour un certain $m \in \mathbb{R}$ on a $\mathbb{E}(X^2) = m^2 = \mathbb{E}(X)^2$ et pourtant X n'est pas indépendante d'elle-même (avouez que ce serait sympathique, mais le monde est injuste).

☞ **Remarque XXVI.9.** Ce résultat se généralise au cas de n variables aléatoires mutuellement indépendantes.

3. Variance

a) C'est quoi ?

Connaître l'espérance d'une variable aléatoire réelle X nous permet d'avoir une idée de l'endroit moyen où se situent ses valeurs : c'est ce que l'on appelle un **paramètre de position**. Une telle donnée, bien qu'utile, n'est en général pas suffisante pour étudier finement le comportement de notre variable aléatoire malgré l'existence d'outils tels que l'inégalité de Markov (théorème [XXVI.8](#)). La question qui nous préoccupe dans ce paragraphe est la suivante : comment diable sont réparties les valeurs de X ? Sont-elles resserrées autour de la moyenne ? Fortement dispersées ? Les quantités nous renseignant à ce sujet sont traditionnellement appelés **paramètres de dispersion**.

Une première idée pourrait être de déterminer l'espérance de la variable aléatoire réelle ou complexe $|X - \mathbb{E}(X)|$; celle se révèle hélas en général d'une étude assez

pénible. C'est pourquoi nous introduisons les quantités *infra*, dont nous dédierons cet ultime paragraphe de l'année à l'étude.

Définition XXVI.4. Soit X une variable aléatoire réelle ; alors on appelle **variance** de X la quantité

$$V(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

et **écart-type** de X la quantité :

$$\sigma(X) = \sqrt{V(X)}.$$

Vocabulaire. Une variable aléatoire réelle de variance égale à 1 est dite **réduite**.

✂ **Remarque XXVI.10.** Si nous souhaitons parler comme les physiciens, nous pourrions dire que $\sigma(X)$ est de même unité/dimension que X . Cette quantité représente l'écart moyen entre deux valeurs prises par la variable aléatoire X .

Le résultat *infra* est nommé en l'honneur du mathématicien (et élève de Jean Bernoulli, premier du nom) allemand Johann Samuel König (1712—1757), connu entre autres pour avoir traduit les éléments d'Euclide et du mathématicien, astronome et physicien néerlandais Christiaan Huygens (1629—1695), inventeur de la première horloge à pendule.

Théorème XXVI.10 (König–Huygens).

Soit X une variable aléatoire réelle ; alors :

$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Démonstration. On a :

$$\begin{aligned} V(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) \\ &= \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2. \end{aligned}$$

□

▣ **Exemple XXVI.6.** Si X correspond à la somme des résultats de deux dés à 6 faces indépendants, on a $V(X) = \frac{35}{6}$ et $\sigma(X) \simeq 2.4$.

Proposition XXVI.11. Soit X une variable aléatoire réelle et soient $a, b \in \mathbb{R}$. Alors :

- (i) $V(aX + b) = a^2V(X)$;
- (ii) $\sigma(aX + b) = |a|\sigma(X)$.

Démonstration. Il s'agit d'un calcul direct. \square

Corollaire XXVI.11.a. Soit X une variable aléatoire réelle telle que $\sigma(X) > 0$; alors la variable aléatoire $\frac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée réduite.

b) Zoologie

◇ Variable presque constante

Soit X une variable aléatoire réelle telle que $X = m$ presque sûrement pour un certain $m \in \mathbb{R}$. Alors, il découle de la formule de König–Huygens (théorème [XXVI.10](#)) que $V(X) = \sigma(X) = 0$. La réciproque est, pour une fois, vraie.

Proposition XXVI.12. Soit X une variable aléatoire réelle telle que $V(X) = 0$. Alors X est constante presque sûrement.

Démonstration. Posons $Y = (X - \mathbb{E}(X))^2$; alors $\mathbb{E}(Y) = V(X) = 0$ et donc $Y = 0$ presque sûrement d'après la proposition [XXVI.6](#), d'où le résultat. \square

◇ Loi de Bernoulli

Proposition XXVI.13. Soit $p \in [0, 1]$ et soit $X \sim \mathcal{B}(p)$. Alors $V(X) = p(1 - p)$.

Démonstration. Il suffit de faire le calcul en remarquant que $X^2 = X$ presque sûrement. \square

◇ Loi binomiale

Proposition XXVI.14. Soit $p \in [0, 1]$ et soit $n \in \mathbb{N}^*$. Alors, si $X \sim \mathcal{B}(n, p)$ on a $V(X) = np(1 - p)$.

Démonstration. On a, par lemme de transfert (proposition [XXVI.7](#)) :

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

De fait, on a :

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

□

✂ **Remarque XXVI.11.** Si $X \sim \mathcal{B}(n, p)$ on a $\sigma(X) = \sqrt{p(1-p)}\sqrt{n} = o(n)$ lorsque $n \rightarrow \infty$. Cela signifie que les écarts entre les valeurs prises par X et l'espérance $\mathbb{E}(X)$ sont faibles devant sa moyenne. À titre d'exemple, pour $n = 10\,000$ et $p = \frac{1}{2}$ on a $\mathbb{E}(X) = 5000$ et $\sigma(X) = 50 \dots$

c) Inégalité de Bienaymé–Tchebychev

Le résultat présenté dans ce paragraphe est nommé en l'honneur du mathématicien français Irénée-Jules Bienaymé (1796–1878), connu pour ses contributions à la théorie des probabilités et à la statistique, et de son homologue russe Pafnouti Lvovitch Tchebychev (1821–1894), dont les travaux ont porté sur les probabilités, la statistique et la théorie des nombres.

Théorème XXVI.15 (Bienaymé–Tchebychev).

Soit X une variable aléatoire réelle ; on pose $m = \mathbb{E}(X)$ et $\sigma = \sigma(X)$. Alors, pour tout $\lambda \in \mathbb{R}_+^*$ on a :

$$\mathbb{P}(|X - m| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}.$$

Démonstration. En appliquant l'inégalité de Markov (théorème [XXVI.8](#)) à la variable aléatoire réelle $(X - m)^2$ avec $t = \lambda^2$, on obtient :

$$\mathbb{P}((X - m)^2 \geq \lambda^2) \leq \frac{\mathbb{E}((X - m)^2)}{\lambda^2}$$

soit

$$\mathbb{P}((X - m)^2 \geq \lambda^2) \leq \frac{V(X)}{\lambda^2}.$$

Il suffit alors pour conclure de remarquer que $V(X) = \sigma^2$ et que, par positivité de λ :

$$\mathbb{P}((X - m)^2 \geq \lambda^2) = \mathbb{P}(|X - m| \geq \lambda).$$

□

☞ **Remarque XXVI.12.** Il peut être intéressant de formuler cette inégalité en remplaçant λ par le produit $\lambda\sigma$. On obtient alors :

$$\mathbb{P}(|X - m| \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$$

ce qui nous permet d'estimer l'écart de la variable aléatoire X à sa moyenne par incréments proportionnels à son écart-type.

▮► **Exemple XXVI.7.** Lançons $n \geq 1$ dés à 6 faces et quantifions le nombre de 6 obtenus par une variable aléatoire X . Il est alors clair que $X \sim \mathcal{B}\left(n, \frac{1}{6}\right)$ et donc, par inégalité de Bienaymé–Tchebychev (théorème [XXVI.15](#)), on obtient pour tout $\lambda > 0$ que :

$$\mathbb{P}\left(\left|X - \frac{n}{6}\right| \geq \lambda\right) \leq \frac{5n}{36\lambda^2}.$$

Pour $\lambda = \frac{n}{6}$, ceci devient

$$\mathbb{P}\left(\left|X - \frac{n}{6}\right| \geq \frac{n}{6}\right) \leq \frac{5}{n}.$$

On vérifie ensuite par un rapide calcul que $\left|X - \frac{n}{6}\right| \geq \frac{n}{6} \Leftrightarrow X \geq \frac{n}{3}$ et donc, *in fine* :

$$\mathbb{P}\left(X \geq \frac{n}{3}\right) \leq \frac{5}{n}$$

i.e

$$\mathbb{P}\left(X < \frac{n}{3}\right) \geq 1 - \frac{5}{n}.$$

Ceci signifie que lorsque $n = 10$, la probabilité que les 6 représentent moins d'un tiers des résultats obtenus est supérieure à $\frac{1}{2}$. Pour $n = 50$, elle dépasse $\frac{9}{10}$.

4. Covariance

a) Co-quoi ?

Définition XXVI.5. Soient X, Y deux variables aléatoires réelles ; on appelle **covariance** de X et Y la quantité

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

☞ **Remarque XXVI.13.** $\text{Cov}(X, X) = V(X)$ pour toute variable aléatoire réelle X .

Proposition XXVI.16. Soient X, Y deux variables aléatoires réelles ; alors :

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

Démonstration. Il s'agit d'un calcul immédiat et remarquablement indolore (affirmation non contractuelle). \square

Corollaire XXVI.16.a. Soient X, Y deux variables aléatoires réelles **indépendantes** ; alors $\text{Cov}(X, Y) = 0$.

✘ **ATTENTION** : la réciproque est **FAUSSE** (avouez, vous êtes surpris). Prenons $X \sim \mathcal{U}(-1, 0, 1)$ et posons $Y = X^2$: il est clair que X et Y ne sont pas indépendantes et pourtant $XY = X^3 = X$ presque sûrement donc

$$\mathbb{E}(XY) = \mathbb{E}(X) = 0 = \mathbb{E}(X)\mathbb{E}(Y).$$

Vocabulaire. Deux variables aléatoires réelles X, Y telles que $\text{Cov}(X, Y) = 0$ sont dites **décorrélées**.

Proposition XXVI.17. Soient X, Y et Z trois variables aléatoires réelles et soient $a, b, c, d \in \mathbb{R}$. Alors :

- (i) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;
- (ii) $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$;
- (iii) $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$.

Démonstration. Il s'agit d'une succession de calculs aussi enrichissants que la formidable activité consistant à regarder de la peinture sécher. \square

b) Sommes de variables aléatoire

Proposition XXVI.18. Soient X, Y deux variables aléatoires réelles. Alors :

- (i) $V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$;
- (ii) $V(X - Y) = V(X) + V(Y) - 2\text{Cov}(X, Y)$.

Démonstration. Il suffit de remarquer que $V(X \pm Y) = \text{Cov}(X \pm Y, X \pm Y)$ et de développer à l'aide des propriétés énoncées dans la proposition **XXVI.17**. \square

Démonstration. On lance deux dés à 6 faces, notant X le résultant du premier et Y celui du second. Alors la variance de la somme est :

$$V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y) = \frac{35}{12} + \frac{35}{12} + 2 \times 0 = \frac{35}{6}.$$

\square

✂ **Remarque XXVI.14.** On peut démontrer par récurrence (youpi) que pour toute famille X_1, \dots, X_n de variables aléatoires réelles, on a :

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

Corollaire XXVI.18.a. Soient X_1, \dots, X_n des variables aléatoires réelles **mutuellement indépendantes**. Alors :

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

✂ **Remarque XXVI.15.** Ce résultat permet de retrouver aisément la variance d'une variable aléatoire réelle ou complexe $X \sim \mathcal{B}(n, p)$ via la proposition **XXIV.13**.

✂ **Remarque XXVI.16.** Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes de même loi. On note m (resp. σ^2) l'espérance (resp. la variance) commune à ces variables. On pose alors, pour tout $n \geq 1$:

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k.$$

On vérifie rapidement que cette nouvelle variable aléatoire est d'espérance m et que :

$$\forall n \geq 1, \quad V(\bar{X}_n) = \sum_{k=1}^n V\left(\frac{X_k}{n}\right) = \sum_{k=1}^n \frac{V(X_k)}{n^2} = \frac{\sigma^2}{n}.$$

On a alors, par inégalité de Bienaymé–Tchebychev, pour tout $\varepsilon > 0$:

$$\forall n \geq 1, \quad \mathbb{P}(|\bar{X}_n - m| \geq \varepsilon) \leq \frac{\sigma^2 \varepsilon^2}{n^2}$$

et donc :

$$\forall \varepsilon > 0, \quad \mathbb{P}(|\bar{X}_n - m| \geq \varepsilon) \rightarrow 0.$$

Ce résultat est connu sous le nom de **loi faible des grands nombres**. Il nous indique que la moyenne empirique (la variable \bar{X}_n) "converge" (on parlera de convergence en probabilité) vers la moyenne "théorique" (l'espérance).