

Espérance et variance

Dans tout le chapitre, $(\Omega, \mathbb{P})$ désigne un espace probabilisé fini.

1 ESPÉRANCE

Définition 1 : Espérance d'une variable aléatoire

Soit X une variable aléatoire réelle finie $X : \Omega \rightarrow E = \{x_1, \dots, x_n\} \subset \mathbb{R}$. On définit l'**espérance de X** , et on le note $E(X)$ par

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

On dit que X est **centrée** si $E(X) = 0$.

Qu'est ce que l'espérance

L'espérance de X est la *somme des valeurs coefficientées par la probabilité d'apparition de cette valeur*, soit la moyenne des valeurs possibles pour X pondérées par leurs probabilités d'apparaître. Cette quantité dépend de la loi de X mais pas de la réalisation de cette loi. Autrement dit deux variables de même loi ont la même espérance, mais pas inversement.

L'espérance est donc un indicateur de position. Plus $P(X = x)$ est proche de 1, plus x a d'importance dans le calcul, et donc plus l'espérance est proche de x .

Exemple 1

Un dé à 6 faces numérotées de 1 à 6 a été truqué de telle sorte que les faces 1,2 et 3 tombent avec probabilité $\frac{1}{6}$, les faces 4 et 5 avec probabilité $\frac{1}{12}$ et la face 6 avec probabilité $\frac{1}{3}$. Quel numéro obtient-on en moyenne?

On peut vérifier que les probabilités sont bien définies, car on a $3 \times \frac{1}{6} + 2 \times \frac{1}{12} + \frac{1}{3} = 1$.

Notons X le numéro obtenu avec un tel dé après un lancer.

Comme $P(X = 1) = P(X = 2) = P(X = 3) = \frac{1}{6}$, $P(X = 4) = P(X = 5) = \frac{1}{12}$ et $P(X = 6) = \frac{1}{3}$, on a

$$E(X) = \sum_{k=1}^6 P(X = k)k = \frac{1+2+3}{6} + \frac{4+5}{12} + \frac{6}{3} = 1 + \frac{3}{4} + 2 = \frac{7}{4}.$$

Exemple 2

On lance deux dés. Soit X la variable aléatoire égale à la somme des chiffres apparus.

On a $\mathbb{P}(X = k) = \begin{cases} \frac{k-1}{36} & \text{si } k \in \llbracket 1, 6 \rrbracket \\ \frac{13-k}{36} & \text{si } k \in \llbracket 7, 13 \rrbracket \end{cases}$ Donc

$$\mathbb{E}(X) = 2\frac{1}{36} + 3\frac{1}{18} + 4\frac{1}{12} + 5\frac{1}{9} + 6\frac{5}{36} + 7\frac{1}{6} + 8\frac{5}{36} + 9\frac{1}{9} + 10\frac{1}{12} + 11\frac{1}{18} + 12\frac{1}{36} = 7.$$

Autre formule pour l'espérance

Soit X une variable aléatoire à valeurs dans $\llbracket 0, n \rrbracket$. Réécrivons la formule donnant l'espérance en remarquant que, pour tout $k \in \llbracket 0, n \rrbracket$, $P(X = k) = P(X \geq k) - P(X \geq k + 1)$. On obtient

$$E(X) = \sum_{k=0}^n k(P(X \geq k) - P(X \geq k + 1)) = \sum_{k=0}^n kP(X \geq k) - \sum_{k=1}^{n+1} (k-1)P(X \geq k) = \sum_{k=1}^n P(X \geq k) .$$

Ce qui donne

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

un formule utile en pratique.

Proposition 1 : Espérance des lois usuelles

Soit X une variable aléatoire complexe.

- ① **Variables aléatoires constantes** X est certaine de valeur c , alors $E(X) = c$.
- ② **Loi uniforme** Soit $E = \{x_1, \dots, x_n\}$ une partie de $\mathbb{C}$. Si $X \sim \mathcal{U}(E)$, alors $E(X) = \frac{1}{n} \sum_{k=1}^n x_k$ est la moyenne arithmétique des valeurs $x_1, \dots, x_n$ prises par X .
- ③ **Loi de Bernoulli** Soit $p \in [0, 1]$. Si $X \sim \mathcal{B}(p)$, alors $E(X) = p$.
- ④ **Loi binomiale** Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. Si $X \sim \mathcal{B}(n, p)$, alors $E(X) = np$.

Preuve.

- ① **Variables aléatoires constantes** $E(X) = c P(X = c) = c$.
- ② **Loi uniforme** $E(X) = \sum_{k=1}^n kP(X = k) = \sum_{k=1}^n k \frac{1}{n} = \frac{n+1}{2}$.
- ③ **Loi de Bernoulli** $E(X) = 1 \times P(X = 1) + 0 \times P(X = 0) = p$.
- ④ **Loi binomiale** $E(X) = \sum_{k=0}^n P(X = k)k = \sum_{k=0}^n \binom{n}{k} kp^k(1-p)^{n-k}$. Pour trouver le résultat, on utilise les polynômes.

$$\text{On a } (X + 1 - p)^n = \sum_{k=0}^n \binom{n}{k} X^k (1-p)^{n-k}.$$

$$\text{En dérivant puis en multipliant par } X, \text{ on obtient } nT(T + 1 - p)^{n-1} = \sum_{k=0}^n \binom{n}{k} kT^k (1-p)^{n-k}.$$

$$\text{Il suffit alors d'évaluer en } p, \text{ et on a } np = \sum_{k=0}^n \binom{n}{k} kp^k (1-p)^{n-k}.$$

□

Espérance de l'indicatrice

Pour tout événement $A \in \mathcal{P}(\Omega)$, on a $\mathbb{1}_A \sim \mathcal{B}(P(A))$, donc

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

Proposition 2 : Propriétés de l'espérance

Soient X et Y deux variables aléatoires complexes sur Ω .

① **Formule par la probabilité des événements élémentaires** $E(X) = \sum_{\omega \in \Omega} P(\{\omega\})X(\omega).$

② **Linéarité** Pour tous $\lambda, \mu \in \mathbb{C}$, $E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$

③ **Inégalité triangulaire** $|E(X)| \leq E(|X|).$

④ **Lien avec les parties réelle et imaginaire** $E(X) = E(\operatorname{Re}(X)) + i E(\operatorname{Im}(X)).$

De plus, si X et Y sont **réelles**.

⑤ **Positivité** Si $X \geq 0$, alors $E(X) \geq 0$.
De plus, si $E(X) = 0$, alors $\mathbb{P}(X = 0) = 1$, et X est dite **presque sûrement nulle**.

⑥ **Croissance** Si $X \leq Y$, alors $E(X) \leq E(Y).$

Preuve.

① Pour tout $k \in X(\Omega)$, les événements $\{X = k\}$ forment un système complet d'événements, donc

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

② La linéarité découle immédiatement de la propriété précédente

$$E(\lambda X + \mu Y) \stackrel{①}{=} \sum_{\omega \in \Omega} P(\{\omega\})(\lambda X + \mu Y)(\omega) = \lambda \sum_{\omega \in \Omega} P(\{\omega\})X(\omega) + \mu \sum_{\omega \in \Omega} P(\{\omega\})Y(\omega) = \lambda E(X) + \mu E(Y).$$

$$③ \quad |E(X)| \stackrel{①}{=} \left| \sum_{\omega \in \Omega} P(\{\omega\})X(\omega) \right| \leq \sum_{\omega \in \Omega} P(\{\omega\})|X(\omega)| \stackrel{①}{=} E(|X|).$$

$$④ \quad E(X) \stackrel{①}{=} \sum_{\omega \in \Omega} P(\{\omega\})X(\omega) = \sum_{\omega \in \Omega} P(\{\omega\})(\operatorname{Re}(X)(\omega) + i \operatorname{Im}(X)(\omega)) = \sum_{\omega \in \Omega} P(\{\omega\}) \operatorname{Re}(X)(\omega) + i \sum_{\omega \in \Omega} P(\{\omega\}) \operatorname{Im}(X)(\omega) \\ \stackrel{①}{=} E(\operatorname{Re}(X)) + i E(\operatorname{Im}(X)).$$

⑤ L'espérance est positive par définition, si $X \geq 0$.

Le cas d'égalité implique que la somme de termes positifs $\sum_{x \in E} xP(X = x)$ est nulle, donc tous les termes positifs doivent être nuls. Pour $x \neq 0$, $xP(X = x) = 0$, implique que $P(X = x) = 0$.
Donc $P(X = 0) = \sum_{x \in E} P(X = x) = P(\Omega) = 1$.

⑥ Pour la croissance, si $X \leq Y$, alors $Y - X \geq 0$, donc par linéarité $E(Y) - E(X) = E(Y - X) \geq 0$.

□

Formule du crible

On a maintenant une nouvelle preuve de la **formule du crible**. Cette formule est hors programme et ne peut donc pas être utilisée.

$$\text{Pour tous } A_1, \dots, A_n \in \mathcal{P}(\Omega), \quad P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}).$$

On a d'abord

$$\mathbb{1}_{A_1 \cup \dots \cup A_n} = 1 - \mathbb{1}_{\overline{A_1 \cup \dots \cup A_n}} = 1 - \mathbb{1}_{\bar{A}_1 \cap \dots \cap \bar{A}_n} = 1 - \prod_{i=1}^n \mathbb{1}_{\bar{A}_i} = 1 - \prod_{i=1}^n (1 - \mathbb{1}_{A_i}) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{1}_{A_{i_1} \cap \dots \cap A_{i_k}}$$

Donc, par linéarité de l'espérance

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \mathbb{E}(\mathbb{1}_{A_1 \cup \dots \cup A_n}) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{E}(\mathbb{1}_{A_{i_1} \cap \dots \cap A_{i_k}}) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}).$$

Ce qui donne la formule voulue.

Proposition 3 : Formule de transfert

Soient X une variable aléatoire sur Ω et $f : X(\Omega) \rightarrow \mathbb{C}$ une fonction.

L'espérance de $f(X)$ est entièrement déterminée par f et la loi de X ,

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

Exemple 3

Soit X une variable de loi binomiale de paramètres n et p . Calculons $\mathbb{E}\left(\frac{1}{X+1}\right)$ avec la formule de transfert.

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

Exemple 4

Soit X_n une variable aléatoire de loi uniforme sur $\llbracket 1, n \rrbracket$. On pose $U_n = \frac{X_n}{n}$.

Donc U_n suit la loi uniforme sur $\left\{\frac{k}{n} \mid k \in \llbracket 1, n \rrbracket\right\}$ et pour toute fonction $f \in \mathcal{C}_m([0, 1], \mathbb{C})$, d'après la formule de transfert, on a

$$\mathbb{E}(f(U_n)) = \sum_{k=1}^n P(X_n = k) f\left(\frac{k}{n}\right) = \frac{1}{n} \sum_{k=1}^n f\left(\frac{k}{n}\right) \xrightarrow{n \rightarrow +\infty} \int_0^1 f(x) dx,$$

d'après le théorème sur les sommes de Riemann.

Dans la pratique

L'intérêt de la formule de transfert est qu'elle permet le calcul de l'espérance de $f(X)$ sans qu'il y ait besoin de déterminer sa loi. Il suffit de connaître celle de X .

La formule s'applique en particulier quand X est un couple ou un n -uplet de variables aléatoires. Le calcul de l'espérance du produit de deux variables aléatoires en donne un premier exemple. La formule de transfert peut donc être vue comme une généralisation de la linéarité de l'espérance.

Exemple 5

Soit X et Y deux variables aléatoires complexes sur l'espace probabilisé $(\Omega, \mathbb{P})$. Supposons que la loi conjointe de (X, Y) est une loi uniforme sur $\llbracket 1, n \rrbracket \times \llbracket 1, n \rrbracket$.

Pour calculer $\mathbb{E}(XY)$, on peut écrire $XY = f(Z)$, où $Z = (X, Y)$ et $f : (x, y) \mapsto xy$. En appliquant la formule de transfert à $f(Z)$, on obtient

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

Donc, en remplaçant la loi de (X, Y) et la fonction f , on a

$$\mathbb{E}(XY) = \frac{1}{n^2} \sum_{(x,y) \in \llbracket 1, n \rrbracket^2} xy = \frac{1}{n^2} \left(\sum_{x=1}^n x \right)^2 = \frac{1}{n^2} \left(\frac{n(n+1)}{2} \right)^2 = \frac{(n+1)^2}{4}.$$

Preuve. Comme précédemment, on réutilise le système complet d'événements, $\{X = x\}$, avec x décrivant $X(\Omega)$.

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

□

Proposition 4 : Espérance d'un produit de variable aléatoires indépendantes

Soient X et Y deux variables aléatoires complexes, indépendantes, sur $(\Omega, \mathbb{P})$. On a alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Preuve. En utilisant la même preuve que dans l'exemple 3,

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

□

Indépendance mutuelle de n variable aléatoire

Bien sûr, une preuve raide par récurrence montre que Si $X_1, \dots, X_n$ sont des variables aléatoires complexes mutuellement indépendantes sur $(\Omega, \mathbb{P})$, alors

$$\mathbb{E} \left(\prod_{k=1}^n X_k \right) = \prod_{k=1}^n \mathbb{E}(X_k).$$

⚠ ATTENTION L'identité $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ est fautive en général. Par exemple, pour toute variable aléatoire X de loi uniforme sur $\{-1, 1\}$, on a $\mathbb{E}(X) = 0$, donc $\mathbb{E}(X)^2 = 0$. Pourtant $\mathbb{E}(X^2) = \mathbb{E}(1) = 1$.

Proposition 5 : Inégalité de Cauchy Schwarz

Soient X et Y deux variables aléatoires réelles. Alors, $\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2) \mathbb{E}(Y^2)$.

Preuve. Considérons la fonction polynomiale $P : \lambda \mapsto \mathbb{E}((\lambda X + Y)^2)$. Cette fonction est positive, comme espérance d'une variable positive, et s'écrit en développant

$$\forall \lambda \in \mathbb{R}, \quad P(\lambda) = \mathbb{E}(X^2) \lambda^2 + 2\mathbb{E}(XY) \lambda + \mathbb{E}(Y^2).$$

Si $\mathbb{E}(X^2) = 0$, la fonction P est de degré au plus 1 et de signe constant donc est une constante, ainsi $\mathbb{E}(XY) = 0$ et l'inégalité est vérifiée.

Sinon, P est de degré 2 et de signe constant, donc son discriminant est négatif soit

$$4 \mathbb{E}(XY)^2 - 4 \mathbb{E}(X^2) \mathbb{E}(Y^2) \leq 0,$$

ce qui est bien l'inégalité attendue. □

Centralisation d'une variable aléatoire

L'espérance d'une variable aléatoire donne la valeur prise en moyenne par X . On peut centrer X , en posant la variable aléatoire $X - \mathbb{E}(X)$ qui a donc pour espérance 0.

2 VARIANCE, ÉCART TYPE, COVARIANCE

Définition 2 : Moments

Pour tout $k \in \mathbb{N}$, le **moment d'ordre k d'une variable aléatoire réelle X** est le réel $\mathbb{E}(X^k)$ et son **moment centré d'ordre k** est $\mathbb{E}((X - \mathbb{E}(X))^k)$.

Vocabulaire

- Le moment d'ordre 2 est appelé **variance**, noté $\mathbb{V}(X)$.
- L'**écart type** est noté $\sigma(X) := \sqrt{\mathbb{V}(X)}$.
- On dit que X est **réduite** si $\mathbb{V}(X) = 1$.

Qu'est ce que la variance et l'écart-type?

L'espérance de X est un indicateur de position. Mais comment peut-on savoir si les valeurs de X sont-elles plutôt proches de cette valeur moyenne ou plutôt éloignées ?

Une mesure de cette proximité à la moyenne est appelée **indicateur de dispersion**.

L'écart le plus naturel entre X et son espérance est $|X - \mathbb{E}(X)|$. On peut donc calculer l'écart moyen de X à sa moyenne $\mathbb{E}(|X - \mathbb{E}(X)|)$.

La variance, c'est-à-dire l'**écart quadratique**, parce qu'on passe au carré, moyen à la moyenne est un autre écart souvent plus utilisé, car il met en exergue les valeurs de X très lointaine de la moyenne, mais aussi parce que, la variance est plus facile à manipuler d'un point de vue calculatoire.

Le problème de la variance, si par exemple X représente une longueur, $\mathbb{V}(X)$ représente une longueur au carré, donc il n'est pas possible de comparer directement la position moyenne $\mathbb{E}(X)$ et sa dispersion moyenne $\mathbb{V}(X)$. L'espérance et la variance ne sont pas homogènes. L'écart-type, au contraire, est homogène et donc comparable à l'espérance.

Proposition 6 : Propriétés de la variance

Soient X et Y deux variable aléatoire réelle sur Ω .

- ① **Positivité** $\mathbb{V}(X) \geq 0$.
- ② **Formule de König-Huygens** $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
- ③ **Constance presque sûre** $\mathbb{V}(X) = 0$ si et seulement si $P(X = \mathbb{E}(X)) = 1$.
 X est dite **presque sûrement constante**.
- ④ **Variance d'une transformation affine de X** Si (a, b) est un couple de réels, $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.
- ⑤ **Variance d'une somme de variables aléatoires indépendante** Si Y est indépendante de X , alors $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$.

Preuve.

- ① Immédiat par positivité de l'espérance.
- ② Par linéarité de l'espérance, $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2) = \mathbb{E}(X^2) - 2\mathbb{E}(X) \cdot \mathbb{E}(X) + \mathbb{E}(X)^2$.
- ③ $\mathbb{V}(X) = 0 \Leftrightarrow \mathbb{E}((X - \mathbb{E}(X))^2) = 0 \xrightarrow{\text{Prop. 2}} P(X - \mathbb{E}(X) = 0) = 1$, d'où le résultat.
- ④ $\mathbb{V}(aX + b) = \mathbb{E}(a^2 X^2 + 2abX + b^2) - (a\mathbb{E}(X) + b)^2 = a^2 (\mathbb{E}(X^2) - (\mathbb{E}(X))^2) + 2ab\mathbb{E}(X) + b^2 - 2ab\mathbb{E}(X) - b^2 = a^2 \mathbb{V}(X)$.
- ⑤ $\mathbb{V}(X + Y) = \mathbb{E}((X + Y)^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2$
 $= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)^2$
 $= \mathbb{V}(X) + \mathbb{V}(Y) + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))$
 $= \mathbb{V}(X) + \mathbb{V}(Y)$.

□

Variable aléatoire centrée réduite

Soit X une variable aléatoire réelle non constante. Alors, la variable $Y = \frac{X - \mathbb{E}(X)}{\sigma(X)}$ est centrée et réduite. En effet,

$$\mathbb{E}(Y) = \frac{1}{\sigma(X)}(\mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X))) = 0 \quad \text{et} \quad \mathbb{V}(Y) = \frac{1}{\sigma(X)^2} \mathbb{V}(X) = 1.$$

Définition 3 : Covariance

Soit X et Y deux variables aléatoires réelles. La **covariance de X et Y** est le réel

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

Les variables X et Y sont dites **décorrélées** si $\text{Cov}(X, Y) = 0$.

Proposition 7 : Propriétés de la covariance

Soit X, X', Y et Y' des variables aléatoires réelles sur $(\Omega, \mathbb{P})$ et λ un réel.

- ① **Formule de König-Huygens** $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
- ② **Lien Covariance-Variance** $\text{Cov}(X, X) = \mathbb{V}(X) \geq 0$
 $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2 \text{Cov}(X, Y)$.
- ③ **Symétrie** $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.
- ④ **Bilinéarité** $\text{Cov}(\lambda X + X', Y) = \lambda \text{Cov}(X, Y) + \text{Cov}(X', Y)$
 $\text{Cov}(X, \lambda Y + Y') = \lambda \text{Cov}(X, Y) + \text{Cov}(X, Y')$.

Preuve.

- ① $\text{Cov}(X, Y) = \mathbb{E}(XY - \overbrace{X \mathbb{E}(Y)}^{\text{Cste}} - \overbrace{\mathbb{E}(X) Y}^{\text{Cste}} + \overbrace{\mathbb{E}(X)\mathbb{E}(Y)}^{\text{Cste}})$
 $= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y)$
 $= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
- ② $\mathbb{V}(X + Y) = \mathbb{E}((X - \mathbb{E}(X) + Y - \mathbb{E}(Y))^2)$
 $= \mathbb{E}((X - \mathbb{E}(X))^2 + 2(X - \mathbb{E}(X))(Y - \mathbb{E}(Y)) + (Y - \mathbb{E}(Y))^2)$
 $= \mathbb{V}(X) + 2 \text{Cov}(X, Y) + \mathbb{V}(Y)$.
- ③ Immédiat.
- ④ Calcul de linéarité immédiat, par linéarité de l'espérance.

□

Application bilinéaire

On vient de montrer que l'application $(X, Y) \mapsto \text{Cov}(X, Y)$ est une forme bilinéaire, symétrique et positive sur l'espace vectoriel des variables aléatoires réelles sur $(\Omega, \mathbb{P})$.

L'utilisation de la bilinéarité peut éviter des calculs de covariance souvent pénibles.

⚠ ATTENTION Deux variables indépendantes sont toujours décorréélées, mais la réciproque est fautive en général. Par exemple, pour toute variable aléatoire X de loi uniforme sur $\{-1, 0, 1\}$, on a

$$\text{Cov}(X, X^2) = \mathbb{E}(X^3) - \mathbb{E}(X)\mathbb{E}(X^2) = 0 - 0 \times \mathbb{E}(X^2) = 0,$$

mais X et X^2 ne sont pas indépendantes car $P(X = 0, X^2 = 1) = 0$, alors que $P(X = 0)P(X^2 = 1) = \frac{1}{3} \times \frac{2}{3} \neq 0$.

Exemple 6

Si la variable X est constante, égale à a , on a $\text{Cov}(X, Y) = 0$. En effet, $\mathbb{E}(XY) = \mathbb{E}(aY) = a\mathbb{E}(Y) = \mathbb{E}(X)\mathbb{E}(Y)$.

Proposition 8 : Variance des lois usuelles

Soit X une variable aléatoire complexe.

- ① **Variables aléatoires constantes** X est certaine de valeur c , alors $V(X) = 0$.
- ② **Loi de Bernoulli** Soit $p \in [0, 1]$. Si $X \sim \mathcal{B}(p)$, alors $V(X) = p(1 - p)$.
- ③ **Loi binomiale** Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. Si $X \sim \mathcal{B}(n, p)$, alors $E(X) = np(1 - p)$.

⚠ ATTENTION Il n'y a pas de formule explicite pour la loi uniforme, parce que cela dépend de l'espace d'arrivée de variable aléatoire.

Preuve.

- ① Si $X = c$, alors $\mathbb{E}(X) = c$ et $\mathbb{E}(X^2) = c^2$, donc $V(X) = 0$.
- ② $V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = p - p^2 = p(1 - p)$, car $X^2 = X$, puisque $X(\Omega) = \{0, 1\}$.
- ③ On réécrit $X = \sum_{k=1}^n X_i$, où les X_i sont des variables aléatoires indépendantes de loi $\mathcal{B}(p)$. On a alors

$$V(X) = V\left(\sum_{k=1}^n X_i\right) = \sum_{k=1}^n V(X_i) = \sum_{k=1}^n p(1 - p) = np(1 - p), \text{ par indépendance des } X_i.$$

□

Exemple 7**Problème des rencontres**

Une urne contient n boules numérotées de 1 à n . On les extrait successivement sans remise. On dit qu'il y a rencontre au i -ème tirage si la boule tirée porte le numéro i . Déterminons le nombre moyen de rencontres et la dispersion moyenne. Posons

- $\forall i \in \llbracket 1, n \rrbracket$, X_i la variable qui prend la valeur 1 s'il y a rencontre au i -ème tirage, et 0 sinon.
- $X = \sum_{i=1}^n X_i$.

Il faut donc calculer l'espérance de la variable X . On a d'abord $\mathbb{E}(X_i) = \mathbb{P}(X_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$, car il y a $n!$ tirages possibles et $(n-1)!$ tirages pour lesquels la boule i est tirée au i -ème tirage.

Donc par linéarité, $\mathbb{E}(X) = \sum_{i=1}^n \mathbb{E}(X_i) = \sum_{i=1}^n \frac{1}{n} = n \frac{1}{n} = 1$.

Déterminons maintenant $V(X)$.

D'après sa définition et son espérance, chaque X_i est une variable de Bernoulli de paramètre $\frac{1}{n}$ et donc de variance

$$V(X_i) = \frac{1}{n} \left(1 - \frac{1}{n}\right) = \frac{n-1}{n^2}.$$

Maintenant, pour calculer la variance de la somme des X_i , nous devons calculer $\text{Cov}(X_i, X_j)$.

Soit $(i, j) \in \llbracket 1, n \rrbracket^2$ tel que $i \neq j$. L'événement $\{X_i = 1\} \cap \{X_j = 1\}$ est réalisé s'il y a rencontre aux i -ème et j -ème tirages. Les résultats des autres tirages sont quelconques. Donc $\mathbb{P}(\{X_i = 1\} \cap \{X_j = 1\}) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$.

Mais $X_i X_j$ est aussi une variable de Bernoulli donc $\mathbb{E}(X_i X_j) = \mathbb{P}(X_i X_j = 1) = \mathbb{P}(\{X_i = 1\} \cap \{X_j = 1\}) = \frac{1}{n(n-1)}$.

Cela nous donne $\text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i) \mathbb{E}(X_j) = \frac{1}{n(n-1)} - \frac{1}{n^2} = \frac{1}{n^2(n-1)}$. Appliquons alors la formule de la variance d'une somme

$$\mathbb{V}(X) = \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j) = n\mathbb{V}(X_1) + 2 \frac{n(n-1)}{2} \text{Cov}(X_1, X_2) = \frac{n-1}{n} + \frac{1}{n} = 1.$$

3 INÉGALITÉS

Proposition 9 : Inégalité de Markov

Soit X une variable aléatoire réelle. Pour tout $a > 0$, $\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}$.

Inégalité large

Comme $\{|X| > a\} \subset \{|X| \geq a\}$, on peut remplacer $P(|X| \geq a)$ par $P(|X| > a)$ si on veut.

Preuve. On a $\mathbb{1}_{\{|X| \geq a\}} a \leq \mathbb{1}_{\{|X| \geq a\}} |X| \leq |X|$.

Donc, par croissance de l'espérance $aP(|X| \geq a) = \mathbb{E}(\mathbb{1}_{\{|X| \geq a\}} a) \leq \mathbb{E}(|X|)$. $\square$

Proposition 10 : Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle. Pour tout $a > 0$, $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}$.

Preuve. D'après l'inégalité de Markov

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

$\square$

Exemple 8

On dispose d'une pièce dont la probabilité d'obtention de pile est notée p . Pour connaître p , on lance cette pièce n fois et on note F la fréquence d'apparition de pile ainsi obtenue. Avec l'inégalité de Bienaymé-Tchebychev, on peut calculer à partir de quelle valeur de n la probabilité pour que F soit une approximation de p à 10^{-2} près est-elle supérieure à 0,9.

Notons N le nombre de piles obtenus, de sorte que $F = \frac{N}{n}$.

Par indépendance des lancers $N \sim \mathcal{B}(n, p)$, donc $\mathbb{E}(N) = np$, et on a $\mathbb{E}(F) = p$.

D'après l'inégalité de Bienaymé-Tchebychev

$$P(|F - p| \geq 10^{-2}) = P(|N - \mathbb{E}(N)| \geq 10^{-2}n) \leq \frac{\mathbb{V}(N)}{(10^{-2}n)^2} = \frac{np(1-p)}{10^{-4}n^2} = \frac{10^4}{n} \times p(1-p) \leq \frac{10^4}{n} \times \frac{1}{4} = \frac{2500}{n}$$

$$\text{Or } P(|F - p| < 10^{-2}) \geq \frac{9}{10} \iff P(|F - p| \geq 10^{-2}) \leq \frac{1}{10}.$$

Par conséquent, il nous suffit de choisir n avec la condition $\frac{2500}{n} \leq \frac{1}{10}$ pour garantir que F soit une approximation de p à 10^{-2} près.

Il faut donc quand même lancer 25000 fois la pièce !