

Chapitre 30

Variables aléatoires sur un univers fini

Dans l'étude d'une expérience aléatoire, il est fréquent qu'on s'intéresse à une quantité numérique qui dépend de l'issue d'une expérience, plutôt qu'à l'issue de l'expérience elle-même.

À titre d'exemple, on considère le jeu suivant, qui consiste à faire un lancer de dé équilibré à 6 faces, avec les gains :

- si le résultat est 1, 2 ou 3 : le joueur perd 1€,
- si le résultat est 4 ou 5 : le joueur ne perd rien, et ne gagne rien,
- si le résultat est 6 : le joueur gagne 2€.

On décrit la situation à l'aide d'une *fonction*, notée X , qui associe à chaque issue $\omega \in \llbracket 1, 6 \rrbracket$ possible une quantité réelle qui correspond au gain algébrique. Ici, $X(\omega) = -1$ si $\omega \in \{1, 2, 3\}$, $X(\omega) = 0$ si $\omega \in \{4, 5\}$, et $X(\omega) = 2$ si $\omega = 6$.

Pour l'étude, c'est la donnée de $\mathbb{P}(\{\omega \in \Omega, X(\omega) = -1\})$, $\mathbb{P}(\{\omega \in \Omega, X(\omega) = 0\})$ et $\mathbb{P}(\{\omega \in \Omega, X(\omega) = 2\})$ qui importe plus que la donnée de $\mathbb{P}$ elle-même. Nous allons voir que cette donnée, qui s'appelle la loi de X , suffit pour l'étude présente.

I Variable aléatoire et loi

1. Définition

Définition - Variable aléatoire

Soit Ω un ensemble fini et E un ensemble. On appelle *variable aléatoire* sur Ω une application X de Ω dans E . On parle de variable aléatoire réelle lorsque $E \subset \mathbb{R}$.

Remarques.

- Si X est une variable aléatoire sur un univers Ω fini, alors l'ensemble image $X(\Omega)$, parfois appelé univers image de X , est fini également.
- Une variable aléatoire est souvent décrite par la quantité qu'elle représente, qui est fonction de l'issue de l'expérience.
 - ♦ Dans le cas du jeu de dé de début de chapitre, X désigne le gain du joueur, et on a $X(\Omega) = \{-1, 0, 2\}$.
 - ♦ Si on effectue deux lancers de dés à 6 faces, on peut choisir $\Omega = \llbracket 1, 6 \rrbracket^2$. La variable aléatoire correspondant à la somme des valeurs obtenues sur les dés est alors $X : (i, j) \in \Omega \mapsto i + j$, et $X(\Omega) = \llbracket 2, 12 \rrbracket$.
- Si X une variable aléatoire sur Ω à valeurs dans E , alors pour toute fonction $f : E \rightarrow F$, la fonction $f \circ X$, notée $f(X)$ définit une variable aléatoire réelle sur Ω à valeurs dans F .

Notation - Événements liés à une variable aléatoire

Soient Ω un univers fini et X une variable aléatoire définie sur Ω à valeurs dans un ensemble E . Si $x \in E$ et $A \subset E$, on introduit les notations suivantes pour les événements liés à X .

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

Si $E = \mathbb{R}$ et $x \in \mathbb{R}$, on note aussi $\{X \leq x\} = X^{-1}([-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\}$. On introduit par ailleurs les notations analogues $\{X < x\}$, $\{X \geq x\}$ et $\{X > x\}$.

Si $\mathbb{P}$ est une probabilité sur Ω , on note généralement $\mathbb{P}(X = x)$ au lieu de $\mathbb{P}(\{X = x\})$. On note de même $\mathbb{P}(X \leq x)$, $\mathbb{P}(X \in A)$, etc.

Exemples.

- *Exemple du jeu de dé* : on a $\{X = -1\} = \{1, 2, 3\}$, $\{X > 0\} = \{6\}$.
- *Exemple de la somme des deux dés* : on a $\{X = 4\} = \{(1, 3), (2, 2), (3, 1)\}$, $\{X \leq 3\} = \{(1, 1), (1, 2), (2, 1)\}$.

Théorème - Système complet d'événements associé à une variable aléatoire

Si Ω est un univers fini et X une variable aléatoire sur Ω , alors $(\{X = x\})_{x \in \Omega}$ est un système complet d'événement.

Démonstration. Si $\omega \in \Omega$, on note $x = X(\omega)$, et on a alors $\omega \in \{X = x\}$. Ceci montre que $\Omega = \bigcup_{x \in \Omega} \{X = x\}$ car l'autre inclusion est claire.

Par ailleurs, si $\omega \in \{X = x\} \cap \{X = y\}$, alors $x = X(\omega) = y$, donc les ensembles $\{X = x\}$ sont disjoints deux à deux, ce qui conclut. $\square$

2. Loi d'une variable aléatoire**Définition-théorème - Loi d'une variable aléatoire**

Soient $(\Omega, \mathbb{P})$ un espace probabilisé fini et X une variable aléatoire sur Ω . On appelle *loi* de la variable aléatoire X l'application $\mathbb{P}_X : A \mapsto \mathbb{P}(X \in A)$, qui définit alors une probabilité sur $X(\Omega)$.

Si X et Y sont deux variables aléatoires ayant même loi : $\mathbb{P}_X = \mathbb{P}_Y$, on note $X \sim Y$.

Démonstration. On a $\mathbb{P}_X(\Omega) = \mathbb{P}(X \in \Omega) = \mathbb{P}(\Omega) = 1$. Par ailleurs, si A, B sont deux parties disjointes de $X(\Omega)$, alors $\mathbb{P}_X(A \sqcup B) = \mathbb{P}(X \in A \sqcup B) = \mathbb{P}(\{X \in A\} \sqcup \{X \in B\}) = \mathbb{P}(X \in A) + \mathbb{P}(X \in B) = \mathbb{P}_X(A) + \mathbb{P}_X(B)$. $\square$

Théorème - Loi et distribution de probabilité

Soient $(\Omega, \mathbb{P})$ un espace probabilisé fini et X une variable aléatoire réelle sur Ω . La loi $\mathbb{P}_X$ est entièrement déterminée par la donnée de $(\mathbb{P}(X = x))_{x \in \Omega}$. Plus précisément, si $A \subset X(\Omega)$ alors

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

Démonstration. Si $A \subset X(\Omega)$, alors $A = \bigsqcup_{x \in A} \{x\}$, donc $\mathbb{P}(A) = \sum_{x \in A} \mathbb{P}_X(\{x\})$ $\square$

Remarques.

- On a en particulier $\sum_{x \in \Omega} \mathbb{P}(X = x) = 1$.
- Pour donner la loi de X dans le cas où Ω est fini, on se contentera généralement de donner :
 - ◊ son ensemble image $X(\Omega)$,
 - ◊ les probabilités $\mathbb{P}(X = x)$ pour tout $x \in X(\Omega)$.

Par ce qui précède, si $\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = 1$, alors ceci définit bien une loi de probabilité, de manière unique.

Exemples.

- *Jeu de dés.* On a vu que $X(\Omega) = \{-1, 0, 2\}$. On a

$$\mathbb{P}(X = -1) = \mathbb{P}(\{1, 2, 3\}) = \frac{1}{2}, \quad \mathbb{P}(X = 0) = \mathbb{P}(\{4, 5\}) = \frac{1}{3}, \quad \mathbb{P}(X = 2) = \mathbb{P}(\{6\}) = \frac{1}{6}.$$

- *Somme de deux dés.* On a vu que $X(\Omega) = \llbracket 2, 12 \rrbracket$. On a

$$\begin{aligned} \mathbb{P}(X = 2) &= \frac{1}{36}, \quad \mathbb{P}(X = 3) = \frac{2}{36}, \quad \mathbb{P}(X = 4) = \frac{3}{36}, \quad \mathbb{P}(X = 5) = \frac{4}{36}, \quad \mathbb{P}(X = 6) = \frac{5}{36}, \quad \mathbb{P}(X = 7) = \frac{6}{36} \\ \mathbb{P}(X = 8) &= \frac{5}{36}, \quad \mathbb{P}(X = 9) = \frac{4}{36}, \quad \mathbb{P}(X = 10) = \frac{3}{36}, \quad \mathbb{P}(X = 11) = \frac{2}{36}, \quad \mathbb{P}(X = 12) = \frac{1}{36}. \end{aligned}$$

Théorème

Soient Ω un univers fini et X, Y des variables aléatoires sur Ω à valeurs dans un ensemble E . Si $X \sim Y$, alors pour toute fonction $f : E \rightarrow F$, on a $f(X) \sim f(Y)$.

Autrement dit, la loi de $f(X)$ est entièrement déterminée par la loi de X .

Démonstration. Soit $y \in f(X(\Omega))$, on remarque que $\{f(X) = y\} = \bigsqcup_{x \in f^{-1}(\{y\})} \{X = x\}$, donc

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

□

Définition - Loi conditionnelle

Soient X une variable aléatoire sur un univers fini Ω et un événement $B \in \mathcal{P}(\Omega)$ tel que $\mathbb{P}(B) \neq 0$. On appelle *loi conditionnelle* de X sachant B la probabilité sur $\mathcal{P}(X(\Omega))$ définie par

$$\mathbb{P}_X(\cdot | B) : A \mapsto \mathbb{P}_X(A|B).$$

3. Lois usuelles**Définition - Loi certaine**

On dit qu'une variable aléatoire X est certaine si X est constante, c'est-à-dire si $X(\Omega)$ est un singleton.

Remarque. Si X est une variable aléatoire sur un espace probabilisé $(\Omega, \mathbb{P})$ telle qu'il existe $x \in X(\Omega)$ vérifiant $\mathbb{P}(X = x) = 1$, on dit que X est une variable aléatoire quasi-certaine. Bien sûr, si X est certaine, elle est quasi-certaine.

Définition - Loi uniforme

Soient Ω un univers fini et E un ensemble. On dit qu'une variable aléatoire X sur Ω suit la loi *uniforme* sur E si X est à valeurs dans E , et pour tout $x \in E$, $\mathbb{P}(X = x) = \frac{1}{|E|}$. Autrement dit, $\mathbb{P}_X$ est la probabilité uniforme sur E :

$$\forall A \in \mathcal{P}(E), \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \frac{|A|}{|E|}.$$

On note alors $X \sim \mathcal{U}(E)$.

Exemple. Si X est la variable aléatoire qui désigne le résultat d'un dé équilibré à 6 faces, alors $X \sim \mathcal{U}(\llbracket 1, 6 \rrbracket)$.

Définition - Loi de Bernoulli

Soient Ω un univers fini et $p \in [0, 1]$. On dit qu'une variable aléatoire X sur Ω suit la loi de Bernoulli de paramètre p si X est à valeurs dans $\{0, 1\}$ et $\mathbb{P}(X = 1) = p$. On note alors $X \sim \mathcal{B}(p)$.

Remarque. Une expérience aléatoire qui n'a que deux issues possibles, appelée *épreuve de Bernoulli*, peut toujours être modélisée à l'aide d'une variable X suivant la loi de Bernoulli. On appelle souvent Échec et Succès les deux issues de Ω , et X associe 0 à l'issue Échec et 1 à l'issue Succès.

Exemple. Si $(\Omega, \mathbb{P})$ est un espace probabilisé et $A \in \mathcal{P}(\Omega)$, la variable aléatoire

$$\mathbb{1}_A : \omega \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

suit la loi de Bernoulli de paramètre $\mathbb{P}(A)$: $\mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(\{\omega \in \Omega, X(\omega) = 1\}) = \mathbb{P}(A)$ et il est clair que X est à valeurs dans $\{0, 1\}$.

Remarques.

- Il faut comprendre la variable aléatoire $\mathbb{1}_A$ comme la variable aléatoire qui vaut 1 si A est réalisé, et 0 sinon.
- Une variable aléatoire qui désigne un nombre de succès lors d'une série d'expériences peut toujours être écrite comme une somme d'indicatrices.

Exemple. On lance n fois une pièce équilibrée, et pour tout $k \in \llbracket 1, n \rrbracket$, on note A_k l'événement "on obtient Pile au k -ème lancer". Si X est la variable aléatoire qui désigne le nombre de Pile obtenus, on peut écrire

$$X = \mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}.$$

Définition-théorème - Loi binomiale

Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. On dit qu'une variable aléatoire X suit la loi binomiale de paramètres n et p si $X(\Omega) = \llbracket 0, n \rrbracket$, et

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

On définit ainsi une loi de probabilité notée $\mathcal{B}(n, p)$, et on note alors $X \sim \mathcal{B}(n, p)$.

Démonstration. On a $\sum_{k=0}^n \mathbb{P}(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1$ par le binôme de Newton. $\square$

**Expérience associée à la loi binomiale**

Si on reproduit n épreuves de Bernoulli identiques et indépendantes, ayant pour probabilité de succès p , la variable aléatoire X saignant le nombre de succès obtenus suit la loi $\mathcal{B}(n, p)$.

Démonstration. Pour tout $i \in \llbracket 1, n \rrbracket$, notons S_i l'événement "on obtient un succès à l'épreuve i ". Pour tout $k \in \llbracket 0, n \rrbracket$,

$$\{X = k\} = \bigsqcup_{I \in \mathcal{P}_k(\llbracket 1, n \rrbracket)} \bigcap_{i \in I} S_i \cap \bigcap_{j \in \bar{I}} \bar{S}_j, \quad \text{où } \bar{I} = \llbracket 1, n \rrbracket \setminus I.$$

Si $I \in \mathcal{P}_k(\llbracket 1, n \rrbracket)$, par indépendance des S_i , $\mathbb{P}\left(\bigcap_{i \in I} S_i \cap \bigcap_{j \in \bar{I}} \bar{S}_j\right) = \prod_{i \in I} \mathbb{P}(S_i) \prod_{j \in \bar{I}} \mathbb{P}(\bar{S}_j) = p^k (1-p)^{n-k}$, donc

$$\mathbb{P}(X = k) = \sum_{I \in \mathcal{P}_k(\llbracket 1, n \rrbracket)} \mathbb{P}\left(\bigcap_{i \in I} S_i \cap \bigcap_{j \in \bar{I}} \bar{S}_j\right) = \sum_{I \in \mathcal{P}_k(\llbracket 1, n \rrbracket)} p^k (1-p)^{n-k} = \binom{n}{k} p^k (1-p)^{n-k},$$

ce qui conclut. $\square$

Remarque. Si $p \in [0, 1]$, la loi de Bernoulli $\mathcal{B}(p)$ n'est autre que la loi binomiale $\mathcal{B}(1, p)$.

II Espérance et variance

1. Définition

Définition - Espérance

Soient $(\Omega, \mathbb{P})$ un espace probabilisé et X une variable aléatoire réelle sur Ω . On appelle *espérance* de X le réel

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

Dans le cas où $\mathbb{E}(X) = 0$, on dit que la variable aléatoire X est *centrée*.

Exemples.

- Si X est une variable quasi-certaine et si $c \in X(\Omega)$ est tel que $\mathbb{P}(X = c) = 1$, alors $\mathbb{E}(X) = c$.
- Si $p \in [0, 1]$ et $X \sim \mathcal{B}(p)$, alors $\mathbb{E}(X) = 0 \mathbb{P}(X = 0) + 1 \mathbb{P}(X = 1) = p$.

Remarques.

- Si $X(\Omega) = \{x_1, \dots, x_n\}$ et $p_i = \mathbb{P}(X = x_i)$ pour tout $i \in \llbracket 1, n \rrbracket$, alors

$$\mathbb{E}(X) = \sum_{i=1}^n x_i p_i = x_1 p_1 + \dots + x_n p_n.$$

Il s'agit de la moyenne des réels x_i pondérée par les probabilités $\mathbb{P}(X = x_i)$. L'espérance de X est donc intuitivement la "valeur moyenne" prise par X .

- On remarque que l'espérance d'une variable aléatoire X ne dépend que de sa loi : deux variables aléatoires ayant même loi ont aussi même espérance. On parle donc de l'espérance d'une loi de probabilité donnée.

Exemple. *Jeu de dés.* Si X désigne le gain du joueur dans le jeu présenté plus haut, alors

$$\mathbb{E}(X) = -1 \mathbb{P}(X = -1) + 0 \mathbb{P}(X = 0) + 2 \mathbb{P}(X = 2) = -\frac{1}{2} + 2 \times \frac{1}{6} = -\frac{1}{6}.$$

2. Espérance des lois usuelles

Théorème - Espérance de la loi $\mathcal{U}(E)$

Si E est un ensemble fini de $\mathbb{R}$ dont les éléments sont $x_1, \dots, x_n$ et X une variable aléatoire suivant la loi uniforme sur E , alors

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

Démonstration. Si $X \sim \mathcal{U}(E)$, alors on a $\mathbb{E}(X) = \sum_{i=1}^n x_i \mathbb{P}(X = x_i) = \sum_{i=1}^n x_i \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n x_i$. $\square$

Exemple. Si $n \in \mathbb{N}^*$ et $X \sim \mathcal{U}([1, n])$, alors $\mathbb{E}(X) = \frac{n+1}{2}$:

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

En particulier, si X est la variable désignant le résultat d'un lancer de dé à 6 faces équilibré, alors $\mathbb{E}(X) = \frac{7}{2}$.

Théorème - Espérance de la loi $\mathcal{B}(p)$

Si $p \in [0, 1]$ et X est une variable aléatoire sur Ω qui suit la loi $\mathcal{B}(p)$, alors $\mathbb{E}(X) = p$.

En particulier, si $A \in \mathcal{P}(\Omega)$, alors $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

Démonstration. Nous avons déjà vu plus haut que si $X \sim \mathcal{B}(p)$, alors $\mathbb{E}(X) = p$. Par ailleurs, si $A \in \mathcal{P}(\Omega)$, on sait que $\mathbb{1}_A \sim \mathcal{B}(\mathbb{P}(A))$, ce qui donne $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$. $\square$

Théorème - Espérance de la loi $\mathcal{B}(n, p)$

Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. Si $X \sim \mathcal{B}(n, p)$, alors $\mathbb{E}(X) = np$.

Démonstration. On rappelle que d'après la formule du capitaine $k \binom{n}{k} = n \binom{n-1}{k-1}$ pour tout $k \in [1, n]$. Ainsi

$$\mathbb{E}(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = n \sum_{k=1}^n \binom{n-1}{k-1} p^k (1-p)^{n-k} \stackrel{k'=k-1}{=} n \sum_{k'=0}^{n-1} \binom{n-1}{k'} p^{k'+1} (1-p)^{n-(k'+1)}.$$

Finalement, par la formule du binôme de Newton, on a $\mathbb{E}(X) = np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = np$. $\square$

3. Propriétés de l'espérance

Lemme

Soient $(\Omega, \mathbb{P})$ un espace probabilisé et X une variable aléatoire réelle sur Ω , on a

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

Démonstration. On somme par paquets dans la somme sur $\omega \in \Omega$: on sait que $\Omega = \bigsqcup_{x \in X(\Omega)} \{X = x\}$, donc on peut regrouper les $\omega \in \Omega$ ayant même image par X dans la somme :

$$\sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) = \sum_{x \in X(\Omega)} \sum_{\omega \in \{X=x\}} \underbrace{X(\omega)}_{=x} \mathbb{P}(\{\omega\}) = \sum_{x \in X(\Omega)} x \sum_{\omega \in \{X=x\}} \mathbb{P}(\{\omega\}) = \sum_{x \in X(\Omega)} x \mathbb{P}(\{X=x\}) = \mathbb{E}(X). \quad \square$$

Théorème - Propriétés de l'espérance

Soient Ω un univers fini, X, Y des variables aléatoires réelles sur Ω et $\lambda, \mu \in \mathbb{R}$.

i. *Linéarité.* $\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$.

ii. *Positivité.* Si X est positive, c'est-à-dire $\forall \omega \in \Omega, X(\omega) \geq 0$, alors $\mathbb{E}(X) \geq 0$.

- iii. *Croissance.* Si $X \leq Y$, c'est-à-dire $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.
 iv. *Inégalité triangulaire.* $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$.

Démonstration. Nous allons utiliser à plusieurs reprises le lemme ci-dessus.

i. D'après le lemme, on a

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

ii. Si X est positive, alors $\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) \geq 0$.

iii. Si $X \leq Y$, alors la variable aléatoire $Y - X$ est positive, donc $\mathbb{E}(Y - X) \geq 0$. On a alors $\mathbb{E}(Y) - \mathbb{E}(X) \geq 0$ par linéarité, ce qui conclut.

iv. On a $|\mathbb{E}(X)| = \left| \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) \right| \leq \sum_{\omega \in \Omega} |X(\omega)| \mathbb{P}(\{\omega\}) = \mathbb{E}(|X|)$. $\square$

Définition-théorème - Variable aléatoire centrée associée à X

Si X est une variable aléatoire réelle alors la variable aléatoire $X - \mathbb{E}(X)$ est centrée. On dit qu'il s'agit de la variable aléatoire centrée associée à X .

Démonstration. Par linéarité de l'espérance, on a $\mathbb{E}(X - \mathbb{E}(X)) = \mathbb{E}(X) - \mathbb{E}(\mathbb{E}(X))$. La variable aléatoire $\mathbb{E}(X)$ étant certaine, on a $\mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X)$, ce qui donne $\mathbb{E}(X - \mathbb{E}(X)) = 0$. $\square$

Théorème - Théorème de transfert

Soient Ω un univers fini et X une variable aléatoire sur Ω . Pour toute fonction $f : X(\Omega) \rightarrow \mathbb{R}$, on a

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

Démonstration. Par le lemme ci-dessus, on a $\mathbb{E}(f(X)) = \sum_{\omega \in \Omega} f(X(\omega)) \mathbb{P}(\{\omega\})$, donc comme $\Omega = \bigsqcup_{x \in X(\Omega)} \{X = x\}$,

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

Exemple. Si X est une variable aléatoire sur un univers fini Ω , alors $\mathbb{E}(X^2) = \sum_{x \in X(\Omega)} x^2 \mathbb{P}(X = x)$.

4. Variance d'une variable aléatoire

Définition - Variance et écart-type d'une variable aléatoire finie

Soit X une variable aléatoire réelle finie. On définit la *variance* de X , notée $\mathbb{V}(X)$, par

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

On appelle *écart-type* de X , noté $\sigma(X)$, le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Si $\mathbb{V}(X) = 1$, on dit que la variable aléatoire X est *réduite*.

Remarques.

- La variance est toujours positive, en tant qu'espérance d'une variable aléatoire positive.
- La variance et l'écart-type mesurent la dispersion des valeurs la variable aléatoire autour de son espérance. Si la variance est faible, les valeurs sont regroupées autour de l'espérance. Si la variance est élevée, les valeurs sont très dispersées.
- Par le théorème de transfert,

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

Comme l'espérance, la variance d'une variable aléatoire ne dépend que de sa loi : deux variables aléatoires de même loi ont aussi même variance. On parle donc de la variance d'une loi de probabilité donnée.

Théorème - Formule de Koenig-Huygens

Soit X une variable aléatoire réelle finie. On a $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Démonstration. Par la linéarité de l'espérance, on a

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

Comme $\mathbb{E}(2\mathbb{E}(X)X) = 2\mathbb{E}(X)\mathbb{E}(X)$, on a $\mathbb{V}(X) = \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. $\square$

5. Propriétés de la variance**Théorème - Variance de $aX + b$**

Si X est une variable aléatoire réelle sur un univers fini et $a, b \in \mathbb{R}$, alors $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.

On a donc aussi $\sigma(aX + b) = |a| \sigma(X)$.

Démonstration. On a $\mathbb{V}(aX + b) = \mathbb{E}((aX + b - \mathbb{E}(aX + b))^2) = \mathbb{E}((aX + b - a\mathbb{E}(X) - b)^2)$. Ainsi,

$$\mathbb{V}(aX + b) = \mathbb{E}((a(X - \mathbb{E}(X)))^2) = \mathbb{E}(a^2(X - \mathbb{E}(X))^2) = a^2 \mathbb{E}((X - \mathbb{E}(X))^2) = a^2 \mathbb{V}(X). \quad \square$$

Remarque. En particulier, si $a, b \in \mathbb{R}$, alors $\diamond \mathbb{V}(aX) = a^2 \mathbb{V}(X)$, et donc $\mathbb{V}(-X) = \mathbb{V}(X)$,
 $\diamond \mathbb{V}(X + b) = \mathbb{V}(X)$.

Théorème - Variable aléatoire centrée réduite

Si X est une variable aléatoire réelle avec $\mathbb{V}(X) \neq 0$, alors la variable aléatoire

$$\frac{X - \mathbb{E}(X)}{\sigma(X)}$$

est centrée réduite et est appelée *variable aléatoire centrée réduite associée à X* .

Démonstration. On note $Z = \frac{X - \mathbb{E}(X)}{\sigma(X)}$. Il suffit de remarquer que $\mathbb{E}(Z) = \frac{1}{\sigma(X)} \mathbb{E}(X - \mathbb{E}(X)) = 0$, et :

$$\mathbb{V}(Z) = \frac{1}{\sigma(X)^2} \mathbb{V}(X - \mathbb{E}(X)) = \frac{1}{\sigma(X)^2} \mathbb{V}(X) = 1. \quad \square$$

Théorème - Variance nulle

Si X est une variable aléatoire sur un univers fini, alors $\mathbb{V}(X) = 0$ si et seulement si X est quasi-certaine.

Démonstration. Par le théorème de transfert, on a $\mathbb{V}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x)$. Ainsi, si $\mathbb{V}(X) = 0$, on a $\mathbb{P}(X = x) = 0$ pour tout $x \neq \mathbb{E}(X)$.

Comme $(\{X = x\})_{x \in X(\Omega)}$ est un système complet d'événements, ceci entraîne que $\mathbb{P}(X = \mathbb{E}(X)) = 1$, et la variable aléatoire X est quasi-certaine.

Réciproquement, si X est quasi-certaine, alors $\mathbb{P}(X = \mathbb{E}(X)) = 1$, donc d'après l'égalité ci-dessus, $\mathbb{V}(X) = 0$. $\square$

Théorème - Variance des lois de Bernoulli et binomiales

Soient $n \in \mathbb{N}^*$, $p \in [0, 1]$ et X une variable aléatoire.

- Si $X \sim \mathcal{B}(p)$, alors $\mathbb{V}(X) = p(1 - p)$.
- Si $X \sim \mathcal{B}(n, p)$, alors $\mathbb{V}(X) = np(1 - p)$.

Démonstration. Traitons le cas binomial, le cas Bernoulli en étant un cas particulier :

$$\mathbb{E}(X^2) = \sum_{k=1}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} + \underbrace{\sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k}}_{=\mathbb{E}(X)=np}$$

↑ ou 2

Or la formule du capitaine donne $k(k-1) \binom{n}{k} = n(n-1) \binom{n-1}{k-1} = n(n-1) \binom{n-2}{k-2}$. Ainsi

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

Le binôme de Newton donne alors $\mathbb{E}(X^2) = n(n-1)p^2 + np$.

Finalement, $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = n(n-1)p^2 + np - n^2p^2 = -np^2 + np = np(1-p)$, ce qui conclut. $\square$

Remarque. Nous verrons plus loin une preuve plus directe, basée sur l'écriture de X comme une somme de variables de Bernoulli indépendantes.

III Couples de variables aléatoires

Nous nous intéressons ici aux couples (X, Y) de variables aléatoires, voire quand nous l'indiquons aux n -uplets $(X_1, \dots, X_n)$ de variables aléatoires.

Remarquons pour commencer que si $X : \Omega \rightarrow E$ et $Y : \Omega \rightarrow F$ sont deux variables aléatoires sur le même univers, alors le couple (X, Y) est encore une variable aléatoire, cette fois à valeurs dans $E \times F$. Il en va bien sûr de même des n -uplets $(X_1, \dots, X_n)$ de variables aléatoires.

1. Loi conjointe, lois marginales

Définition - Couple de variables aléatoires

Si X, Y sont deux variables aléatoires sur un même univers fini Ω à valeurs dans des ensembles respectifs E et F , alors $(X, Y) : \omega \mapsto (X(\omega), Y(\omega))$ est une variable aléatoire sur Ω à valeurs dans $E \times F$ appelée couple des variables aléatoires X et Y .

Notation. Si (X, Y) est un couple de variables aléatoires sur un univers Ω et (x, y) est un élément de l'univers image $(X, Y)(\Omega)$, on note $\{X = x, Y = y\}$ l'événement $\{X = x\} \cap \{Y = y\}$.

Si $\mathbb{P}$ est une probabilité sur Ω , on note par ailleurs $\mathbb{P}(X = x, Y = y) = \mathbb{P}(\{X = x\} \cap \{Y = y\})$, et plus généralement, si $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(\{X \in A\} \cap \{Y \in B\})$.

Définition-théorème - Loi conjointe d'un couple – lois marginales

Si (X, Y) est un couple de variables aléatoires sur un univers Ω , alors la loi de (X, Y) , appelée *loi conjointe* de X et Y est entièrement déterminée par la donnée de la famille $(\mathbb{P}(X = x, Y = y))_{(x,y) \in (X,Y)(\Omega)}$.

On appelle par ailleurs *lois marginales* de (X, Y) les lois de X et Y .

Démonstration. Si l'on se donne $(\mathbb{P}(X = x, Y = y))_{(x,y) \in (X,Y)(\Omega)}$, alors pour tout $A \subset (X, Y)(\Omega)$,

$$\mathbb{P}_{(X,Y)}(A) = \mathbb{P}((X, Y) \in A) = \sum_{(x,y) \in A} \mathbb{P}(X = x, Y = y). \quad \square$$

Exemple. On lance successivement deux dés équilibrés à 6 faces, on note X la variable aléatoire correspondant au plus petit des deux résultats obtenus, et Y celle correspondant au plus grand. Pour tout $(i, j) \in \llbracket 1, 6 \rrbracket^2$, si on note D_1 la variable aléatoire donnant le résultat du premier dé, et D_2 celle donnant le résultat du second dé, alors :

$$\mathbb{P}(X = i, Y = j) = \begin{cases} 0 & \text{si } i > j, \\ \mathbb{P}(D_1 = i, D_2 = i) & \text{si } i = j, \\ \mathbb{P}(D_1 = i, D_2 = j) + \mathbb{P}(D_1 = j, D_2 = i) & \text{si } i < j. \end{cases}$$

Comme D_1 et D_2 sont des variables aléatoires indépendantes, on a $\mathbb{P}([X = i] \cap [Y = j]) = \begin{cases} 0 & \text{si } i > j, \\ \frac{1}{36} & \text{si } i = j, \\ \frac{1}{18} & \text{si } i < j. \end{cases}$

Nous allons voir qu'on peut obtenir les lois de X et Y (lois marginales) à partir de la loi conjointe ou de lois conditionnelles. Il faut retenir que tout repose naturellement sur la formule des probabilités totales.

Théorème - Lois marginales à partir de la loi conjointe

Si (X, Y) est un couple de variable aléatoire sur un espace probabilisé $(\Omega, \mathbb{P})$. Les lois marginales sont entièrement déterminées par la loi conjointe $\mathbb{P}_{(X,Y)}$:

$$\forall x \in X(\Omega), \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y) \quad \text{et} \quad \forall y \in Y(\Omega), \mathbb{P}(Y = y) = \sum_{x \in X(\Omega)} \mathbb{P}(X = x, Y = y).$$

Démonstration. Montrons la première égalité, la deuxième est similaire. Si $x \in X(\Omega)$, on écrit la formule des probabilités totales avec le système complet d'événements $(\{Y = y\})_{y \in Y(\Omega)}$:

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

Remarque. On peut aussi trouver les lois marginales à partir de la connaissance de lois conditionnelles : si on connaît la loi de Y et la loi de X sachant $\{Y = y\}$ pour tout $y \in Y(\Omega)$, alors toujours par la formule des probabilités totales,

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

Un couple (X, Y) étant aussi une variable aléatoire, le théorème de transfert s'étend immédiatement à ce cadre.

Théorème - Théorème de transfert pour un couple

Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé $(\Omega, \mathbb{P})$. Pour toute fonction $f : (X, Y)(\Omega) \rightarrow \mathbb{R}$, on a

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

Exemple. Si X, Y sont des variables aléatoires réelles sur un espace probabilisé $(\Omega, \mathbb{P})$, alors

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

Remarque. Plus généralement, on peut appliquer la formule de transfert au cas de n -uplets de variables aléatoires sur un même espace probabilisé.

2. Variables aléatoires indépendantes

Définition - Variables aléatoires indépendantes

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathbb{P})$. On dit que X et Y sont indépendantes si pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B).$$

On note alors $X \perp\!\!\!\perp Y$.

Remarques.

- De manière équivalente, les variables aléatoires X et Y de l'énoncé sont indépendantes si pour tous $x \in X(\Omega)$ et $y \in Y(\Omega)$,

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \mathbb{P}(Y = y).$$

- Cette définition se généralise au cas de n variables aléatoires $X_1, \dots, X_n$ sur un même espace probabilisé : on dit que $X_1, \dots, X_n$ sont (mutuellement) indépendantes si pour tous $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,

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

Théorème - Somme de lois de Bernoulli indépendantes

Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. Si pour tout $i \in \llbracket 1, n \rrbracket$, $X_i \sim \mathcal{B}(p)$ et si $X_1, \dots, X_n$ sont indépendantes, alors $X_1 + \dots + X_n \sim \mathcal{B}(n, p)$.

Démonstration. Ce résultat a déjà été montré pour décrire l'expérience associée à la loi binomiale. Il suffit de poser $S_i = \{X_i = 1\}$, et on retrouve la preuve rencontrée plus haut : pour tout $i \in \llbracket 1, n \rrbracket$, on a $\mathbb{P}(S_i) = p$, et les S_i sont indépendants par indépendance des X_i . $\square$

Théorème - Espérance d'un produit de variables aléatoires indépendantes

Si X et Y sont deux variables aléatoires réelles indépendantes sur un même espace probabilisé $(\Omega, \mathbb{P})$, alors

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

Démonstration. Par le théorème de transfert : on a

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

$$\text{Ainsi, } \mathbb{E}(XY) = \sum_{x \in X(\Omega)} x \mathbb{P}(X = x) \sum_{y \in Y(\Omega)} y \mathbb{P}(Y = y) = \mathbb{E}(X) \mathbb{E}(Y). \quad \square$$

! La réciproque est fautive : on peut avoir $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$, mais X et Y ne sont pas indépendantes.

Exemple. Si $X(\Omega) = \{-1, 0, 1\}$ et $\mathbb{P}(X = -1) = \mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{3}$, et si $Y = \mathbb{1}_{\{X=0\}}$,

– $\mathbb{E}(X) = 0$, et $\mathbb{E}(XY) = 0$ car $\mathbb{P}(XY = 0) = 1$, donc $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y) = 0$.

– $\mathbb{P}(X = 0, Y = 0) = 0$ mais $\mathbb{P}(X = 0) \mathbb{P}(Y = 0) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9} \neq 0$, donc X et Y ne sont pas indépendantes.

Le résultat ci-dessous permet de répondre à des questions du type : si X, Y, Z sont trois variables aléatoires indépendantes sur un même espace probabilisé, les variables aléatoires $X + Y$ et Z sont-elles indépendantes ? (Oui !)

Théorème - Lemme des coalitions

Soient $X_1, \dots, X_n$ des variables aléatoires sur un même espace probabilisé $(\Omega, \mathbb{P})$, et deux fonctions $f : \mathbb{R}^k \rightarrow \mathbb{R}$, $g : \mathbb{R}^{n-k} \rightarrow \mathbb{R}$. Si $X_1, \dots, X_n$ sont indépendantes, alors $f(X_1, \dots, X_k)$ et $g(X_{k+1}, \dots, X_n)$ sont indépendantes.

Démonstration.

– Montrons d'abord que si $X \perp\!\!\!\perp Y$ et $f : X(\Omega) \rightarrow E$, $g : Y(\Omega) \rightarrow F$, alors $f(X) \perp\!\!\!\perp g(Y)$. Pour tous $u \in E$ et $v \in F$, on a

$$\mathbb{P}(f(X) = u, g(Y) = v) = \mathbb{P}(X \in f^{-1}(u), Y \in g^{-1}(\{v\})) = \mathbb{P}(X \in f^{-1}(u)) \mathbb{P}(Y \in g^{-1}(\{v\}))$$

par indépendance, donc $\mathbb{P}(f(X) = u, g(Y) = v) = \mathbb{P}(f(X) = u) \mathbb{P}(g(Y) = v)$, ce qui conclut.

– Montrons ensuite que si $X = (X_1, \dots, X_k)$ et $Y = (X_{k+1}, \dots, X_n)$, on a $X \perp\!\!\!\perp Y$. Si pour tout $i \in \llbracket 1, n \rrbracket$, $x_i \in X_i(\Omega)$, on a par indépendance de $X_1, \dots, X_n$:

$$\begin{aligned} \mathbb{P}(X = (x_1, \dots, x_k), Y = (y_{k+1}, \dots, y_n)) &= \mathbb{P}\left(\bigcap_{i=1}^n \{X_i = x_i\}\right) = \prod_{i=1}^k \mathbb{P}(X_i = x_i) \prod_{i=k+1}^n \mathbb{P}(X_i = x_i) \\ &= \mathbb{P}(X = (x_1, \dots, x_k)) \mathbb{P}(Y = (y_{k+1}, \dots, y_n)) \\ \mathbb{P}(X = (x_1, \dots, x_k)) \mathbb{P}(Y = (y_{k+1}, \dots, y_n)) &= \prod_{i=1}^k \mathbb{P}(X_i = x_i) \prod_{i=k+1}^n \mathbb{P}(X_i = x_i) = \prod_{i=1}^n \mathbb{P}(X_i = x_i). \end{aligned}$$

Ainsi, $X \perp\!\!\!\perp Y$. En joignant les deux résultats, on a bien le résultat de l'énoncé. $\square$

Exemple. Si X, Y, Z, T, U sont des variables (mutuellement) indépendantes sur un même espace probabilité, alors $\sqrt{X^2 + Y^2}$ et $e^{Z(T+U)}$ sont indépendantes.

3. Covariance

Définition-théorème - Covariance

Soient X et Y deux variables aléatoires réelles. On définit la *covariance* de X et Y , notée $\text{Cov}(X, Y)$ par

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

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

Si $\text{Cov}(X, Y) = 0$, on dit que les variables aléatoires X et Y sont *décorrélées*.

Démonstration. On a $\text{Cov}(X, Y) = \mathbb{E}(XY - \mathbb{E}(X)Y - \mathbb{E}(Y)X + \mathbb{E}(X)\mathbb{E}(Y))$. Comme $\mathbb{E}(X) \in \mathbb{R}$, $\mathbb{E}(Y) \in \mathbb{R}$ et $\mathbb{E}(X)\mathbb{E}(Y) \in \mathbb{R}$, la linéarité de l'espérance donne

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

Remarques.

- Si X est une variable aléatoire réelle, alors $\text{Cov}(X, X) = \mathbb{V}(X)$.
- Si X et Y sont indépendantes, on sait que $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, donc $\text{Cov}(X, Y) = 0$: X et Y sont *décorrélées*. Notons que, comme on l'a vu plus haut, X et Y peuvent vérifier $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ sans être indépendantes, donc X et Y peuvent être *décorrélées* sans être indépendantes.

Théorème - Variance d'une somme de variables aléatoires

Soient X et Y deux variables aléatoires réelles. On a

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + 2\text{Cov}(X, Y) + \mathbb{V}(Y).$$

Ainsi, si $\text{Cov}(X, Y) = 0$, alors $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$. C'est en particulier le cas si $X \perp Y$.

Démonstration. On a en utilisant la linéarité de l'espérance :

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

d'où $\mathbb{V}(X + Y) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)) + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = \mathbb{V}(X) + 2\text{Cov}(X, Y) + \mathbb{V}(Y)$. $\square$

Remarque. Si $X \sim \mathcal{B}(n, p)$, alors on a vu que $X \sim X_1 + \dots + X_n$, où $X_1, \dots, X_n$ sont des variables aléatoires indépendantes suivant la loi de Bernoulli $\mathcal{B}(p)$. Par conséquent, on retrouve :

$$\mathbb{V}(X) = \mathbb{V}(X_1) + \dots + \mathbb{V}(X_n) = np.$$

Théorème - Inégalité de Cauchy-Schwarz

Si X et Y sont deux variables aléatoires réelles, alors

$$|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y).$$

Démonstration. Montrons l'inégalité $|\mathbb{E}(XY)| \leq \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}$, encore appelée inégalité de Cauchy-Schwarz : on note

$$f : t \mapsto \mathbb{E}((tX + Y)^2) = \mathbb{E}(t^2X^2 + 2tXY + Y^2) = t^2\mathbb{E}(X^2) + 2t\mathbb{E}(XY) + \mathbb{E}(Y^2).$$

On remarque que la fonction f est polynomiale de degré au plus 2, et positive sur $\mathbb{R}$ entier, par conséquent son discriminant associé $\Delta = 4\mathbb{E}(XY)^2 - 4\mathbb{E}(X^2)\mathbb{E}(Y^2)$ est négatif, ce qui donne $\mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$ ce qui conclut.

Ainsi, $|\text{Cov}(X, Y)| = |\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))| \leq \sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)}\sqrt{\mathbb{E}((Y - \mathbb{E}(Y))^2)} = \sigma(X)\sigma(Y)$. $\square$

Remarque. On en déduit que, si X et Y sont deux variables réelles de variance non nulle, alors

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)} \in [-1, 1].$$

Ce réel $\rho(X, Y)$ est appelé coefficient de corrélation linéaire de X et Y . On remarque d'après le cas d'égalité dans l'inégalité de Cauchy-Schwarz que $\rho(X, Y)$ vaut 1 si et seulement si X et Y sont positivement liées, et $\rho(X, Y) = -1$ si et seulement si X et Y sont négativement liées.

IV Inégalités probabilistes

1. Inégalité de Markov

Théorème - Inégalité de Markov

Soit X une variable aléatoire réelle à valeurs dans $\mathbb{R}_+$. Pour tout $a > 0$, on a

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

Démonstration. On remarque que $X \geq a \mathbb{1}_{\{X \geq a\}}$. En effet, si $\omega \in \{X \geq a\}$, alors $X(\omega) \geq a = a \mathbb{1}_{\{X \geq a\}}(\omega)$, et si $\omega \notin \{X \geq a\}$, alors $X(\omega) \geq 0 = a \mathbb{1}_{\{X \geq a\}}(\omega)$. Ainsi, par croissance de l'espérance, on a

$$\mathbb{E}(X) \geq \mathbb{E}(a \mathbb{1}_{\{X \geq a\}}) = a \mathbb{E}(\mathbb{1}_{\{X \geq a\}}) = a \mathbb{P}(X \geq a). \quad \square$$

Remarques.

- Notons que la majoration de l'inégalité de Markov se base seulement sur la valeur de l'espérance de X . Il s'agit donc d'une majoration souvent grossière de $\mathbb{P}(X \geq a)$.
- On remarque qu'on peut obtenir d'autres inégalités du même type en remarquant qu'en appliquant l'inégalité de Markov à e^X .

$$\mathbb{P}(X \geq a) = \mathbb{P}(e^X \geq e^a) \leq \frac{\mathbb{E}(e^X)}{e^a}.$$

2. Inégalité de Bienaymé-Tchebychev

Théorème - Inégalité de Bienaymé-Tchebychev

Si X est une variable aléatoire réelle, alors pour tout $\varepsilon > 0$,

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

Démonstration. On commence par remarquer que si $\varepsilon > 0$, alors $\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) = \mathbb{P}((X - \mathbb{E}(X))^2 \geq \varepsilon^2)$. Il suffit donc d'appliquer l'inégalité de Markov à la variable aléatoire $(X - \mathbb{E}(X))^2$:

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

La loi des grands nombres qui suit (dans sa version faible), énonce un résultat familier : si on simule un grand nombre de fois la même variable aléatoire réelle X , de manière indépendante, et qu'on prend la moyenne des valeurs obtenues, cette moyenne s'approchera de l'espérance de X .

Voici deux exemples simples.

- Si on lance 10 000 fois une pièce de monnaie équilibrée, et qu'on gagne un euro à chaque *pile* obtenu, il y a toutes les chances qu'on ait gagné une somme proche de 5 000 €. Ceci correspond à simuler 10 000 fois une variable aléatoire $X \hookrightarrow \mathcal{B}(\frac{1}{2})$. Si on fait la moyenne des gains obtenus par lancer, on obtiendra alors une valeur proche de 0,5, qui est l'espérance de X .
- Si on joue un grand nombre n de fois à une machine à sous dans un casino, et qu'on fait la moyenne des gains sur n parties, on approche assez bien l'espérance de gain de la machine. En effet, ceci revient à faire n simulations d'une variable aléatoire X qui correspond au gain pour une partie, et à effectuer la moyenne de ces simulations.

Ceci n'est valide bien sûr que si le nombre de simulations est assez grand. Si on reprend le premier exemple ci-dessus avec 3 lancers, il est tout à fait possible que les trois lancers aient donné *face*, ce qui donne un gain moyen de 0...

Théorème - Loi faible des grands nombres

Soit $(X_i)_{i \in \mathbb{N}^*}$ une suite de variables aléatoires *indépendantes* qui suivent la même loi. On note m l'espérance commune de ces variables aléatoires, et σ leur écart-type commun.

Si on note $S_n = \sum_{i=1}^n X_i$ pour tout $n \in \mathbb{N}^*$, alors pour tout $\varepsilon > 0$,

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

Démonstration. Pour tout $n \in \mathbb{N}^*$, on peut appliquer l'inégalité de Bienaymé-Tchebychev à la variable aléatoire $\frac{S_n}{n}$:

$$\mathbb{E}(S_n) = \sum_{i=1}^n \mathbb{E}(X_i) = nm, \quad \text{et} \quad \mathbb{V}(S_n) \stackrel{\text{indép.}}{=} \sum_{i=1}^n \mathbb{V}(X_i) = n\sigma^2,$$

donc $\mathbb{E}\left(\frac{S_n}{n}\right) = m$, et $\mathbb{V}\left(\frac{S_n}{n}\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$. Alors, par inégalité de Bienaymé-Tchebychev,

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

Ainsi, on a $\mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0$, ce qui conclut. $\square$

Remarques.

- Si X est une variable aléatoire réelle et si on dispose de $X_1, \dots, X_n$ variables aléatoires indépendantes et de même loi que X , alors la loi faible des grands nombres exprime qu'en un certain sens, $\frac{S_n}{n}$ s'approche de $\mathbb{E}(X)$ lorsque n est grand. Dans la pratique, une réalisation de $\frac{S_n}{n}$ (par exemple obtenue par simulation) donnera une approximation $\mathbb{E}(X)$ d'autant meilleure que n est grand.
- Cas $X \sim \mathcal{B}(p)$. Si $X \sim \mathcal{B}(p)$, alors $\mathbb{E}(X) = p$, et si $S_n = \sum_{i=1}^n X_i$, la loi faible de grands nombre entraîne

$$\mathbb{P} \left(\left| \frac{S_n}{n} - p \right| \geq \varepsilon \right) \xrightarrow{n \rightarrow +\infty} 0$$

car $\mathbb{E}(X) = p$. On a donc par ailleurs la limite ci-dessus dès que $S_n \sim \mathcal{B}(n, p)$.

- Si A est un événement, on peut donc approcher $\mathbb{P}(A)$ en utilisant le résultat ci-dessus appliqué à $X = \mathbb{1}_A$, qui suit la loi $\mathcal{B}(\mathbb{P}(A))$:

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