

Espérance, Variance & Co

Plan du chapitre

I CALCULS D'ESPÉRANCES	1
I.1 CAS GÉNÉRAL : ESPÉRANCE DE $Z = g(X, Y)$	1
I.2 ESPÉRANCE D'UNE SOMME	5
I.3 ESPÉRANCE D'UN PRODUIT	6
II COVARIANCE D'UN COUPLE	10
II.1 VARIANCE D'UNE SOMME	12
II.2 COEFFICIENT DE CORRÉLATION LINÉAIRE	15
III BILAN ET COMPÉTENCES ATTENDUES	18
INDEX	19

Dans tout ce chapitre, $(\Omega, \mathcal{A}, \mathbf{P})$ désigne un espace probabilisé.

I. CALCULS D'ESPÉRANCES

I.1 – Cas général : espérance de $Z = g(X, Y)$

Pour commencer, on retrouve le fameux :

Théorème 1 – Théorème de transfert :

Soit g une fonction définie sur $\mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)$.

La variable aléatoire $g(X, Y)$ admet une *espérance* si, et seulement si la série double

$$\sum_{(x,y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)} g(x, y) \mathbf{P}([X = x] \cap [Y = y])$$

À retenir

Sous réserve de convergence absolue, on a :

$$\mathbf{E}(g(X, Y)) = \sum_{\substack{x \in \mathcal{X}(\Omega) \\ y \in \mathcal{Y}(\Omega)}} g(x, y) \mathbf{P}([X = x] \cap [Y = y]).$$

est absolument convergente et dans ce cas, $\mathbf{E}(g(\mathbb{X}, \mathbb{Y}))$ est la somme de cette série :

$$\begin{aligned}\mathbf{E}(g(\mathbb{X}, \mathbb{Y})) &= \sum_{x \in \mathbb{X}(\Omega)} \left(\sum_{y \in \mathbb{Y}(\Omega)} g(x, y) \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) \right) \\ &= \sum_{y \in \mathbb{Y}(\Omega)} \left(\sum_{x \in \mathbb{X}(\Omega)} g(x, y) \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) \right).\end{aligned}$$

Précisons l'expression « sous réserve de convergence absolue ».

1. Si $\mathbb{X}$ et $\mathbb{Y}$ sont finies :

Dans ce cas, $\mathbb{X}(\Omega) = \{x_1, \dots, x_m\}$ et $\mathbb{Y}(\Omega) = \{y_1, \dots, y_n\}$.

La variable aléatoire $g(\mathbb{X}, \mathbb{Y})$ admet alors une espérance donnée par :

$$\mathbf{E}(g(\mathbb{X}, \mathbb{Y})) = \sum_{i=1}^m \sum_{j=1}^n g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j]).$$

2. Si $\mathbb{X}$ est finie et $\mathbb{Y}$ infinie :

Dans ce cas, $\mathbb{X}(\Omega) = \{x_1, \dots, x_m\}$ et $\mathbb{Y}(\Omega) = \{y_j \mid j \in \mathbb{N}\}$.

Si pour tout $i \in \llbracket 1, m \rrbracket$, la série $\sum_{j \geq 0} g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j])$ est absolument convergente alors $g(\mathbb{X}, \mathbb{Y})$ admet une espérance et :

$$\mathbf{E}(g(\mathbb{X}, \mathbb{Y})) = \sum_{i=1}^m \sum_{j=0}^{+\infty} g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j]).$$

Le cas $\mathbb{X}$ infinie et $\mathbb{Y}$ finie est similaire :

$$\mathbb{X}(\Omega) = \{x_i \mid i \in \mathbb{N}\} \text{ et } \mathbb{Y}(\Omega) = \{y_1, \dots, y_n\}.$$

Si pour tout $j \in \llbracket 1, n \rrbracket$, la série $\sum_{i \geq 0} g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j])$ est absolument convergente alors $g(\mathbb{X}, \mathbb{Y})$ admet une espérance et :

$$\mathbf{E}(g(\mathbb{X}, \mathbb{Y})) = \sum_{i=0}^{+\infty} \sum_{j=1}^n g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j]).$$

3. Si $\mathbb{X}$ est infinie et $\mathbb{Y}$ infinie :

Dans ce cas, $\mathbb{X}(\Omega) = \{x_i \mid i \in \mathbb{N}\}$ et $\mathbb{Y}(\Omega) = \{y_j \mid j \in \mathbb{N}\}$.

Il faut alors démontrer :

- $\forall i \in \mathbb{N}$, la série $\sum_{j \geq 0} g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j])$ est absolument convergente.
- la série $\sum_{i \geq 0} \left(\sum_{j=0}^{+\infty} g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j]) \right)$ est absolument convergente.

Alors $g(\mathbb{X}, \mathbb{Y})$ admet une espérance donnée par :

$$\mathbf{E}(g(\mathbb{X}, \mathbb{Y})) = \sum_{i=0}^{+\infty} \sum_{j=0}^{+\infty} g(x_i, y_j) \mathbf{P}([\mathbb{X} = x_i] \cap [\mathbb{Y} = y_j]).$$

Exemple 1 – Espérance d'un maximum :

On reprend les notations d'un exemple du chapitre précédent où :

- $\mathbb{X} \hookrightarrow \mathcal{G}(p_1)$ et $\mathbb{Y} \hookrightarrow \mathcal{G}(p_2)$ avec $p_1, p_2 \in]0; 1[$ et $q_1 = 1 - p_1, q_2 = 1 - p_2$,
- $\mathbb{V} = \max(\mathbb{X}, \mathbb{Y})$,
- avec $\mathbf{P}([\mathbb{V} \leq n]) = 1 - q_1^n - q_2^n + (q_1 q_2)^n$.

Montrons que $\mathbb{V}$ admet une espérance et déterminons-la.

$\mathbb{V}$ admet une espérance si, et seulement si la série $\sum_{n \geq 1} n\mathbf{P}([\mathbb{V} = n])$ est absolument convergente i.e. convergente puisque cette série est à termes positifs.

Soit $m \in \mathbb{N}$.

Montrons tout d'abord que $\sum_{n=1}^m n\mathbf{P}([\mathbb{V} = n]) = \sum_{n=0}^{m-1} \mathbf{P}([\mathbb{V} > n]) - m\mathbf{P}([\mathbb{V} > m])$.

En effet,

$$\begin{aligned} \sum_{n=1}^m n\mathbf{P}([\mathbb{V} = n]) &= \sum_{n=1}^m n(\mathbf{P}([\mathbb{V} > n-1]) - \mathbf{P}([\mathbb{V} > n])) \\ &= \sum_{n=1}^m n\mathbf{P}([\mathbb{V} > n-1]) - \sum_{n=1}^m n\mathbf{P}([\mathbb{V} > n]) \quad (\text{par linéarité de } \Sigma) \\ &= \sum_{n=0}^{m-1} (n+1)\mathbf{P}([\mathbb{V} > n]) - \sum_{n=1}^m n\mathbf{P}([\mathbb{V} > n]) \quad (\text{par décalage d'indice}) \\ &= \sum_{n=0}^{m-1} n\mathbf{P}([\mathbb{V} > n]) + \sum_{n=0}^{m-1} \mathbf{P}([\mathbb{V} > n]) - \sum_{n=1}^m n\mathbf{P}([\mathbb{V} > n]) \\ &= \sum_{n=0}^{m-1} \mathbf{P}([\mathbb{V} > n]) - m\mathbf{P}([\mathbb{V} > m]). \end{aligned}$$

Comme $\mathbf{P}([\mathbb{V} \leq n]) = (1 - q_1^n)(1 - q_2^n) = 1 - q_1^n - q_2^n + (q_1 q_2)^n$, on en déduit,

$$\mathbf{P}([\mathbb{V} > n]) = 1 - \mathbf{P}([\mathbb{V} \leq n]) = q_1^n + q_2^n - (q_1 q_2)^n.$$

$$\text{D'où, } \sum_{n=1}^m n\mathbf{P}([\mathbb{V} = n]) = \sum_{n=0}^{m-1} q_1^n + \sum_{n=0}^{m-1} q_2^n - \sum_{n=0}^{m-1} (q_1 q_2)^n - m\mathbf{P}([\mathbb{V} > m]).$$

Or,

- $q_1, q_2 \in]0; 1[$ entraîne la convergence des séries $\sum_{n \in \mathbb{N}} q_1^n$ et $\sum_{n \in \mathbb{N}} q_2^n$.

$$\text{De plus, } \sum_{n=0}^{m-1} q_1^n \xrightarrow{m \rightarrow +\infty} \frac{1}{1 - q_1} = \frac{1}{p_1} \quad \text{et} \quad \sum_{n=0}^{m-1} q_2^n \xrightarrow{m \rightarrow +\infty} \frac{1}{1 - q_2} = \frac{1}{p_2}.$$

- De même, $q_1 q_2 \in]0; 1[$, entraîne la convergence de la série $\sum_{n \in \mathbb{N}} (q_1 q_2)^n$ et

$$\sum_{n=0}^{m-1} (q_1 q_2)^n \xrightarrow{m \rightarrow +\infty} \frac{1}{1 - q_1 q_2} = \frac{1}{p_1 + p_2 - p_1 p_2}.$$

- Enfin, $m\mathbf{P}([\mathbb{V} > m]) = m(q_1^m + q_2^m - (q_1 q_2)^m) \xrightarrow{m \rightarrow +\infty} 0$ par croissances comparées.

En conclusion, la série $\sum_{n \in \mathbb{N}} n \mathbf{P}(\mathbb{V} = n)$ est convergente *i.e.* $\mathbb{V}$ admet une espérance et

$$\mathbf{E}(\mathbb{V}) = \frac{1}{p_1} + \frac{1}{p_2} - \frac{1}{p_1 + p_2 - p_1 p_2}.$$

Exemple 2 :

Considérons deux variables aléatoires $\mathbb{X}$ et $\mathbb{Y}$ dont la loi jointe est définie par :

$$\mathbb{X}(\Omega) = \mathbb{Y}(\Omega) = \mathbb{N} \quad \text{et} \quad \forall (j, k) \in \mathbb{N}^2, \quad \mathbf{P}([\mathbb{X} = j] \cap [\mathbb{Y} = k]) = \frac{j+k}{e 2^{j+k} j! k!}.$$

Justifions que la variable aléatoire $2^{\mathbb{X}+\mathbb{Y}}$ admet une espérance et déterminons-la.

— D'après le théorème de transfert :

$$\begin{aligned} 2^{\mathbb{X}+\mathbb{Y}} \text{ admet une espérance} &\iff \text{la série double} \\ &\sum_{(j,k) \in \mathbb{N}^2} 2^{j+k} \mathbf{P}([\mathbb{X} = j] \cap [\mathbb{Y} = k]) \text{ est absolu-} \\ &\text{ment convergente} \\ &\iff \text{la série double} \\ &\sum_{(j,k) \in \mathbb{N}^2} 2^{j+k} \mathbf{P}([\mathbb{X} = j] \cap [\mathbb{Y} = k]) \text{ est conver-} \\ &\text{gente, car il s'agit d'une série à termes posi-} \\ &\text{tifs.} \end{aligned}$$

— Soit $j \in \mathbb{N}$. Soit $N \in \mathbb{N}$, on a :

$$\begin{aligned} \sum_{k=0}^N 2^{j+k} \mathbf{P}([\mathbb{X} = j] \cap [\mathbb{Y} = k]) &= \frac{1}{e j!} \sum_{k=0}^N \frac{j+k}{k!} = \frac{1}{e j!} \left(j \sum_{k=0}^N \frac{1}{k!} + \sum_{k=0}^N \frac{k}{k!} \right) \\ &= \frac{1}{e j!} \left(j \sum_{k=0}^N \frac{1}{k!} + \sum_{k=1}^N \frac{1}{(k-1)!} \right) \\ &= \frac{1}{e j!} \left(j \sum_{k=0}^N \frac{1}{k!} + \sum_{k=0}^{N-1} \frac{1}{k!} \right) \end{aligned}$$

Or, $\sum_{k \geq 0} \frac{1}{k!}$ est une série exponentielle convergente de somme e .

Par conséquent, la série $\sum_{k \geq 0} 2^{j+k} \mathbf{P}([\mathbb{X} = j] \cap [\mathbb{Y} = k])$ est convergente et :

$$\begin{aligned} \sum_{k=0}^{+\infty} 2^{j+k} \mathbf{P}([\mathbb{X} = j] \cap [\mathbb{Y} = k]) &= \frac{1}{e j!} \left(j \sum_{k=0}^{+\infty} \frac{1}{k!} + \sum_{k=0}^{+\infty} \frac{1}{k!} \right) \\ &= \frac{1}{e j!} (j e + e) \\ &= \frac{j+1}{j!}. \end{aligned}$$

— Ensuite, de la même façon que ci-dessus, la série $\sum_{j \geq 0} \frac{j+1}{j!}$ est convergente, de somme égale à $2e$.

Idée
 $\mathbb{X}$ et $\mathbb{Y}$ ont des rôles symétriques dans la loi jointe donc les deux lois marginales sont identiques.

— Par conséquent, la série double $\sum_{(j,k) \in \mathbb{N}^2} 2^{j+k} \mathbf{P}([X = j] \cap [Y = k])$ est convergente.

On en déduit que 2^{X+Y} admet une espérance et :

$$\mathbf{E}(2^{X+Y}) = \sum_{j=0}^{+\infty} \sum_{k=0}^{+\infty} 2^{j+k} \mathbf{P}([X = j] \cap [Y = k]) = \sum_{j=0}^{+\infty} \left(\frac{j}{j!} + \frac{1}{j!} \right) = 2e.$$

I.2 – Espérance d'une somme

Proposition 2 – Linéarité de l'espérance :

Soit $n \in \mathbb{N}_{\geq 2}$. Si les variables aléatoires $X_1, X_2, \dots, X_n$ admettent toutes une espérance, alors, pour tous $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$, la variable aléatoire $\lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_n X_n$ admet une espérance et :

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

En particulier, lorsque pour tout $k \in \llbracket 1, n \rrbracket$, $\lambda_k = 1$, on obtient :

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

Preuve : Commençons par le montrer pour deux variables aléatoires X et Y et posons, pour cela, $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ la fonction définie par $g(x, y) = \lambda x + \mu y$ où $\lambda, \mu \in \mathbb{R}$.

D'après le **théorème (1)** de transfert, $Z = g(X, Y)$ admet une espérance si, et seulement si, la famille des

$$g(x, y) \mathbf{P}([X = x] \cap [Y = y]) = \lambda x \mathbf{P}([X = x] \cap [Y = y]) + \mu y \mathbf{P}([X = x] \cap [Y = y])$$

est absolument sommable.

Considérons la famille $\left(x \mathbf{P}([X = x] \cap [Y = y]) \right)_{(x,y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)}$:

— Pour tout $x \in \mathcal{X}(\Omega)$, la série $\sum_{y \in \mathcal{Y}(\Omega)} |x| \mathbf{P}([X = x] \cap [Y = y])$ converge d'après la formule des probabilités totales, et on a :

$$\sum_{y \in \mathcal{Y}(\Omega)} |x| \mathbf{P}([X = x] \cap [Y = y]) = |x| \mathbf{P}([X = x]).$$

— La série $\sum_{x \in \mathcal{X}(\Omega)} |x| \mathbf{P}([X = x])$ est convergente car X admet une espérance.

La famille $\left(x \mathbf{P}([X = x] \cap [Y = y]) \right)_{(x,y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)}$ est donc absolument convergente (sommable) et :

$$\sum_{(x,y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)} x \mathbf{P}([X = x] \cap [Y = y]) = \mathbf{E}(X).$$

Le même raisonnement s'applique à la famille $\left(y \mathbf{P}([X = x] \cap [Y = y]) \right)_{(x,y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)}$.

i Remarque

On dit indifféremment absolument convergente ou sommable.

On peut donc conclure que la famille $\left(g(x, y) \mathbf{P}([X = x] \cap [Y = y]) \right)_{(x, y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)}$ est sommable comme combinaison linéaire de familles sommables, donc $\mathbf{E}(Z)$ existe et vaut :

$$\begin{aligned} \mathbf{E}(Z) &= \sum_{(x, y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)} g(x, y) \mathbf{P}([X = x] \cap [Y = y]) \\ &= \lambda \sum_{(x, y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)} x \mathbf{P}([X = x] \cap [Y = y]) + \mu \sum_{(x, y) \in \mathcal{X}(\Omega) \times \mathcal{Y}(\Omega)} y \mathbf{P}([X = x] \cap [Y = y]) \\ &= \lambda \mathbf{E}(X) + \mu \mathbf{E}(Y). \end{aligned}$$

Le résultat général s'obtient ensuite par récurrence sur $n \in \mathbb{N}_{\geq 2}$. 

Proposition 3 – Croissance de l'espérance :

- P1** • $\left. \begin{array}{l} \forall \omega \in \Omega, X(\omega) \geq 0 \text{ ou } \mathbf{P}([X \geq 0]) = 1 \\ X \text{ admet une espérance} \end{array} \right\} \Rightarrow \mathbf{E}(X) \geq 0.$
- P2** • $\left. \begin{array}{l} \forall \omega \in \Omega, X(\omega) \leq Y(\omega) \text{ ou } \mathbf{P}([X \leq Y]) = 1 \\ X \text{ et } Y \text{ admettent une espérance} \end{array} \right\} \Rightarrow \mathbf{E}(X) \leq \mathbf{E}(Y).$

 **Preuve :**

- P1** • Supposons que $\mathbf{P}([X \geq 0]) = 1$ et que X admet une espérance.
Puisque X admet une espérance, la série $\sum_{x \in \mathcal{X}(\Omega)} x \mathbf{P}([X = x])$ est absolument convergente et :

$$\mathbf{E}(X) = \sum_{x \in \mathcal{X}(\Omega)} x \mathbf{P}([X = x]) = \sum_{\substack{x \in \mathcal{X}(\Omega) \\ x \geq 0}} x \mathbf{P}([X = x]) \geq 0$$

car $\mathbf{P}([X \geq 0]) = 1$ entraîne $\mathbf{P}([X = x]) = 0$ pour tout $x < 0$.

- P2** • On applique le point précédent à la variable aléatoire $Z = Y - X$, par linéarité de l'espérance. 

I.3 – Espérance d'un produit

Théorème 4 – Cas général :

Soit (X, Y) un couple de variables aléatoires réelles discrètes admettant chacune un moment d'ordre 2.

Alors :

— XY admet une espérance.

— Et :

$$\begin{aligned}
\mathbf{E}(\mathbb{X}\mathbb{Y}) &= \sum_{(x,y) \in \mathbb{X}(\Omega) \times \mathbb{Y}(\Omega)} xy \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) \\
&= \sum_{x \in \mathbb{X}(\Omega)} \sum_{y \in \mathbb{Y}(\Omega)} xy \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) \\
&= \sum_{y \in \mathbb{Y}(\Omega)} \sum_{x \in \mathbb{X}(\Omega)} xy \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]).
\end{aligned}$$

 **Preuve :****Lemme 5 :**

$$\forall a, b \in \mathbb{R}, \quad |ab| \leq \frac{a^2 + b^2}{2}.$$

 **Preuve :** Pour tous $a, b \in \mathbb{R}$, $0 \leq (a + b)^2 \Rightarrow -ab \leq \frac{a^2 + b^2}{2}$ et $0 \leq (a - b)^2 \Rightarrow ab \leq \frac{a^2 + b^2}{2}$.
D'où $|ab| \leq \frac{a^2 + b^2}{2}$. 

D'après le théorème de transfert, $\mathbb{X}\mathbb{Y}$ admet une espérance ssi la série double $\sum_{(x,y) \in \mathbb{X}(\Omega) \times \mathbb{Y}(\Omega)} xy \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y])$ est absolument convergente.

Par le lemme : $\forall (x, y) \in \mathbb{X}(\Omega) \times \mathbb{Y}(\Omega)$,

$$0 \leq |xy| \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) \leq \frac{1}{2} (x^2 \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) + y^2 \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]))$$

Soit $x \in \mathbb{X}(\Omega)$. Par la formule des probabilités totales, $\sum_{y \in \mathbb{Y}(\Omega)} \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) = \mathbf{P}([\mathbb{X} = x])$, d'où
 $x^2 \mathbf{P}([\mathbb{X} = x]) = \sum_{y \in \mathbb{Y}(\Omega)} x^2 \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]).$

Comme $\mathbb{X}$ et $\mathbb{Y}$ admettent un moment d'ordre 2, les séries $\sum_{x \in \mathbb{X}(\Omega)} x^2 \mathbf{P}([\mathbb{X} = x])$ et $\sum_{y \in \mathbb{Y}(\Omega)} y^2 \mathbf{P}([\mathbb{Y} = y])$ convergent.

Par comparaison de séries à termes positifs, la série double $\sum_{(x,y) \in \mathbb{X}(\Omega) \times \mathbb{Y}(\Omega)} |xy| \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y])$ converge, ce qui prouve l'existence de $\mathbf{E}(\mathbb{X}\mathbb{Y})$. 

Exemple 3 :

On considère le jet de deux dés à 3 faces et on note :

- $\mathbb{X}$ le minimum des deux nombres obtenus,
- $\mathbb{Y}$ le maximum des deux nombres obtenus.

Déterminons $\mathbf{E}(\mathbb{X}\mathbb{Y})$.

La loi conjointe de $(\mathbb{X}, \mathbb{Y})$ est résumée ci-dessous ($\mathbb{X}(\Omega) = \mathbb{Y}(\Omega) = \llbracket 1, 3 \rrbracket$) :

$\mathbb{X} \backslash \mathbb{Y}$	1	2	3
1	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$
2	0	$\frac{1}{9}$	$\frac{2}{9}$
3	0	0	$\frac{1}{9}$

$\mathbb{X}$ et $\mathbb{Y}$ étant à support fini, elles admettent un moment d'ordre 2. Par le théorème de transfert :

$$\begin{aligned}
 \mathbf{E}(\mathbb{X}\mathbb{Y}) &= \sum_{x \in \mathbb{X}(\Omega) \times \mathbb{Y}(\Omega)} xy \mathbf{P}([\mathbb{X} = x] \cap [\mathbb{Y} = y]) \\
 &= \frac{1}{9}(1 \cdot 1) + \frac{2}{9}(1 \cdot 2) + \frac{2}{9}(1 \cdot 3) + \frac{1}{9}(2 \cdot 2) + \frac{2}{9}(2 \cdot 3) + \frac{1}{9}(3 \cdot 3) \\
 &= \frac{1 + 4 + 6 + 4 + 12 + 9}{9} = \frac{36}{9} = 4.
 \end{aligned}$$

Exercice 1 – (Calcul d'espérance par transfert)

Soit $(\mathbb{X}, \mathbb{Y})$ un couple de variables aléatoires dont la loi conjointe est donnée par :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, \quad \mathbf{P}([\mathbb{X} = i] \cap [\mathbb{Y} = j]) = \frac{4ij}{n^2(n+1)^2}$$

1. Vérifier que la somme des probabilités vaut 1.
2. En utilisant le théorème de transfert, calculer l'espérance $\mathbf{E}(\mathbb{Z})$ de la variable $\mathbb{Z} = \mathbb{X}\mathbb{Y}$.

1. On calcule la somme double des probabilités :

$$\sum_{i=1}^n \sum_{j=1}^n \frac{4ij}{n^2(n+1)^2} = \frac{4}{n^2(n+1)^2} \left(\sum_{i=1}^n i \right) \left(\sum_{j=1}^n j \right)$$

Comme $\sum_{i=1}^n i = \frac{n(n+1)}{2}$, on a :

$$\sum_{i=1}^n \sum_{j=1}^n \mathbf{P}([\mathbb{X} = i] \cap [\mathbb{Y} = j]) = \frac{4}{n^2(n+1)^2} \times \frac{n^2(n+1)^2}{4} = 1$$

2. Par le théorème de transfert :

$$\mathbf{E}(\mathbb{X}\mathbb{Y}) = \sum_{i=1}^n \sum_{j=1}^n (ij) \cdot \mathbf{P}([\mathbb{X} = i] \cap [\mathbb{Y} = j]) = \frac{4}{n^2(n+1)^2} \left(\sum_{i=1}^n i^2 \right) \left(\sum_{j=1}^n j^2 \right)$$

Comme $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$:

$$\mathbf{E}(\mathbb{Z}) = \frac{4}{n^2(n+1)^2} \left(\frac{n(n+1)(2n+1)}{6} \right)^2 = \frac{(2n+1)^2}{9}$$

Proposition 6 – Espérance d'un produit de variables aléatoires indépendantes :

- P1** • Si X et Y sont deux variables aléatoires **indépendantes** admettant une espérance, alors la variable aléatoire XY admet une espérance et

$$E(XY) = E(X)E(Y).$$

- P2** • Si $(X_k)_{k \in \mathbb{N}^*}$ est une suite de variables aléatoires **indépendantes** admettant une espérance, alors pour tout $n \in \mathbb{N}^*$, la variable aléatoire $\prod_{k=1}^n X_k$ admet une espérance et

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

Attention

La réciproque est fausse !
confer le contre-
exemple (6)

Rappel

Si X admet une espérance, alors pour tous $a, b \in \mathbb{R}$, $aX + b$ admet une espérance et $E(aX + b) = aE(X) + b$.

Preuve : Le **théorème (1)** de transfert s'applique directement avec

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

ce qui permet de séparer les sommes et de prouver **P1**.

Le point **P2** s'en déduit par récurrence sur $n \in \mathbb{N}^*$. ✓

Exemple 4 :

Soit $p \in]0; 1[$. On considère une variable aléatoire X dont la loi est donnée par :

$$X(\Omega) = \{-1; 1\} \quad \text{et} \quad P([X = 1]) = p, \quad P([X = -1]) = 1 - p.$$

Pour $n \in \mathbb{N}^*$, on considère $X_1, \dots, X_n$ des variables aléatoires indépendantes ayant toutes la même loi que X et on pose $Y_n = \prod_{k=1}^n X_k$.

Déterminons l'espérance de Y_n puis déduisons-en la loi de Y_n .

— Puisque $X(\Omega)$ est fini, la variable aléatoire X admet une espérance et :

$$E(X) = -1 \cdot (1 - p) + 1 \cdot p = 2p - 1.$$

— Les variables $X_1, \dots, X_n$ admettent une espérance et sont indépendantes, donc Y_n admet une espérance et :

$$E(Y_n) = \prod_{k=1}^n E(X_k) = (2p - 1)^n.$$

— Pour tout $k \in \mathbb{N}^*$, $X_k(\Omega) = \{-1; 1\}$ d'où $Y_n(\Omega) = \{-1; 1\}$ et :

$$P([Y_n = -1]) + P([Y_n = 1]) = 1. \quad (\text{V.1})$$

Or $E(Y_n) = -P([Y_n = -1]) + P([Y_n = 1])$, d'où :

$$-P([Y_n = -1]) + P([Y_n = 1]) = (2p - 1)^n. \quad (\text{V.2})$$

En résolvant le système formé par (V.1) et (V.2), on obtient :

$$P([Y_n = 1]) = \frac{1 + (2p - 1)^n}{2} \quad \text{et} \quad P([Y_n = -1]) = \frac{1 - (2p - 1)^n}{2}.$$

Remarque

Comme $Y_n(\Omega) = \{-1; 1\}$ est fini, l'existence de l'espérance était immédiate.

Exemple 5 :

Soient X_1, X_2, X_3 trois variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre $p \in]0; 1[$.

On pose $Y = X_1 X_2$ et $Z = X_2 X_3$.

Montrons que Y et Z ne sont pas indépendantes.

— Par indépendance de X_1 et X_2 : $E(Y) = E(X_1)E(X_2) = p^2$. De même, $E(Z) = p^2$.

— On a $YZ = X_1 X_2^2 X_3$. Comme $X_2 \hookrightarrow \mathcal{B}(p)$, on a $X_2^2 = X_2$.

D'où $YZ = X_1 X_2 X_3$. Par indépendance de X_1, X_2, X_3 :

$$E(YZ) = E(X_1)E(X_2)E(X_3) = p^3.$$

Comme $p \in]0; 1[$, $p^3 \neq p^4 = E(Y)E(Z)$, Y et Z ne sont pas indépendantes.

À retenir
 $X \hookrightarrow \mathcal{B}(p) \implies X^2 = X$.

Contre-exemple 6 – $E(XY) = E(X)E(Y)$ mais X et Y non indépendantes :

On choisit au hasard et de façon équiprobable un point dans le plan parmi les points de coordonnées $(0; 1)$, $(1; 0)$, $(-1; 0)$ et $(0; -1)$.

On note X l'abscisse du point et Y son ordonnée.

Par équiprobabilité, $P([X = -1]) = \frac{1}{4}$, $P([X = 0]) = \frac{1}{2}$, $P([X = 1]) = \frac{1}{4}$, d'où $E(X) = 0$.

De même $E(Y) = 0$.

Pour tout point possible, le produit de l'abscisse et de l'ordonnée est nul, donc $XY = 0$ presque sûrement et $E(XY) = 0$.

Cependant, $P([X = 0] \cap [Y = 0]) = 0 \neq P([X = 0])P([Y = 0]) = \frac{1}{4}$.

Les variables X et Y ne sont donc pas indépendantes, bien que $E(XY) = E(X)E(Y)$.

II. COVARIANCE D'UN COUPLE**Définition 1 – Covariance :**

La *covariance* du couple (X, Y) est le réel, noté $\text{Cov}(X, Y)$, défini par (si existence) :

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

À retenir
 $\text{Cov}(X, X) = V(X)$.

Théorème 7 – Condition suffisante d'existence de la covariance :

Si les variables aléatoires X et Y admettent un moment d'ordre 2, alors $\text{Cov}(X, Y)$ existe.

Preuve : Supposons que X et Y admettent un moment d'ordre 2. En développant : $(X - E(X))(Y - E(Y)) = XY - E(X)Y - E(Y)X + E(X)E(Y)$. Comme X et Y ont un moment d'ordre 2, XY , X et Y admettent une espérance. Par linéarité, la combinaison linéaire $(X - E(X))(Y - E(Y))$ admet une espérance, donc $\text{Cov}(X, Y)$ existe. 

En pratique, on utilisera assez peu la **définition (1)**, mais plutôt la formule qui suit :

Proposition 8 – Formule de Koenig-Huygens :

Si les variables aléatoires X et Y admettent un moment d'ordre 2, alors :

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

 **Preuve :** Par linéarité de l'espérance :

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

**Proposition 9 :**

Soient X, Y, X_1, X_2, Y_1, Y_2 des variables aléatoires admettant un moment d'ordre 2.

P1 • $\text{Cov}(X, X) = V(X).$

P2 • $\text{Cov}(X, Y) = \text{Cov}(Y, X).$ (symétrie)

P3 • $\forall \lambda_1, \lambda_2 \in \mathbb{R}, \text{Cov}(\lambda_1 X_1 + \lambda_2 X_2, Y) = \lambda_1 \text{Cov}(X_1, Y) + \lambda_2 \text{Cov}(X_2, Y)$ (linéarité à gauche)

P4 • $\forall \lambda_1, \lambda_2 \in \mathbb{R}, \text{Cov}(X, \lambda_1 Y_1 + \lambda_2 Y_2) = \lambda_1 \text{Cov}(X, Y_1) + \lambda_2 \text{Cov}(X, Y_2)$ (linéarité à droite)

P5 • $\forall a \in \mathbb{R}, \text{Cov}(X, a) = 0.$

P6 • Si X et Y sont indépendantes, alors $\text{Cov}(X, Y) = 0.$

⚠ Attention

La réciproque de **P6** est fautive mais la contraposée peut servir pour montrer que deux variables aléatoires ne sont pas indépendantes !

 **Preuve :**

P1 • Immédiat par définition de $\text{Cov}(X, X)$ et $V(X).$

P2 • Immédiat par définition de $\text{Cov}(X, Y)$ et $\text{Cov}(Y, X).$

P3 • Soient $\lambda_1, \lambda_2 \in \mathbb{R}.$

On sait que X_1 et X_2 admettent un moment d'ordre 2, donc $\lambda_1 X_1 + \lambda_2 X_2$ également.

Et on sait que Y admet un moment d'ordre 2.

Par conséquent, $\text{Cov}(\lambda_1 X_1 + \lambda_2 X_2, Y)$ existe et d'après la formule de Koenig-Huygens :

$$\begin{aligned} \text{Cov}(\lambda_1 X_1 + \lambda_2 X_2, Y) &= E((\lambda_1 X_1 + \lambda_2 X_2)Y) - E(\lambda_1 X_1 + \lambda_2 X_2)E(Y) \\ &= E(\lambda_1 X_1 Y + \lambda_2 X_2 Y) - (\lambda_1 E(X_1) + \lambda_2 E(X_2))E(Y) \\ &= \lambda_1 E(X_1 Y) + \lambda_2 E(X_2 Y) - \lambda_1 E(X_1)E(Y) - \lambda_2 E(X_2)E(Y) \\ &= \lambda_1 \text{Cov}(X_1, Y) + \lambda_2 \text{Cov}(X_2, Y) \end{aligned}$$

P4 • Immédiat, par symétrie et linéarité à gauche de la covariance.

P5 • Soit $a \in \mathbb{R}.$ On a :

$$\text{Cov}(X, a) = E((X - E(X))(a - a)) = 0.$$

P6 • D'après la formule de Koenig-Huygens :

$$\begin{aligned} \text{Cov}(X, Y) &= E(XY) - E(X)E(Y) \\ &= E(X)E(Y) - E(X)E(Y) \quad (\text{indépendance de } X \text{ et } Y) \\ &= 0 \end{aligned}$$

🏠 Rappel

La somme de variables aléatoires admettant un moment d'ordre r admet également un moment d'ordre r . Immédiat en utilisant : $\forall a, b \in \mathbb{R},$

$$|(a+b)^r| \leq |a|^r + |b|^r.$$

(avec formule du binôme)



II.1 – Variance d'une somme

Proposition 10 – Variance d'une somme :

- P1** • Si $\mathbb{X}$ et $\mathbb{Y}$ admettent une variance, alors $\mathbb{X} + \mathbb{Y}$ admet une variance et :

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

- P2** • Si $\mathbb{X}$ et $\mathbb{Y}$ admettent une variance et sont indépendantes, alors $\mathbb{X} + \mathbb{Y}$ admet une variance et :

$$\mathbf{V}(\mathbb{X} + \mathbb{Y}) = \mathbf{V}(\mathbb{X}) + \mathbf{V}(\mathbb{Y}).$$

- P3** • Si $(\mathbb{X}_k)_{k \in \mathbb{N}^*}$ est une suite de variables aléatoires mutuellement indépendantes admettant une variance, alors pour tout $n \in \mathbb{N}^*$, $\sum_{k=1}^n \mathbb{X}_k$ admet une variance et :

$$\mathbf{V} \left(\sum_{k=1}^n \mathbb{X}_k \right) = \sum_{k=1}^n \mathbf{V}(\mathbb{X}_k).$$

Remarque : On retiendra en particulier que :

$$\mathbf{Cov}(\mathbb{X}, \mathbb{Y}) = \frac{1}{2} \left(\mathbf{V}(\mathbb{X} + \mathbb{Y}) - \mathbf{V}(\mathbb{X}) - \mathbf{V}(\mathbb{Y}) \right),$$

qui permet de calculer la covariance à partir des variances.

Preuve :

- P1** • Supposons que $\mathbb{X}$ et $\mathbb{Y}$ admettent une variance. Ainsi, $\mathbb{X}$ et $\mathbb{Y}$ admettent un moment d'ordre 2, et donc $\mathbb{X} + \mathbb{Y}$ également.

Par conséquent, $\mathbb{X} + \mathbb{Y}$ admet une variance et

$$\begin{aligned} \mathbf{V}(\mathbb{X} + \mathbb{Y}) &= \mathbf{Cov}(\mathbb{X} + \mathbb{Y}, \mathbb{X} + \mathbb{Y}) \\ &= \mathbf{Cov}(\mathbb{X}, \mathbb{X} + \mathbb{Y}) + \mathbf{Cov}(\mathbb{Y}, \mathbb{X} + \mathbb{Y}) \\ &= \mathbf{Cov}(\mathbb{X}, \mathbb{X}) + \mathbf{Cov}(\mathbb{X}, \mathbb{Y}) + \mathbf{Cov}(\mathbb{Y}, \mathbb{X}) + \mathbf{Cov}(\mathbb{Y}, \mathbb{Y}) \end{aligned}$$

Par linéarité à gauche et à droite et par symétrie.

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

- P2** • Immédiat d'après **P1** et la proposition (9).

- P3** • Supposons que $(\mathbb{X}_k)_{k \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes admettant une variance. On raisonne par récurrence sur $n \in \mathbb{N}^*$.

— Initialisation. Pour $n = 1$: rien à faire.

— Hérédité. Soit $n \in \mathbb{N}^*$. Supposons « $\sum_{k=1}^n \mathbb{X}_k$ admet une variance et $\mathbf{V} \left(\sum_{k=1}^n \mathbb{X}_k \right) = \sum_{k=1}^n \mathbf{V}(\mathbb{X}_k)$ » et montrons le résultat pour $n + 1$ variables indépendantes.

$$\text{Il suffit d'écrire } \sum_{k=1}^{n+1} \mathbb{X}_k = \sum_{k=1}^n \mathbb{X}_k + \mathbb{X}_{n+1}.$$

Alors,

Remarque

Si on initialise à $n = 2$, c'est alors **P2**.

- par hypothèse de récurrence, $\sum_{k=1}^n \mathbb{X}_k$ admet une variance,
- $\mathbb{X}_{n+1}$ admet une variance,
- les variables aléatoires $\mathbb{X}_1, \dots, \mathbb{X}_n, \mathbb{X}_{n+1}$ sont indépendantes, donc par lemme des coalitions, les variables aléatoires $\mathbb{X}_1 + \dots + \mathbb{X}_n$ et $\mathbb{X}_{n+1}$ sont également indépendantes.

D'après **P2**, la variable aléatoire $\sum_{k=1}^n \mathbb{X}_k + \mathbb{X}_{n+1}$ admet une variance et on a :

$$\begin{aligned}
 \mathbf{V} \left(\sum_{k=1}^{n+1} \mathbb{X}_k \right) &= \mathbf{V} \left(\sum_{k=1}^n \mathbb{X}_k + \mathbb{X}_{n+1} \right) \\
 &= \mathbf{V} \left(\sum_{k=1}^n \mathbb{X}_k \right) + \mathbf{V}(\mathbb{X}_{n+1}) \\
 &= \sum_{k=1}^n \mathbf{V}(\mathbb{X}_k) + \mathbf{V}(\mathbb{X}_{n+1}) \quad (\text{hypothèse de récurrence}) \\
 &= \sum_{k=1}^{n+1} \mathbf{V}(\mathbb{X}_k).
 \end{aligned}$$

L'hérédité est ainsi établie.

- Conclusion : pour tout $n \in \mathbb{N}^*$, $\sum_{k=1}^n \mathbb{X}_k$ admet une variance et

$$\mathbf{V} \left(\sum_{k=1}^n \mathbb{X}_k \right) = \sum_{k=1}^n \mathbf{V}(\mathbb{X}_k).$$



Exemple 7 :

Dans l'exemple du jet de deux dés à 3 faces, déterminons $\mathbf{V}(\mathbb{X} + \mathbb{Y})$ avec $\mathbb{X} = \min(\mathbb{Z}_1, \mathbb{Z}_2)$ et $\mathbb{Y} = \max(\mathbb{Z}_1, \mathbb{Z}_2)$ où $\mathbb{Z}_1, \mathbb{Z}_2 \hookrightarrow \mathcal{U}([1, 3])$ et sont indépendantes.

Afin de s'épargner tout calcul, on remarque :

$$\mathbb{X} + \mathbb{Y} = \min(\mathbb{Z}_1, \mathbb{Z}_2) + \max(\mathbb{Z}_1, \mathbb{Z}_2) = \mathbb{Z}_1 + \mathbb{Z}_2.$$

Comme $\mathbb{Z}_1$ et $\mathbb{Z}_2$ sont indépendantes et admettent une variance :

$$\mathbf{V}(\mathbb{Z}_1 + \mathbb{Z}_2) = \mathbf{V}(\mathbb{Z}_1) + \mathbf{V}(\mathbb{Z}_2) = \frac{3^2 - 1}{12} + \frac{3^2 - 1}{12} = \frac{2}{3} + \frac{2}{3} = \frac{4}{3}.$$

Ainsi, $\mathbb{X} + \mathbb{Y}$ admet une variance qui vaut :

$$\mathbf{V}(\mathbb{X} + \mathbb{Y}) = \mathbf{V}(\mathbb{Z}_1 + \mathbb{Z}_2) = \frac{4}{3}.$$

On peut calculer facilement la variance d'une $\mathcal{B}(n, p)$ à partir de la **proposition (10)**.

En effet, une variable aléatoire réelle de loi binomiale $\mathcal{B}(n, p)$ peut s'écrire comme la somme de n variables aléatoires indépendantes de Bernoulli de paramètre p :

Rappel
Si $\mathbb{Z} \hookrightarrow \mathcal{U}([1, n])$ alors

$$\mathbf{V}(\mathbb{Z}) = \frac{n^2 - 1}{12}.$$

Corollaire 11 – Variance d'une loi binomiale :

Soit $X \hookrightarrow \mathcal{B}(n, p)$. Alors :

$$V(X) = np(1 - p).$$

Preuve : On redécompose X sous la forme d'une somme de n variables aléatoires indépendantes X_i de variance $p(1 - p)$. On a alors :

$$V(X) = V(X_1) + V(X_2) + \dots + V(X_n) = p(1 - p) + p(1 - p) + \dots + p(1 - p) = np(1 - p).$$

**Théorème 12 – Inégalité de Cauchy-Schwarz :**

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbf{P})$ admettant un moment d'ordre 2.

Alors :

$$(\text{Cov}(X, Y))^2 \leq V(X)V(Y).$$

Remarque

L'inégalité de Cauchy-Schwarz s'écrit aussi :

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

Rappel

Si X admet une variance, alors pour tous $a, b \in \mathbb{R}$, $aX + b$ admet une variance et $V(aX + b) = a^2 V(X)$.

Preuve : Supposons que X et Y admettent une variance.

Soit $t \in \mathbb{R}$. La variable aléatoire $tX + Y$ est une combinaison linéaire de deux variables aléatoires admettant une variance, par conséquent, elle admet également une variance et

$$\begin{aligned} V(tX + Y) &= V(tX) + 2\text{Cov}(tX, Y) + V(Y) \\ &= t^2 V(X) + 2t \text{Cov}(X, Y) + V(Y). \quad (\text{linéarité à gauche}) \end{aligned}$$

Distinguons deux cas :

1. Si $V(X) = 0$.

Dans ce cas, d'après ce qui précède :

$$\forall t \in \mathbb{R}, V(tX + Y) = 2t \text{Cov}(X, Y) + V(Y).$$

Fonction affine en t qui doit être toujours positive. C'est le cas si, et seulement si son coefficient directeur est nul i.e. $\text{Cov}(X, Y) = 0$.

L'inégalité de Cauchy-Schwarz est alors vérifiée.

2. Si $V(X) \neq 0$.

Dans ce cas, la fonction $t \mapsto t^2 V(X) + 2t \text{Cov}(X, Y) + V(Y)$ est une fonction polynomiale de degré 2 positive sur $\mathbb{R}$, car égale à la fonction $t \mapsto V(tX + Y)$ et qu'une variance est toujours positive.

Par conséquent, son discriminant est négatif ou nul :

$$\begin{aligned} 4(\text{Cov}(X, Y))^2 - 4V(X)V(Y) &\leq 0 \\ ((\text{Cov}(X, Y))^2 &\leq V(X)V(Y). \end{aligned}$$



Remarque importante : Dans le cas d'égalité i.e. $((\text{Cov}(X, Y))^2 = V(X)V(Y)$, la fonction $t \mapsto V(tX + Y)$ admet alors une unique racine $\alpha \in \mathbb{R}$ telle que $V(\alpha X + Y) = 0$ i.e. la variable aléatoire $\alpha X + Y$ est presque sûrement constante.

On reviendra sur ce point à la **proposition (13)**.

Pour terminer, le coefficient de corrélation linéaire et son interprétation :

II.2 – Coefficient de corrélation linéaire

Définition 2 – Coefficient de corrélation linéaire :

Soient X et Y deux variables aléatoires d'écarts-types non nuls. Le coefficient de corrélation linéaire de (X, Y) est le réel, noté $\rho(X, Y)$, défini par :

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

où $\sigma(X)$ et $\sigma(Y)$ désignent les écarts-types respectifs de X et Y .

Rappel

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

Remarque

$$\rho(X, Y) = \rho(Y, X).$$

Proposition 13 – Interprétation du coefficient de corrélation linéaire :

Soient X et Y deux variables aléatoires de variance non nulle.

P1 • $-1 \leq \rho(X, Y) \leq 1$.

P2 • (1) $\rho(X, Y) = 1$ si, et seulement si, l'une des variables aléatoires est presque-sûrement fonction affine strictly croissante de l'autre.

(2) $\rho(X, Y) = -1$ si, et seulement si, l'une des variables aléatoires est presque-sûrement fonction affine strictement décroissante de l'autre.

Remarques :

- Le coefficient de corrélation linéaire permet donc de mesurer la dépendance linéaire qui peut exister entre un couple de variables aléatoires (X, Y) .
 - Si $\rho(X, Y)$ est proche de 1 ou de -1 , la dépendance linéaire entre X et Y est importante.
Exemple : X = nombre de remises promotionnelles, Y = volume des ventes.
 - Si $\rho(X, Y)$ est proche de 0 les deux variables sont linéairement indépendantes.
Exemple : X = nombre de décès sur les routes de Franche Comté, Y = production de comté.
- Le signe de $\rho(X, Y)$ (et de $\text{Cov}(X, Y)$) indique le sens d'évolution d'une variable par rapport à l'autre :
 - Si $\rho(X, Y) > 0$, alors les deux variables évoluent en moyenne dans le même sens.
Exemple : X = température extérieure, Y = consommation de crèmes glacées.
 - Si $\rho(X, Y) < 0$, les deux variables évoluent en moyenne en sens inverse.
Exemple : X = température extérieure, Y = facture de chauffage.

Vocabulaire : Si $\rho(X, Y) = 0$ i.e. $\text{Cov}(X, Y) = 0$, on dit que X et Y sont non corrélés.

Preuve :

P1 • Découle immédiatement de l'inégalité de Cauchy-Schwarz.

P2 • Supposons que X et Y admettent une variance non nulle.

Remarquons déjà que Y est presque-sûrement fonction affine de X si, et seulement si, X est presque-sûrement fonction affine de Y .

⚠ Attention
 indépendantes $\Rightarrow$ non
 corrélées mais la réciproque
 est fausse ! *confer* le contre-
 exemple (9) .

Dans la suite, on considère donc, par exemple, que $\mathbb{Y}$ est fonction affine de $\mathbb{X}$.

$$\begin{aligned}
 |\rho(\mathbb{X}, \mathbb{Y})| = 1 &\iff \rho(\mathbb{X}, \mathbb{Y})^2 = 1 \\
 &\iff \left(\frac{\mathbf{Cov}(\mathbb{X}, \mathbb{Y})}{\sigma(\mathbb{X})\sigma(\mathbb{Y})} \right)^2 = 1 \\
 &\iff \frac{(\mathbf{Cov}(\mathbb{X}, \mathbb{Y}))^2}{\mathbf{V}(\mathbb{X})\mathbf{V}(\mathbb{Y})} = 1 \\
 &\iff (\mathbf{Cov}(\mathbb{X}, \mathbb{Y}))^2 = \mathbf{V}(\mathbb{X})\mathbf{V}(\mathbb{Y}) \\
 &\iff \text{il y a égalité dans l'inégalité de Cauchy-Schwarz} \\
 &\iff \text{le discriminant de } t \mapsto \mathbf{V}(t\mathbb{X} + \mathbb{Y}) \text{ est nul} \\
 &\iff t \mapsto \mathbf{V}(t\mathbb{X} + \mathbb{Y}) \text{ possède une unique racine } \alpha \in \mathbb{R}. \\
 &\iff \text{il existe un unique réel } \alpha \text{ tel que } \mathbf{V}(\alpha\mathbb{X} + \mathbb{Y}) = 0 \\
 &\iff \text{il existe un unique réel } \alpha \text{ tel que la variable aléatoire } \alpha\mathbb{X} + \mathbb{Y} \text{ est presque-} \\
 &\quad \text{sûrement constante} \\
 &\iff \exists ! \alpha, b \in \mathbb{R} \text{ tels que } \mathbf{P}([\alpha\mathbb{X} + \mathbb{Y} = b]) = 1 \\
 &\iff \exists ! a, b \in \mathbb{R} \text{ tels que } \mathbf{P}([\mathbb{Y} = a\mathbb{X} + b]) = 1.
 \end{aligned}$$

Considérons ensuite ces réels a et b . On a ainsi, d'après la formule de Koenig-Huygens :

$$\begin{aligned}
 \mathbf{Cov}(\mathbb{X}, \mathbb{Y}) &= \mathbf{E}(\mathbb{X}\mathbb{Y}) - \mathbf{E}(\mathbb{X})\mathbf{E}(\mathbb{Y}) \\
 &= \mathbf{E}(\mathbb{X}(a\mathbb{X} + b)) - \mathbf{E}(\mathbb{X})\mathbf{E}(a\mathbb{X} + b) \\
 &= a\mathbf{E}(\mathbb{X}^2) + b\mathbf{E}(\mathbb{X}) - a\mathbf{E}(\mathbb{X})^2 - b\mathbf{E}(\mathbb{X}) \\
 &= a(\mathbf{E}(\mathbb{X}^2) - \mathbf{E}(\mathbb{X})^2) \\
 &= a\mathbf{V}(\mathbb{X})
 \end{aligned}
 \quad \text{(formule de Koenig-Huygens)}$$

Et ainsi :

$$\begin{aligned}
 \rho(\mathbb{X}, \mathbb{Y}) &= \frac{a\mathbf{V}(\mathbb{X})}{\sigma(\mathbb{X})\sigma(\mathbb{Y})} = \frac{a\mathbf{V}(\mathbb{X})}{\sqrt{\mathbf{V}(\mathbb{X})}\sqrt{\mathbf{V}(a\mathbb{X} + b)}} = \frac{a\mathbf{V}(\mathbb{X})}{|a|\sqrt{\mathbf{V}(\mathbb{X})}\sqrt{\mathbf{V}(\mathbb{X})}} = \frac{a}{|a|} \\
 &= \text{signe de } a.
 \end{aligned}$$

À retenir

$$\mathbf{P}([\mathbb{X} = \mathbb{Y}]) = 1$$

↓

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

Exemple 8 :

Reprenons une dernière fois les données de l'exemple (3) et déterminons $\rho(\mathbb{X}, \mathbb{Y})$.

$\mathbb{X} \backslash \mathbb{Y}$	1	2	3
1	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$
2	0	$\frac{1}{9}$	$\frac{2}{9}$
3	0	0	$\frac{1}{9}$

Rapidement, on a :

$$\text{— } \mathbf{E}(\mathbb{X}) = \frac{14}{9} \quad \text{et} \quad \mathbf{E}(\mathbb{Y}) = \frac{22}{9}.$$

$$— \mathbf{E}(XY) = 4.$$

$$— \mathbf{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y) = 4 - \frac{14}{9} \frac{22}{9} = \frac{16}{81}.$$

$$\begin{aligned} — \mathbf{V}(X) &= \mathbf{E}(X^2) - \mathbf{E}(X)^2 \\ &= \frac{1^2 \times 5 + 2^2 \times 3 + 3^2 \times 1}{9} - \left(\frac{14}{9}\right)^2 = \frac{26}{9} - \left(\frac{14}{9}\right)^2 \\ &= \frac{9 \times 26}{81} - \frac{14^2}{81} = \frac{234 - 196}{81} = \frac{38}{81} \end{aligned}$$

$$— \text{De même, } \mathbf{V}(Y) = \frac{9 \times 58 - 22^2}{81} = \frac{38}{81}.$$

Finalement,

$$\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\sqrt{\mathbf{V}(X)}\sqrt{\mathbf{V}(Y)}} = \frac{\frac{16}{81}}{\sqrt{\frac{38}{81}}\sqrt{\frac{38}{81}}} = \frac{16}{81} \times \frac{81}{38} = \frac{16}{38} = \frac{8}{19}.$$

Contre-exemple 9 – Suite de l'exemple (5) :

On a trouvé :

$$— \mathbf{E}(XY) = 0 \text{ et } \mathbf{E}(X) = \mathbf{E}(Y) = 0.$$

D'après la formule de Koenig-Huygens, on a alors $\mathbf{Cov}(X, Y) = 0$.

— les variables X et Y ne sont pas indépendantes.

Les variables aléatoires X et Y sont non corrélées, mais ne sont pas indépendantes !

Dernière remarque : Par la formule des racines des polynômes de second degré, on obtient que l'unique racine α de

$$\mathbf{V}(tX + Y) = t^2\mathbf{V}(X) + 2t\mathbf{Cov}(X, Y) + \mathbf{V}(Y)$$

s'écrit sous la forme :

$$\begin{aligned} \alpha &= -\frac{\cancel{X}\mathbf{Cov}(X, Y)}{\cancel{X}\mathbf{V}(X)} = -\frac{\mathbf{Cov}(X, Y)}{\sigma(X)\sigma(Y)} \times \frac{\sigma(X)\sigma(Y)}{\mathbf{V}(X)} \\ &= -\rho(X, Y) \times \frac{\sigma(X)\sigma(Y)}{\sigma^2(X)} \\ &= -\rho(X, Y) \frac{\sigma(Y)}{\sigma(X)} \end{aligned}$$

On en déduit finalement, d'après ce qui précède :

$$\rho(X, Y) = 1 \iff \exists \beta \in \mathbb{R}, Y = \frac{\sigma(Y)}{\sigma(X)}X + \beta \text{ presque sûrement}$$

et

$$\rho(X, Y) = -1 \iff \exists \beta \in \mathbb{R}, Y = -\frac{\sigma(Y)}{\sigma(X)}X + \beta \text{ presque sûrement.}$$

Les variables X et Y sont liées par une relation affine.

Rappel

Dans le cas où « $\Delta = 0$ », l'unique racine α de

$$at^2 + bt + c$$

est

$$\alpha = -\frac{b}{2a}.$$

⚠ Attention

Attention encore une fois de ne pas confondre « indépendance linéaire » et « indépendance » : si $\rho(X, Y) = 0$ (ou $\text{Cov}(X, Y) = 0$), alors X et Y sont linéairement indépendantes.

Mais elles ne sont pas forcément indépendantes : il peut exister une dépendance non linéaire entre les variables, par exemple $Y = e^X, \ln(X), \dots$

Un peu d'histoire : Corrélation vs Indépendance

L'indépendance et la corrélation ou plus exactement la dépendance et la corrélation sont deux notions en statistique qui sont souvent confondues. Formellement, l'indépendance entre deux variables X et Y indique que la connaissance de la réalisation de l'une des variables n'a aucune incidence sur la probabilité de réalisation de l'autre variable. La corrélation quant à elle renvoie à l'existence d'une liaison linéaire entre les deux variables.

Donc l'indépendance est une notion liée à la « chance » de réalisation des événements des variables X et Y tandis que la corrélation est une notion liée à la fonction linéaire qui existerait entre les valeurs des deux variables.

Il est théoriquement démontré que si deux variables sont indépendantes alors elles ne sont pas corrélées. Autrement dit, si les probabilités de réalisation des événements liés aux variables X et Y vérifient la condition d'indépendance alors les valeurs de ces variables n'ont pas de liaison linéaire entre elles. Par conséquent, si deux variables sont corrélées alors elles sont nécessairement dépendantes (c'est-à-dire elles ne sont pas indépendantes).

Ce résultat laisse alors la possibilité aux situations où des variables seraient dépendantes sans être corrélées.

III. BILAN ET COMPÉTENCES ATTENDUES

Les compétences attendues sur le chapitre couple de variables aléatoires réelles sont les suivantes :

- savoir déterminer l'espérance d'une somme (linéarité !), d'un produit (théorème de transfert).
- connaître la définition de la covariance et savoir calculer $\text{Cov}(X, Y)$. Il faut connaître les propriétés (linéarité à droite et à gauche) de l'opérateur Cov .
- savoir déterminer la variance d'une somme (formule dans le cas général et formule dans le cas de l'indépendance).
- savoir calculer un coefficient de corrélation linéaire et savoir commenter sa valeur (proximité de 1, -1 et signe).
- savoir démontrer que deux variables aléatoires ne sont pas indépendantes via le calcul de l'espérance du produit / de la covariance / du coefficient de corrélation linéaire (3 formulations différentes de la même propriété).

INDEX

- Coefficient
 - de corrélation linéaire, 15
- Corrélation, 15
 - Interprétation, 15
- Covariance, 10, 12
- Espérance, 1, 2, 8
 - Croissance de l' , 6
 - d'un maximum, 3
 - d'un produit, 6, 9
 - d'une somme, 5
 - Linéarité de l' , 5
- Formule
 - de Koenig-Huygens, 11
- Inégalité
 - de Cauchy-Schwarz, 14
- Loi
 - binomiale, 14
 - de Bernoulli, 10, 13
- Série
 - absolument convergente, 2
 - double, 1
- Théorème
 - de Cauchy-Schwarz, 14
 - de transfert, 1
- Transfert, 1
- Variable
 - indépendante, 9–12
 - non corrélée, 15
- Variance
 - d'une somme, 12