# 2 | Couples de variables aléatoires

Dans tout ce chapitre $\Omega$ désigne un univers **fini**. Ainsi les variables aléatoires réelles considérées ne prennent qu'un nombre **fini** de valeurs.

## I – Lois de probabilités

### 1 – Loi d'un couple de variables aléatoires

**Définition 2.1 –** On appelle **couple de variables aléatoires**, tout couple $(X, Y)$ où $X$ et $Y$ désignent deux variables aléatoires définies sur un même ensemble $\Omega$ *(l'univers)*.

**Exemple 2.2 –** Tout au long de ce chapitre, on s'appuie sur les deux exemples suivants pour illustrer les différentes notions rencontrées :

1. On lance deux dés équilibrés à six faces, l'un rouge, l'autre noir. On appelle $X$ (resp. $Y$) le numéro obtenu avec le dé rouge (resp. noir).
   Comme $X$ et $Y$ sont des variables aléatoires, alors $(X, Y)$ est un couple de variables aléatoires.
2. On lance les mêmes dés que dans l'exemple précédent. Cette fois, on appelle $X$ le plus petit des deux numéros obtenus et $Y$ le plus grand numéro obtenu (si les numéros sont égaux, $X$ et $Y$ prennent la valeur commune).
   Comme $X$ et $Y$ sont des variables aléatoires, alors $(X, Y)$ est un couple de variables aléatoires.

**Définition 2.3 –** Soit $(X, Y)$ un couple de variables aléatoires. On appelle **loi conjointe** du couple $(X, Y)$ la donnée de toutes les probabilités $P\big([X = x] \cap [Y = y]\big)$ pour tout couple $(x, y) \in X(\Omega) \times Y(\Omega)$.

**Méthode 2.4 – Déterminer la loi conjointe d'un couple de variables aléatoires**

1. On détermine les supports $X(\Omega)$ et $Y(\Omega)$, ensembles des valeurs prises par $X$ et $Y$.
2. On calcule les probabilités $P\big([X = x] \cap [Y = y]\big)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.

Comme pour une variable aléatoire unique, on résume souvent la loi sous la forme d'un tableau, cette fois à double entrée. La somme de toutes les probabilités est toujours égale à 1.

**Exemple 2.5 –** Donner la loi conjointe des couples $(X, Y)$ dans les deux exemples précédents.

1. Les variables aléatoires étant chacune égale au résultat d'un dé, les deux supports sont donnés par $[\![1, 6]\!]$. En outre, chaque dé étant équilibré les $6 \times 6 = 36$ possibilités sont équiprobables.
   Ainsi pour tout $(x, y) \in [\![1, 6]\!] \times [\![1, 6]\!]$, $P\big([X = x] \cap [Y = y]\big) = \dfrac{1}{36}$.
2. Les supports sont les mêmes, mais pas les probabilités.
   En effet, $X$ ne peut pas être plus grand que $Y$ : si $x > y$, alors $P\big([X = x] \cap [Y = y]\big) = 0$.
   Puis en notant $R$ le résultat du dé rouge et $N$ celui du dé noir, alors pour $x < y$,
   $$P\big([X = x] \cap [Y = y]\big) = P\big([R = x] \cap [N = y]\big) + P\big([R = y] \cap [N = x]\big) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36} = \frac{1}{18}.$$
   Enfin si $x = y$, alors $P\big([X = x] \cap [Y = y]\big) = \dfrac{1}{36}$.

J'en déduis les deux tableaux suivants pour les deux lois conjointes :

1.

| | $Y=1$ | $Y=2$ | $Y=3$ | $Y=4$ | $Y=5$ | $Y=6$ |
|---|---|---|---|---|---|---|
| $X=1$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ |
| $X=2$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ |
| $X=3$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ |
| $X=4$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ |
| $X=5$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ |
| $X=6$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ | $\frac{1}{36}$ |

2.

| | $Y=1$ | $Y=2$ | $Y=3$ | $Y=4$ | $Y=5$ | $Y=6$ |
|---|---|---|---|---|---|---|
| $X=1$ | $\frac{1}{36}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ |
| $X=2$ | $0$ | $\frac{1}{36}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ |
| $X=3$ | $0$ | $0$ | $\frac{1}{36}$ | $\frac{1}{18}$ | $\frac{1}{18}$ | $\frac{1}{18}$ |
| $X=4$ | $0$ | $0$ | $0$ | $\frac{1}{36}$ | $\frac{1}{18}$ | $\frac{1}{18}$ |
| $X=5$ | $0$ | $0$ | $0$ | $0$ | $\frac{1}{36}$ | $\frac{1}{18}$ |
| $X=6$ | $0$ | $0$ | $0$ | $0$ | $0$ | $\frac{1}{36}$ |

**Remarque 2.6 –**

- On abrège souvent "loi conjointe du couple" en "loi du couple".
- On note parfois $P\big([X=x],[Y=y]\big)$ au lieu de $P\big([X=x]\cap[Y=y]\big)$, même simplement $P(X=x,Y=y)$.

## 2 – Lois marginales

**Définition 2.7 –** Soit $(X,Y)$ un couple de variables aléatoires. La loi de $X$ est appelée **première loi marginale** du couple et celle de $Y$ est appelée **deuxième loi marginale** du couple.

**Proposition 2.8**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$.

- Pour tout réel $x\in X(\Omega)$,
$$P(X=x) = \sum_{y\in Y(\Omega)} P(X=x,Y=y).$$
- Pour tout réel $y\in Y(\Omega)$,
$$P(Y=y) = \sum_{x\in X(\Omega)} P(X=x,Y=y).$$

**Méthode 2.9 – Déterminer les lois marginales avec la loi du couple**
Une fois que l'on a déterminé la loi du couple, on peut déterminer les lois marginales.
La loi de $X$ s'écrit par exemple

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

Lorsque la loi d'un couple $(X, Y)$ est donnée sous la forme d'un tableau à double entrée, on obtient les lois de $X$ et de $Y$ en sommant les éléments d'une même ligne ou d'une même colonne, selon les cas.

**Exemple 2.10 –** Déterminer les lois marginales de $X$ et $Y$ dans les deux exemples précédents.

1. Pour obtenir la loi de $X$, je calcule la somme de chacune des lignes. J'obtiens :

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(X = x)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

Pour obtenir la loi de $Y$, je calcule la somme de chacune des colonnes. J'obtiens :

| $y$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(Y = y)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

2. Pour obtenir la loi de $X$, je calcule la somme de chacune des lignes. J'obtiens :

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(X = x)$ | $\frac{11}{36}$ | $\frac{9}{36} = \frac{1}{4}$ | $\frac{7}{36}$ | $\frac{5}{36}$ | $\frac{3}{36} = \frac{1}{12}$ | $\frac{1}{36}$ |

Pour obtenir la loi de $Y$, je calcule la somme de chacune des colonnes. J'obtiens :

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(Y = x)$ | $\frac{1}{36}$ | $\frac{3}{36} = \frac{1}{12}$ | $\frac{5}{36}$ | $\frac{7}{36}$ | $\frac{9}{36} = \frac{1}{4}$ | $\frac{11}{36}$ |

**Remarque 2.11 –** Si l'établissement des lois marginales découle directement de la donnée de la loi conjointe, il est en revanche impossible, en général, d'obtenir la loi conjointe à partir des deux lois marginales.

## 3 – Lois conditionnelles

**Définition 2.12 –** Soit $(X, Y)$ un couple de variables aléatoires. Pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, on appelle loi de $X$ **conditionnellement à l'événement** $[Y = y]$ la donnée, pour tout $x \in X(\Omega)$, de

$$P_{[Y=y]}(X = x) = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

**Remarque 2.13 –**

- On dit aussi "loi conditionnelle de $X$ sachant que $[Y = y]$ est réalisé", ou plus simplement "loi de $X$ sachant $[Y = y]$".
- On définit de manière similaire la loi conditionnelle de $Y$ sachant $[X = x]$.

**Exemple 2.14 –** Dans les deux exemples précédents, $P(Y=1) \neq 0$.
Déterminer alors la loi conditionnelle de $X$ sachant $[Y=1]$ dans les deux cas.

1. Je sais déjà que $P(Y=1) = \frac{1}{6}$ et que pour tout $x \in [\![1,6]\!]$, $P(X=x, Y=1) = \frac{1}{36}$.

   Donc pour tout entier $x \in [\![1,6]\!]$,

   $$P_{[Y=1]}(X=x) = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{36} \times \frac{6}{1} = \frac{1}{6}.$$

   Je résume cela dans le tableau suivant :

   | $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
   |---|---|---|---|---|---|---|
   | $P_{[Y=1]}(X=x)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

2. Je sais déjà que $P(Y=1) = \frac{1}{36}$ et que pour tout $x \in [\![2,6]\!]$, $P(X=x, Y=1) = 0$.

   Donc pour tout entier $x \in [\![2,6]\!]$,

   $$P_{[Y=1]}(X=x) = \frac{0}{\frac{1}{36}} = 0.$$

   Par ailleurs $P(X=1, Y=1) = \frac{1}{36}$ donc $P_{[Y=1]}(X=1) = \frac{\frac{1}{36}}{\frac{1}{36}} = 1$.
   Je résume cela dans le tableau suivant :

   | $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
   |---|---|---|---|---|---|---|
   | $P_{[Y=1]}(X=x)$ | 1 | 0 | 0 | 0 | 0 | 0 |

   Cette loi conditionnelle est une loi certaine : si $Y=1$, nécessairement $X=1$.

**Proposition 2.15 – Loi marginale et loi conditionnelle**

Soit $(X,Y)$ un couple de variables aléatoires. Si l'on connaît la loi marginale de $Y$, ainsi que toutes les lois conditionnelles de $X$ sachant $[Y=y]$ pour tous les $y \in Y(\Omega)$, alors la loi de $X$ est déterminée par

$$\forall x \in X(\Omega), \quad P(X=x) = \sum_{y \in Y(\Omega)} P(Y=y) \times P_{[Y=y]}(X=x).$$

**Exemple 2.16 –** J'ai calculé la loi conditionnelle de $X$ sachant $[Y=1]$. Si je calculais les lois conditionnelles de $X$ sachant $[Y=2]$, $[Y=3]$, etc., dans les deux exemples précédents, alors je pourrais retrouver la loi marginale de $X$ grâce à la proposition ci-dessus.

## 4 – Indépendance de deux variables aléatoires

**Définition 2.17 –** On dit que deux variables aléatoires finies $X$ et $Y$ sont **indépendantes** lorsque

$$\forall (x,y) \in X(\Omega) \times Y(\Omega), \quad P(X=x, Y=y) = P(X=x) \times P(Y=y).$$

**Remarque 2.18 – Dans le cas de deux variables aléatoires indépendantes**, on peut déterminer la loi du couple $(X,Y)$ à partir des lois de $X$ et de $Y$.

**Exemple 2.19 –** Tester l'indépendance des variables aléatoires $X$ et $Y$ dans les exemples précédents.

1. Pour tout $x \in [\![1,6]\!]$ et tout $y \in [\![1,6]\!]$,

$$P(X=x) = \frac{1}{6}, \qquad P(Y=y) = \frac{1}{6} \quad \text{et} \quad P(X=x, Y=y) = \frac{1}{36}.$$

   En particulier, pour tout couple $(x,y) \in X(\Omega) \times Y(\Omega)$,

$$P(X=x, Y=y) = P(X=x) \times P(Y=y).$$

   Ainsi les variables aléatoires $X$ et $Y$ sont indépendantes.

2. Dans le second exemple, $P(X=2, Y=1) = 0$ alors que

$$P(X=2) = \frac{9}{36} = \frac{1}{4} \quad \text{et} \quad P(Y=1) = \frac{1}{36}.$$

   En particulier, $P(X=2, Y=1) \neq P(X=2) \times P(Y=1)$.
   Donc les variables aléatoires $X$ et $Y$ ne sont pas indépendantes.

**Proposition 2.20**

Si l'une des deux variables aléatoires $X$ ou $Y$ est constante, alors $X$ et $Y$ sont indépendantes.

# II – Espérance

## 1 – Espérance d'une somme

**Proposition 2.21**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$. Alors

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

**Exemple 2.22 –** Soient $X$ une variable aléatoire qui suit la loi uniforme sur $[\![1,9]\!]$ et $Y$ une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(8, \frac{1}{4}\right)$. Calculer l'espérance de la variable aléatoire $Z = X+Y$.

Puisque $X$ suit la loi uniforme sur $[\![1,9]\!]$, alors $E(X) = \frac{9+1}{2} = 5$.

De même, puisque $Y$ suit la loi binomiale $\mathcal{B}\left(8, \frac{1}{4}\right)$, alors $E(Y) = 8 \times \frac{1}{4} = 2$.

Ainsi

$$E(Z) = E(X) + E(Y) = 5 + 2 = 7.$$

**Proposition 2.23 – Linéarité de l'espérance**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$ et $a$ et $b$ deux réels. Alors

$$E(aX + bY) = aE(X) + bE(Y).$$

**Exemple 2.24 –** Soient $X$ une variable aléatoire qui suit la loi uniforme sur $[\![1,12]\!]$ et $Y$ une variable aléatoire qui suit la loi binomiale $\mathcal{B}\left(7,\frac{1}{3}\right)$. Calculer l'espérance de la variable aléatoire $Z = 2X - Y$.

Puisque $X$ suit une loi uniforme, alors $E(X) = \frac{12+1}{2} = \frac{13}{2}$ et puisque $Y$ suit une loi binomiale, alors $E(Y) = 7 \times \frac{1}{3} = \frac{7}{3}$.
Ainsi

$$E(Z) = E(2X-Y) = 2E(X) - E(Y) = 2 \times \frac{13}{2} - \frac{7}{3} = 13 - \frac{7}{3} = \frac{39-7}{3} = \frac{32}{3}.$$

## 2 – Espérance d'un produit

**Proposition 2.25**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$. Alors l'espérance du produit est définie grâce à la loi conjointe par

$$E(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X=x, Y=y).$$

**Exemple 2.26 –**

1. Un sac contient quatre boules numérotées de 1 à 4. On effectue deux tirages successifs d'une boule, avec remise. On note $X_1$ le numéro de la première boule, $X_2$ le numéro de la deuxième boule et $Y$ le plus grand des deux numéros obtenus.
   Compléter les tableaux suivants, donnant les lois des couples $(X_1, X_2)$ et $(X_1, Y)$.

| | $X_2=1$ | $X_2=2$ | $X_2=3$ | $X_2=4$ |
|---|---|---|---|---|
| $X_1=1$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ |
| $X_1=2$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ |
| $X_1=3$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ |
| $X_1=4$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ |

| | $Y=1$ | $Y=2$ | $Y=3$ | $Y=4$ |
|---|---|---|---|---|
| $X_1=1$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ | $\frac{1}{16}$ |
| $X_1=2$ | $0$ | $\frac{2}{16}=\frac{1}{8}$ | $\frac{1}{16}$ | $\frac{1}{16}$ |
| $X_1=3$ | $0$ | $0$ | $\frac{3}{16}$ | $\frac{1}{16}$ |
| $X_1=4$ | $0$ | $0$ | $0$ | $\frac{4}{16}=\frac{1}{4}$ |

   En déduire $E(X_1X_2)$ et $E(X_1Y)$.

   J'applique la formule précédente en sommant, pour chaque $x \in [\![1,4]\!]$ et chaque $y \in [\![1,4]\!]$, les produits $x \times y \times P(X=x, Y=y)$ :

$$\begin{aligned} E(X_1X_2) &= 1\times 1\times \frac{1}{16} + 1\times 2\times\frac{1}{16} + \cdots + 4\times 4\times \frac{1}{16} \\ &= \frac{1+2+3+4+2+4+6+8+3+6+9+12+4+8+12+16}{16} = \frac{100}{16} = \frac{25}{4}. \end{aligned}$$

$$\begin{aligned} E(X_1Y) &= 1\times 1\times \frac{1}{16} + 1\times 2\times\frac{1}{16} + \cdots + 4\times 4\times \frac{4}{16} \\ &= \frac{1+2+3+4+8+6+8+27+12+64}{16} = \frac{135}{16}. \end{aligned}$$

2. Soit $(X,Y)$ un couple de variables aléatoires finies tel que

$$\forall (i,j) \in [\![1,2]\!]^2, \quad P(X=i, Y=j) = \begin{cases} \frac{i}{4} & \text{si } i = j, \\ \frac{1}{4} & \text{si } i < j, \\ 0 & \text{si } i > j. \end{cases}$$

Calculer $E(XY)$.

Afin de mieux comprendre la loi conjointe du couple $(X,Y)$, je calcule les quatre probabilités et récapitule la loi dans un tableau :

| | $Y=1$ | $Y=2$ |
|---|---|---|
| $X=1$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $X=2$ | $0$ | $\frac{2}{4}=\frac{1}{2}$ |

Donc en additionnant les quatre produits correspondants, j'obtiens

$$E(XY) = 1\times 1\times \frac{1}{4} + 1\times 2\times \frac{1}{4} + 2\times 1\times 0 + 2\times 2\times \frac{2}{4} = \frac{1+2+8}{4} = \frac{11}{4}.$$

**Proposition 2.27**

Soient $X$ et $Y$ deux variables aléatoires **indépendantes** définies sur $\Omega$. Alors

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

**Exemple 2.28 –** Même exemple que précédemment : un sac contient quatre boules numérotées. On effectue deux tirages successifs d'une boule, avec remise. On note $X_1$ le numéro de la première boule, $X_2$ le numéro de la deuxième boule et $Y$ le plus grand des deux numéros obtenus.

1. Déterminer les lois marginales de $X_1$, $X_2$ et $Y$.

Grâce aux tableaux des lois conjointes de $(X_1,X_2)$ et de $(X_1,Y)$ réalisés précédemment, j'obtiens les tableaux des lois marginales de $X_1$, $X_2$ et $Y$ en sommant les probabilités des lignes/colonnes correspondantes :

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P(X_1=x)$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P(X_2=x)$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

| $y$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $P(Y=y)$ | $\frac{1}{16}$ | $\frac{3}{16}$ | $\frac{5}{16}$ | $\frac{7}{16}$ |

2. En déduire les valeurs de $E(X_1)$, $E(X_2)$ et $E(Y)$.

Je me sers cette fois de la définition de l'espérance d'une variable aléatoire unique, j'additionne les produits des valeurs multipliées par les probabilités associées :

$$E(X_1) = 1\times\frac{1}{4} + 2\times\frac{1}{4} + 3\times\frac{1}{4} + 4\times\frac{1}{4} = \frac{10}{4} = \frac{5}{2}$$

$$E(X_2) = 1\times\frac{1}{4} + 2\times\frac{1}{4} + 3\times\frac{1}{4} + 4\times\frac{1}{4} = \frac{10}{4} = \frac{5}{2}$$

$$E(Y) = 1\times\frac{1}{16} + 2\times\frac{3}{16} + 3\times\frac{5}{16} + 4\times\frac{7}{16} = \frac{50}{16} = \frac{25}{8}$$

3. Les variables aléatoires $X_1$ et $X_2$ sont-elles indépendantes ? Et les variables aléatoires $X_1$ et $Y$ ?
Pour tout $x \in [\![1,4]\!]$ et tout $y \in [\![1,4]\!]$,
$$P(X_1 = x) = \frac{1}{4}, \qquad P(X_2 = y) = \frac{1}{4} \quad \text{et} \quad P(X_1 = x, X_2 = y) = \frac{1}{16}.$$
En particulier, pour tout couple $(x, y) \in X(\Omega) \times Y(\Omega)$,
$$P(X = x, Y = y) = P(X = x) \times P(Y = y).$$
Donc les variables aléatoires $X_1$ et $X_2$ sont indépendantes. Par ailleurs, je sais que
$$E(X_1Y) = \frac{135}{16} \quad \text{et} \quad E(X_1) \times E(Y) = \frac{5}{2} \times \frac{25}{8} = \frac{125}{16}.$$
Ainsi $E(XY) \neq E(X) \times E(Y)$, donc les variables aléatoires $X_1$ et $Y$ ne sont pas indépendantes.

**ATTENTION !** L'égalité $E(XY) = E(X)E(Y)$ peut être vérifiée sans que les variables aléatoires $X$ et $Y$ ne soient indépendantes.

# III – Covariance, corrélation linéaire

## 1 – Covariance de deux variables aléatoires

**Définition 2.29 –** Soit $(X, Y)$ un couple de variables aléatoires. On appelle **covariance de** $X$ **et** $Y$, le réel, noté $\text{Cov}(X, Y)$, défini par
$$\text{Cov}(XY) = E\Big(\big(X - E(X)\big)\big(Y - E(Y)\big)\Big).$$

**Théorème 2.30 – Formule de König-Huygens**

Soit $(X, Y)$ un couple de variables aléatoires. Alors
$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

*Démonstration.*
On développe le produit à l'intérieur de l'espérance dans la définition :
$$\begin{aligned} E\Big(\big(X - E(X)\big)\big(Y - E(Y)\big)\Big) &= E\big(XY - X \times E(Y) - E(X) \times Y + E(X)E(Y)\big) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$
□

**Méthode 2.31 – Calculer directement une covariance**
Pour calculer la covariance de deux variables aléatoires $X$ et $Y$ :

1. On calcule les trois espérances $E(X)$, $E(Y)$ et $E(XY)$ si ce n'est pas déjà fait.
2. On applique la formule de König-Huygens : $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.

**Exemple 2.32 –** On reprend les deux exemples précédents.

1. Calculer $\mathrm{Cov}(X_1, X_2)$ et $\mathrm{Cov}(X_1, Y)$.

   Je connais déjà toutes les espérances impliquées dans les calculs, je peux donc appliquer la formule de König-Huygens directement :

$$\mathrm{Cov}(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2) = \frac{25}{4} - \frac{5}{2} \times \frac{5}{2} = 0,$$
$$\mathrm{Cov}(X_1, Y) = E(X_1Y) - E(X_1)E(Y) = \frac{135}{16} - \frac{5}{2} \times \frac{25}{8} = \frac{135}{16} - \frac{125}{16} = \frac{10}{16} = \frac{5}{8}.$$

2. Calculer $\mathrm{Cov}(X, Y)$.

   Cette fois, je ne connais que l'espérance de la loi conjointe, je dois donc calculer les espérances des lois marginales. Je rappelle le tableau de la loi conjointe afin de déterminer en premier lieu les lois marginales de $X$ et de $Y$ :

| | $Y = 1$ | $Y = 2$ |
|---|---|---|
| $X = 1$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $X = 2$ | $0$ | $\frac{1}{2}$ |

   Je somme alors les probabilités des lignes pour obtenir la loi marginale de $X$ et les colonnes pour celle de $Y$. J'obtiens :

| $x$ | $1$ | $2$ |
|---|---|---|
| $P(X = x)$ | $\frac{1}{2}$ | $\frac{1}{2}$ |

| $x$ | $1$ | $2$ |
|---|---|---|
| $P(Y = x)$ | $\frac{1}{4}$ | $\frac{3}{4}$ |

   J'en déduis alors les espérances :

$$E(X) = 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{1+2}{2} = \frac{3}{2} \quad \text{et} \quad E(Y) = 1 \times \frac{1}{4} + 2 \times \frac{3}{4} = \frac{1+6}{4} = \frac{7}{4}.$$

   Puis à l'aide de la formule de König-Huygens,

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{11}{4} - \frac{3}{2} \times \frac{7}{4} = \frac{11}{4} - \frac{21}{8} = \frac{22-21}{8} = \frac{1}{8}.$$

**Proposition 2.33 – Propriétés de la covariance**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$.

- La covariance est symétrique :
$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X).$$
- La covariance d'une variable aléatoire avec elle-même est égale à sa variance :
$$\mathrm{Cov}(X, X) = V(X).$$
- Cas d'une variable aléatoire constante : si $a$ est un réel, alors
$$\mathrm{Cov}(X, a) = 0.$$

**Proposition 2.34 – Linéarité à gauche et à droite de la covariance**

Soient $X$, $X_1$, $X_2$, $Y$, $Y_1$ et $Y_2$ des variables aléatoires définies sur $\Omega$. Pour tout couple $(a,b) \in \mathbb{R}^2$,

$$\mathrm{Cov}(aX_1 + bX_2, Y) = a\,\mathrm{Cov}(X_1, Y) + b\,\mathrm{Cov}(X_2, Y),$$

$$\mathrm{Cov}(X, aY_1 + bY_2) = a\,\mathrm{Cov}(X, Y_1) + b\,\mathrm{Cov}(X, Y_2).$$

**Proposition 2.35**

Soient $X$ et $Y$ deux variables aléatoires **indépendantes** définies sur $\Omega$. Alors

$$\mathrm{Cov}(X, Y) = 0.$$

**Remarque 2.36 –**

- C'est une conséquence directe de la Proposition 2.27.
- La réciproque est fausse : il se peut que $\mathrm{Cov}(X,Y) = 0$ sans que les variables aléatoires $X$ et $Y$ ne soient indépendantes.

**Méthode 2.37 – Montrer que deux variables aléatoires ne sont pas indépendantes**

Ceci est un récapitulatif des résultats à disposition pour montrer que deux variables aléatoires **ne sont pas** indépendantes. Pour rappel, si elles sont indépendantes, il n'y a aucun autre moyen que de montrer par le calcul que chaque probabilité de la loi conjointe s'obtient comme le produit des probabilités des lois marginales.

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$.

- Si la covariance $\mathrm{Cov}(X,Y)$ est non nulle, alors les variables aléatoires $X$ et $Y$ ne sont pas indépendantes.
- Si les espérances ne satisfont pas l'égalité $E(XY) = E(X)E(Y)$, alors les variables aléatoires $X$ et $Y$ ne sont pas indépendantes.
- S'il existe un couple $(x,y) \in X(\Omega) \times Y(\Omega)$ pour qui l'égalité $P(X = x, Y = y) = P(X = x) \times P(Y = y)$ n'est pas vérifiée, alors les variables aléatoires $X$ et $Y$ ne sont pas indépendantes.
  Ce dernier point est souvent le plus facile à utiliser lorsqu'un couple présente une probabilité nulle dans le tableau de la loi conjointe.

## 2 – Variance d'une somme

**Proposition 2.38**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$. Alors

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

**Méthode 2.39 – Calculer la variance d'une somme**

Pour calculer la variance d'une somme de variables aléatoires, il y a deux possibilités :

- Si on connaît la loi de la somme $X+Y$, on utilise la **formule de König-Huygens** :
$$V(X+Y) = E\left((X+Y)^2\right) - E(X+Y)^2.$$
- Si on ne connaît pas la loi de la somme $X+Y$, on utilise plutôt la formule précédente :
$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{Cov}(X,Y).$$

**Exemple 2.40 –** On reprend les deux exemples précédents.

1. Calculer $V(X_1+X_2)$ et $V(X_1+Y)$.

   Comme je ne connais pas la loi de $X_1+X_2$, j'utilise la formule
   $$V(X_1+X_2)=V(X_1)+V(X_2)+2\,\mathrm{Cov}(X_1,X_2).$$
   Je connais déjà $\mathrm{Cov}(X_1,X_2)$. Il me reste à calculer $V(X_1)$ et $V(X_2)$.
   D'après la formule de König-Huygens,
   $$V(X_1)=E\left(X_1^2\right)-E(X_1)^2 \quad \text{et} \quad V(X_2)=E\left(X_2^2\right)-E(X_2)^2.$$
   Je dois alors calculer $E\left(X_1^2\right)$ et $E\left(X_2^2\right)$. Or d'après le théorème de transfert,
   $$E\left(X_1^2\right)=E\left(X_2^2\right)=1^2\times\frac{1}{4}+2^2\times\frac{1}{4}+3^2\times\frac{1}{4}+4^2\times\frac{1}{4}=\frac{1+4+9+16}{4}=\frac{30}{4}=\frac{15}{2}.$$
   Ainsi $V(X_1)=V(X_2)=\frac{15}{2}-\left(\frac{5}{2}\right)^2=\frac{15}{2}-\frac{25}{4}=\frac{30-25}{4}=\frac{5}{4}$. Et donc
   $$V(X_1+X_2)=\frac{5}{4}+\frac{5}{4}+2\times 0=\frac{10}{4}=\frac{5}{2}.$$
   De même, pour calculer $V(X_1+Y)$, il me faut calculer $V(Y)$.
   Pour cela, je commence par calculer $E\left(Y^2\right)$ :
   $$E\left(Y^2\right)=1^2\times\frac{1}{16}+2^2\times\frac{3}{16}+3^2\times\frac{5}{16}+4^2\times\frac{7}{16}=\frac{1+12+45+112}{16}=\frac{170}{16}=\frac{85}{8}.$$
   Alors d'après la formule de König-Huygens,
   $$V(Y)=E\left(Y^2\right)-E(Y)^2=\frac{85}{8}-\left(\frac{25}{8}\right)^2=\frac{85}{8}-\frac{625}{64}=\frac{680-625}{64}=\frac{55}{64}.$$
   Et donc
   $$V(X_1+Y)=\frac{5}{4}+\frac{55}{64}+2\times\frac{5}{8}=\frac{80+55+80}{64}=\frac{215}{64}.$$

2. Calculer $V(X+Y)$.

   Je raisonne de manière similaire. Par le théorème de transfert,
   $$E\left(X^2\right)=1^2\times\frac{1}{2}+2^2\times\frac{1}{2}=\frac{1+4}{2}=\frac{5}{2} \quad \text{et} \quad E\left(Y^2\right)=1^2\times\frac{1}{4}+2^2\times\frac{3}{4}=\frac{1+12}{4}=\frac{13}{4}.$$
   Alors d'après la formule de König-Huygens,
   $$V(X)=E\left(X^2\right)-E(X)^2=\frac{5}{2}-\left(\frac{3}{2}\right)^2=\frac{10-9}{4}=\frac{1}{4} \quad \text{et} \quad V(Y)=E\left(Y^2\right)-E(Y)^2=\frac{13}{4}-\left(\frac{7}{4}\right)^2=\frac{3}{16}.$$
   Et finalement
   $$V(X+Y)=V(X)+V(Y)+2\,\mathrm{Cov}(X,Y)=\frac{1}{4}+\frac{3}{16}+2\times\frac{1}{8}=\frac{4+3+4}{16}=\frac{11}{16}.$$

**Remarque 2.41 –** Cette formule permet également de calculer la covariance de $X$ et $Y$ à l'aide des trois variances $V(X)$, $V(Y)$ et $V(X+Y)$ puisque
$$\mathrm{Cov}(X,Y)=\frac{V(X+Y)-V(X)-V(Y)}{2}.$$

**Proposition 2.42**

Soit $(X, Y)$ un couple de variables aléatoires **indépendantes**. Alors

$$V(X+Y) = V(X) + V(Y).$$

## 3 – Coefficient de corrélation linéaire

**Définition 2.43 –** On appelle **coefficient de corrélation linéaire** de $X$ et $Y$, le réel, noté $\rho(X, Y)$, défini par

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}.$$

**Exemple 2.44 –** On reprend les deux exemples de l'exemple 2.26.

1. Calculer $\rho(X_1, X_2)$ et $\rho(X_1, Y)$.
   Toutes les valeurs impliquées ont déjà été calculées précédemment, je peux donc appliquer la formule directement :

$$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{V(X_1)}\sqrt{V(X_2)}} = \frac{0}{\sqrt{\frac{5}{4}} \times \sqrt{\frac{5}{4}}} = 0,$$

$$\rho(X_1, Y) = \frac{\text{Cov}(X_1, Y)}{\sqrt{V(X_1)}\sqrt{V(Y)}} = \frac{\frac{5}{8}}{\sqrt{\frac{5}{4}} \times \sqrt{\frac{55}{64}}} = \frac{\frac{5}{8}}{\frac{5\sqrt{11}}{16}} = \frac{5}{8} \times \frac{16}{5\sqrt{11}} = \frac{2}{\sqrt{11}}.$$

2. Calculer $\rho(X, Y)$.
   Toutes les valeurs impliquées ont déjà été calculées précédemment, je peux donc appliquer la formule directement :

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{\frac{1}{8}}{\sqrt{\frac{1}{4}} \times \sqrt{\frac{3}{16}}} = \frac{\frac{1}{8}}{\frac{\sqrt{3}}{8}} = \frac{1}{\sqrt{3}}.$$

**Proposition 2.45**

Soient $X$ et $Y$ deux variables aléatoires définies sur $\Omega$. Alors

$$\left|\rho(X, Y)\right| \leqslant 1.$$

**Remarque 2.46 –** Le coefficient de corrélation linéaire mesure la dépendance linéaire entre deux variables aléatoires :

- Si le coefficient de corrélation linéaire est égal à 1 ou $-1$, $X$ et $Y$ sont corrélées linéairement.
  Le signe indique si les variations vont dans le même sens ou dans le sens opposé.
- Si le coefficient de corrélation linéaire est égal à 0, $X$ et $Y$ sont dites "non corrélées linéairement".
  Cela ne dit rien en revanche quant à une autre forme de corrélation.