

## Chapitre 16

# Couples de variables aléatoires, indépendance

$X \backslash Y$	$y_1$	$y_2$	$y_3$	Loi de X
$x_1$				$p[X = x_1]$
$x_2$				$p[X = x_2]$
Loi de Y	$p[Y = y_1]$	$p[Y = y_2]$	$p[Y = y_3]$	1

# Table des matières

<b>I</b>	<b>Couples de variables aléatoires finies</b>	<b>1</b>
1)	Généralités . . . . .	1
2)	Loi conditionnelle . . . . .	8
3)	Généralisation : vecteurs aléatoires . . . . .	10
<b>II</b>	<b>Variables aléatoires indépendantes</b>	<b>11</b>
1)	Indépendance d'un couple de variables aléatoires . . . . .	11
2)	Espérance et variables indépendantes . . . . .	15
3)	Variables aléatoires mutuellement indépendantes . . . . .	18
4)	Sommes de lois de Bernoulli indépendantes . . . . .	22
<b>III</b>	<b>Covariance et coefficient de corrélation linéaire</b>	<b>24</b>
1)	Covariance de deux variables aléatoires . . . . .	24
2)	Coefficient de corrélation linéaire de deux variables aléatoires . . . . .	34

Dans ce chapitre, les variables aléatoires sont supposées **finies**, c'est-à-dire définies sur un **espace probabilisé fini**  $(\Omega, \mathbb{P})$ . L'espérance et de variance de telles variables aléatoires seront donc toujours bien définies.

## I Couples de variables aléatoires finies

### 1) Généralités

Un couple de variables aléatoires est également une variable aléatoire, au sens suivant :

**Définition 1 (Couple de variables aléatoires finies).**

*Soient  $X$  et  $Y$  deux variables aléatoires réelles définies sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ . On appelle **couple des variables aléatoires  $X$  et  $Y$**  la variable aléatoire  $Z = (X, Y) : \Omega \rightarrow \mathbb{R}^2$  définie par  $(X, Y)(\omega) = (X(\omega), Y(\omega))$ . La loi du couple  $(X, Y)$ , notée  $\mathbb{P}_{(X,Y)}$ , est appelée **loi conjointe** de  $X$  et  $Y$ .*

**ATTENTION !**

Un couple de variables aléatoires réelles est une variable aléatoire **vectorielle**.

**Proposition 2 (Détermination de la loi conjointe).**

*La loi d'un couple  $(X, Y) : \Omega \rightarrow \mathbb{R}^2$  est entièrement déterminée par les valeurs :*

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

*pour  $(x, y) \in X(\Omega) \times Y(\Omega)$ .*

**Preuve :** Cela résulte du fait qu'une probabilité sur un ensemble  $E$  est entièrement déterminée par ses valeurs sur les singletons de  $E$ . Ici, on l'applique à l'ensemble des valeurs de  $(X, Y)$ , c'est-à-dire l'ensemble des couples  $(x, y) \in X(\Omega) \times Y(\Omega)$  (on rappelle que la loi d'une variable aléatoire  $Z$  est une probabilité sur l'image  $Z(\Omega)$ ).

□

**Remarque**

Les événements élémentaires  $((X, Y) = (x, y)) = (X = x) \cap (Y = y)$ , où  $(x, y)$  décrit  $X(\Omega) \times Y(\Omega)$ , forment un système complet d'événements.

Il est tout à fait possible que certains de ces événements soient vides.

**Exemple (Deux lancers d'un dé)**

On lance deux fois un dé équilibré. On a alors  $\Omega = \{1, \dots, 6\}^2$ .

On peut définir la variable aléatoire  $X$  qui donne le résultat du premier lancer et  $Y$  donnant le résultat du deuxième lancer. La loi du couple  $(X, Y)$  est alors donnée par :

$$\forall (i, j) \in \{1, \dots, 6\}^2, \quad \mathbb{P}((X, Y) = (i, j)) = \frac{1}{36}.$$

La loi conjointe de  $X$  et  $Y$  est donc ici ..... la loi uniforme sur l'ensemble  $\{1, \dots, 6\}^2$  .....

**Exemple (Urne)**

Une urne contient 2 boules blanches, 1 rouge et 1 noire, toutes indiscernables.

On tire simultanément deux boules de cette urne et on nomme  $X$  la variable aléatoire donnant le nombre de boules blanches tirées et  $Y$  celle donnant le nombre boules rouges tirées.

1. Construire un espace probabilisé  $\Omega$  qui modélise l'expérience aléatoire.
2. Déterminer la loi conjointe de  $X$  et  $Y$ .

1. Ici, le tirage se fait simultanément donc on peut définir  $\Omega$  comme l'ensemble des parties à 2 éléments d'un ensemble à 4 éléments (puisque l'ordre ne compte pas).  
On a donc  $\# \Omega = \binom{4}{2} = 6$ .

2. On a  $X(\Omega) = \{0, 1, 2\}$  et  $Y(\Omega) = \{0, 1\}$ , donc il suffit de calculer  
 $\mathbb{P}((X, Y) = (a, b))$  pour tout  $(a, b) \in \{0, 1, 2\} \times \{0, 1\}$ .

$$\mathbb{P}((X, Y) = (0, 0)) = 0$$

$$\mathbb{P}((X, Y) = (0, 1)) = \frac{1}{6}$$

$$\mathbb{P}((X, Y) = (1, 0)) = \frac{2}{6}$$

$$\mathbb{P}((X, Y) = (1, 1)) = \frac{2}{6}$$

$$\mathbb{P}((X, Y) = (2, 0)) = \frac{1}{6}$$

$$\mathbb{P}((X, Y) = (2, 1)) = 0$$

**Définition 3 (Lois marginales).**

Étant donné un couple de variables aléatoires réelles  $(X, Y)$  sur un espace  
 probabilisé fini  $(\Omega, \mathbb{P})$ , on appelle **lois marginales** les lois de  $X$  et  $Y$ ,

La loi de  $X$  est appelée **première loi marginale de  $(X, Y)$**  et celle de  $Y$   
 est appelée **deuxième loi marginale de  $(X, Y)$** .

**Méthode (Déterminer les lois marginales à partir de la loi conjointe)**

Si on connaît la loi du couple  $(X, Y)$ , alors on peut retrouver celles de  $X$  et de  $Y$ .

Par exemple, pour déterminer la loi de  $X$  : on a, pour tout  $x \in X(\Omega)$ ,

$$\mathbb{P}(X = x) = \mathbb{P}\left(\bigcup_{y \in Y(\Omega)} ((X = x) \cap (Y = y))\right) = \sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y)) \quad ,$$

puisque  $(Y = y)_{y \in Y(\Omega)}$  est un système complet d'événements.

**On peut présenter cette idée sous forme d'un tableau.**

En notant  $X(\Omega) = \{x_1, \dots, x_m\}$  et  $Y(\Omega) = \{y_1, \dots, y_n\}$ , on forme un tableau à  $m$  lignes et  $n$  colonnes et dont les cases sont les probabilités  $p_{i,j} = \mathbb{P}((X = x_i) \cap (Y = y_j))$ . En sommant tous les éléments d'une même ligne, on obtient les probabilités  $\mathbb{P}(X = x_i)$  et, en sommant tous les éléments d'une même colonne, on obtient les  $\mathbb{P}(Y = y_j)$ .

**ATTENTION !****Les lois marginales ne suffisent pas en général à retrouver la loi conjointe !**

En effet si  $X$  peut prendre  $m$  valeurs différentes et  $Y$  peut en prendre  $n$ , alors les  $p_{i,j} = \mathbb{P}((X = x_i) \cap (Y = y_j))$  sont au nombre de  $m \times n$ .

Or la connaissance des lois marginales, donc de  $\mathbb{P}(X = x_i)$  et  $\mathbb{P}(Y = y_j)$  pour  $1 \leq i \leq m$  et  $1 \leq j \leq n$  ne donnent que  $m + n$  équations du type :

$$\sum_{i=1}^m p_{ij} = \mathbb{P}(Y = y_j), \quad \sum_{j=1}^n p_{ij} = \mathbb{P}(X = x_i) \quad .$$

On a donc seulement un système linéaire de  $m + n$  équations à  $m \times n$  inconnues, il ne possède donc pas une unique solution (puisque  $m \times n > m + n$  en général), ce qui empêche la détermination des  $p_{ij}$ , et donc de la loi conjointe.

**Exemple (Urne (suite))**

Reprenons l'exemple précédent :

Une urne contient 2 boules blanches, 1 rouge et 1 noire, toutes indiscernables. On tire simultanément deux boules de cette urne et on nomme  $X$  la variable aléatoire donnant le nombre de boules blanches tirées et  $Y$  celle donnant le nombre boules rouges tirées.

Calculons les lois marginales à partir de la loi du couple  $(X, Y)$  :

$X \backslash Y$	0	1	Loi de $X$
0	0	$\frac{1}{6}$	$\mathbb{P}(X = 0) = \frac{1}{6}$
1	$\frac{2}{6}$	$\frac{2}{6}$	$\mathbb{P}(X = 1) = \frac{4}{6}$
2	$\frac{1}{6}$	0	$\mathbb{P}(X = 2) = \frac{1}{6}$
Loi de $Y$	$\mathbb{P}(Y = 0) = \frac{1}{2}$	$\mathbb{P}(Y = 1) = \frac{1}{2}$	

**2) Loi conditionnelle****Définition 4 (Loi conditionnelle de  $Y$  sachant  $(X = x)$ ).**

Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un espace probabilisé fini  $(\Omega, \mathbb{P})$ .

Pour tout  $x \in X(\Omega)$  tel que  $\mathbb{P}(X = x) \neq 0$ , on appelle loi conditionnelle

de  $Y$  sachant  $(X = x)$  la probabilité  $\mathbb{P}_{Y|(X=x)} : \mathcal{P}(Y(\Omega)) \rightarrow [0; 1]$  définie par :

$$\forall B \subset Y(\Omega), \quad \mathbb{P}_{Y|(X=x)}(B) = \mathbb{P}_{(X=x)}(Y \in B) = \frac{\mathbb{P}((Y \in B) \cap (X = x))}{\mathbb{P}(X = x)}.$$

**Remarque**

- On vérifie aisément que  $\mathbb{P}_{Y|(X=x)}$  est une probabilité sur  $Y(\Omega)$ .
- On peut bien sûr définir de même  $\mathbb{P}_{X|(Y=y)}$  (la loi conditionnelle de  $X$  sachant  $(Y = y)$ ).
- Etant donnée une partie  $A \subset X(\Omega)$  telle que  $\mathbb{P}(X \in A) \neq 0$ , on peut également définir de manière générale la loi conditionnelle de  $Y$  sachant  $(X \in A)$  par :

$$\mathbb{P}_{(X \in A)}(Y \in B) = \frac{\mathbb{P}((Y \in B) \cap (X \in A))}{\mathbb{P}(X \in A)}.$$

**Exemple (Lancer de deux dés)**

On lance deux dés équilibrés,  $X$  est la variable résultat du premier dé et  $Y$  celle du second dé. On note  $S = X + Y$ . On a, par exemple,

$$\begin{aligned}\mathbb{P}_{S|(X=1)}(\{4\}) &= \mathbb{P}_{(X=1)}(S=4) = \frac{\mathbb{P}((X=1) \cap (S=4))}{\mathbb{P}(X=1)} = \frac{\mathbb{P}((X,Y)=(1,3))}{\mathbb{P}(X=1)} = \frac{1/36}{1/6} = \frac{1}{6}; \\ \mathbb{P}_{X|(S=4)}(\{1\}) &= \mathbb{P}_{(S=4)}(X=1) = \frac{\mathbb{P}((X=1) \cap (S=4))}{\mathbb{P}(S=4)} = \frac{\mathbb{P}((X,Y)=(1,3))}{\mathbb{P}(S=4)} = \frac{1/36}{3/36} = \frac{1}{3}.\end{aligned}$$

**Méthode (Calcul de la loi de  $Y$  à l'aide des lois conditionnelles)**

Grâce aux lois conditionnelles de  $Y$  sachant  $(X = x)$  (lorsque  $x$  décrit  $X(\Omega)$ ), on peut calculer la loi de  $Y$  à l'aide de la formule des probabilités totales :

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

puisque les événements  $(X = x)_{x \in X(\Omega)}$  forment un syst. complet d'événements de  $\Omega$ .

**3) Généralisation : vecteurs aléatoires****Définition 5 (Vecteur aléatoire).**

Soient  $X_1, \dots, X_n$  des variables aléatoires réelles sur l'espace probabilisé fini  $(\Omega, \mathbb{P})$ .

On appelle **vecteur aléatoire discret** défini à partir de  $X_1, \dots, X_n$  la variable aléatoire  $Z : \Omega \rightarrow \mathbb{R}^n$  définie par :  $\forall \omega \in \Omega, Z(\omega) = (X_1(\omega), \dots, X_n(\omega))$ .

La loi de  $Z$  est appelée **loi conjointe** des variables  $X_1, \dots, X_n$  tandis que les lois de  $X_1, \dots, X_n$  sont les **lois marginales** de  $Z$ .

**Remarque**

Bien sûr, comme dans le cas  $n = 2$ , la loi conjointe permet de déterminer la loi marginale (en sommant), mais les lois marginales ne déterminent pas en général la loi conjointe.

## II Variables aléatoires indépendantes

### 1) Indépendance d'un couple de variables aléatoires

**Définition 6 (Indépendance de deux variables aléatoires discrètes).**

Deux variables aléatoires  $X$  et  $Y$  sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$  sont dites **indépendantes** lorsque pour toutes parties  $A \subset X(\Omega)$  et  $B \subset Y(\Omega)$ ,  
 les événements  $(X \in A)$  et  $(Y \in B)$  sont indépendants, c'est-à-dire :  

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

On a comme résultat fondamental :

**Proposition 7 (Caractérisation de deux VA indépendantes).**

Soient  $X$  et  $Y$  deux variables aléatoires définies sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ . Alors :  $X$  et  $Y$  sont indépendantes si et seulement si

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

**Remarque**

$X, Y$  indépendantes  $\iff \forall (x, y) \in X(\Omega) \times Y(\Omega), (X = x)$  et  $(Y = y)$  indépendants.

**Preuve :  $\implies$**  : Evident en choisissant  $A = \{x\}$  et  $B = \{y\}$ .

**$\impliedby$**  : S'obtient en écrivant  $A$  et  $B$  comme réunions disjointes de leurs singletons :

$$\mathbb{P}((X \in A) \cap (Y \in B)) = \sum_{(x,y) \in A \times B} \mathbb{P}((X = x) \cap (Y = y)) = \sum_{(x,y) \in A \times B} \mathbb{P}(X = x) \mathbb{P}(Y = y).$$

On somme "par paquets" :

$$\mathbb{P}((X \in A) \cap (Y \in B)) = \sum_{x \in A} \left( \sum_{y \in B} \mathbb{P}(X = x) \mathbb{P}(Y = y) \right) = \sum_{x \in A} \mathbb{P}(X = x) \underbrace{\left( \sum_{y \in B} \mathbb{P}(Y = y) \right)}_{=\mathbb{P}(Y \in B)},$$

donc

$$\mathbb{P}((X \in A) \cap (Y \in B)) = \left( \sum_{x \in A} \mathbb{P}(X = x) \right) \times \mathbb{P}(Y \in B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B).$$

□



**Remarque**

Contrairement au cas général, on constate que pour un couple de VA **indépendantes**, les lois marginales déterminent la loi conjointe.

**Exemple (Deux lancers d'un dé)**

Comme on peut s'en douter, lorsqu'on lance deux fois un dé équilibré, avec  $\Omega = \{1, \dots, 6\}^2$ , les variables aléatoires  $X = \text{"résultat du premier lancer"}$  et  $Y = \text{"résultat du deuxième lancer"}$  sont indépendantes.

En effet,  $\forall (i, j) \in \{1, \dots, 6\}^2 : \mathbb{P}((X, Y) = (i, j)) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \mathbb{P}(X = i)\mathbb{P}(Y = j)$ .

**Proposition 8 (Indépendances de VA composées).**

Soient  $X, Y$  deux variables aléatoires réelles **indépendantes** sur un même univers fini  $(\Omega, \mathbb{P})$ . Pour toutes fonctions  $f : X(\Omega) \rightarrow \mathbb{R}$  et  $g : Y(\Omega) \rightarrow \mathbb{R}$ ,

les variables aléatoires  $f(X)$  et  $g(Y)$  sont indépendantes.

**Preuve :** Soient  $a \in f(X(\Omega))$  et  $b \in g(Y(\Omega))$ . Alors :

$$\begin{aligned} \mathbb{P}((f(X) = a) \cap (g(Y) = b)) &= \mathbb{P}((X \in f^{-1}(\{a\})) \cap (Y \in g^{-1}(\{b\}))) \\ &= \mathbb{P}(X \in f^{-1}(\{a\})) \times \mathbb{P}(Y \in g^{-1}(\{b\})), \end{aligned}$$

par indépendance de  $X$  et  $Y$ , d'où :

$$\mathbb{P}((f(X) = a) \cap (g(Y) = b)) = \mathbb{P}(f(X) = a) \times \mathbb{P}(g(Y) = b),$$

ce qui montre que  $f(X)$  et  $g(Y)$  sont indépendantes. □

**Exemple**

Si  $X$  et  $Y$  sont indépendantes, alors  $X^2$  et  $Y^3$  sont indépendantes.

## 2) Espérance et variables indépendantes

On rappelle que l'on peut définir le produit de deux variables aléatoires réelles par

$$\forall \omega \in \Omega, \quad (XY)(\omega) = X(\omega) \times Y(\omega).$$

### **Théorème 9 (Produit de variables aléatoires indépendantes).**

*Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un univers fini  $(\Omega, \mathbb{P})$ .*

**Si  $X$  et  $Y$  sont indépendantes, alors  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ .**

**Preuve :** Soit  $U = XY$ . On a  $U : \Omega \rightarrow \mathbb{R}$ . Par définition de l'espérance, on a

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

Par indépendance de  $X$  et  $Y$ , ceci se réécrit :

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

On peut alors grouper les termes de la somme selon les valeurs du produit  $u = xy$  :

$$\begin{aligned} \mathbb{E}(X)\mathbb{E}(Y) &= \sum_{u \in U(\Omega)} \left( \sum_{(x,y) \in X(\Omega) \times Y(\Omega), \, xy=u} \underbrace{xy}_{=u} \mathbb{P}((X, Y) = (x, y)) \right) \\ &= \sum_{u \in U(\Omega)} u \left( \sum_{(x,y) \in X(\Omega) \times Y(\Omega), \, xy=u} \mathbb{P}((X, Y) = (x, y)) \right) \\ &= \sum_{u \in U(\Omega)} u \mathbb{P}(U = u) = \mathbb{E}(U), \end{aligned}$$

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

**ATTENTION !**

La réciproque est fausse en général.

**Exemple**

Soit  $X \sim \mathcal{U}(\{-1, 0, 1\})$ , et  $Y = X^2$ .

Montrer que  $X$  et  $Y$  ne sont pas indépendantes mais que  $\mathbb{E}(XY) = 0$ .

On a  $\mathbb{E}(X) = 0$ .

On pose  $Y = X^2$ . Les variables  $X$  et  $Y$  ne sont **pas indépendantes**, puisque :  $\mathbb{P}(X = 0 \cap Y = 0) = \mathbb{P}(X = 0) = \frac{1}{3} \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0) = \mathbb{P}(X = 0)^2 = \frac{1}{9}$ .

Pourtant  $\mathbb{E}(XY) = \mathbb{E}(X^3) = \mathbb{E}(X) = 0 = \mathbb{E}(X)\mathbb{E}(Y)$ .

**Corollaire 10.**

Soient  $X, Y$  deux variables aléatoires réelles **indépendantes** sur un même univers fini  $(\Omega, \mathbb{P})$ . Pour toutes fonctions  $f : X(\Omega) \rightarrow \mathbb{R}$  et  $g : Y(\Omega) \rightarrow \mathbb{R}$ , on a  $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$ .

**Preuve :** Vu que  $X$  et  $Y$  sont indépendantes,  $f(X)$  et  $g(Y)$  aussi (d'après la prop. 8), donc l'espérance de  $f(X)g(Y)$  vaut le produit  $\mathbb{E}(f(X)) \times \mathbb{E}(g(Y))$ .  $\square$

**3) Variables aléatoires mutuellement indépendantes**

La notion d'indépendance se généralise à  $n$  variables aléatoires :

**Définition 11 (Variables aléatoires mutuellement indépendantes).**

On dit que des variables aléatoires réelles  $X_1, \dots, X_n$  sur un même univers fini  $(\Omega, \mathbb{P})$  sont **mutuellement indépendantes** lorsque

pour toutes parties  $A_1 \subset X_1(\Omega), \dots, A_n \subset X_n(\Omega)$  :

$$\mathbb{P}((X_1 \in A_1) \cap \dots \cap (X_n \in A_n)) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i).$$

On dispose de différentes propriétés, que nous admettrons :

**Proposition 12 (Caractérisation de l'indépendance mutuelle de VA).**

*Des variables aléatoires  $X_1, \dots, X_n$  sur un même univers fini  $(\Omega, \mathbb{P})$*

*sont mutuellement indépendantes ssi  $\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega),$*

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

**Remarque**

Lorsqu'on répète une expérience et qu'on suppose que le résultat d'une expérience est sans effet sur les autres, cela signifie que les variables aléatoires  $X_1, \dots, X_n$  associées respectivement aux expériences numéro  $1, \dots, n$  sont mutuellement indépendantes.

C'est notamment le cas lorsqu'on lance plusieurs fois une même pièce de monnaie, lorsqu'on tire au hasard, **avec remise**, des boules dans une même urne ...

**Proposition 13.**

*Si les variables  $X_1, \dots, X_n$  sont mutuellement indépendantes, alors*

*pour tout  $m \in \{1, \dots, n-1\}$  et toutes fonctions  $f : (X_1, \dots, X_m)(\Omega) \rightarrow \mathbb{R}$*

*et  $g : (X_{m+1}, \dots, X_n)(\Omega) \rightarrow \mathbb{R}$ , les variables  $f(X_1, \dots, X_m)$  et  $g(X_{m+1}, \dots, X_n)$*

*sont indépendantes .*

**Exemple**

1. Si on a 3 variables indépendantes  $X, Y$  et  $Z$ , alors on en déduit notamment que  $Z$  est indépendante de  $X + Y$ , ou encore que  $Z^2$  est indépendante de  $XY$  ...
2. Si on répète une expérience décrite par une suite de variables aléatoires  $(X_n)_{n \geq 1}$  et que l'on suppose chaque réalisation d'une expérience indépendante des autres, alors on a en particulier, pour tout  $n \geq 1$ ,  $X_{n+1}$  qui est indépendante de  $X_1 + \dots + X_n$  ainsi que de  $X_1 \times \dots \times X_n$ .

On a enfin le résultat suivant, concernant l'espérance d'un produit de variables mutuellement indépendantes :

**Proposition 14 (Espérance d'un produit de VA mutuellement indép.).**

Soient  $X_1, \dots, X_n$  des variables aléatoires réelles sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ . Si  $X_1, \dots, X_n$  sont **mutuellement indépendantes**, alors

$$\mathbb{E}(X_1 \times \dots \times X_n) = \mathbb{E}(X_1) \times \dots \times \mathbb{E}(X_n) .$$

**Preuve :** Par récurrence sur  $n$  : c'est vrai pour  $n = 2$  (cf. th 9), et si  $X_1, \dots, X_{n+1}$  sont mutuellement indépendantes, alors  $X_{n+1}$  est indépendante de  $Y = X_1 \times \dots \times X_n$  d'après ce qui précède, donc  $\mathbb{E}(X_1 \times \dots \times X_n \times X_{n+1}) = \mathbb{E}(X_1 \times \dots \times X_n) \times \mathbb{E}(X_{n+1})$ , puis on utilise l'hypothèse de récurrence sur  $X_1, \dots, X_n$  qui sont mutuellement indépendantes.  $\square$

#### 4) Sommes de lois de Bernoulli indépendantes

On a le résultat important suivant :

**Théorème 15 (Somme de lois de Bernoulli indépendantes).**

Soit  $p \in [0, 1]$ . Si  $X_1, \dots, X_n$  sont des variables aléatoires

**mutuellement indépendantes** sur un même univers fini  $(\Omega, \mathbb{P})$ , et suivant toutes la loi de Bernoulli  $\mathcal{B}(p)$ , alors  $X_1 + \dots + X_n$  suit la loi binomiale  $\mathcal{B}(n, p)$ .

**Preuve :** Les  $X_i$  sont à valeurs dans  $\{0, 1\}$  donc la variable aléatoire  $S = X_1 + \dots + X_n$  est à valeurs dans  $\{0, \dots, n\}$ . Ensuite, pour tout  $k \in \{0, \dots, n\}$ , on a

$$(S = k) = \bigcup_{(x_1, \dots, x_n) \in A_k} ((X_1, \dots, X_n) = (x_1, \dots, x_n)),$$

(la réunion étant disjointe) où  $A_k$  est l'ensemble des listes binaires  $(x_1, \dots, x_n)$  qui comportent exactement  $k$  "1" et  $n - k$  "0". Il y a exactement  $\binom{n}{k}$  suites de ce type, donc  $\#A_k = \binom{n}{k}$ . Par additivité de la probabilité  $\mathbb{P}$ , on en déduit que

$$\mathbb{P}(S = k) = \sum_{(x_1, \dots, x_n) \in A_k} \mathbb{P}((X_1, \dots, X_n) = (x_1, \dots, x_n)).$$

L'indépendance mutuelle de  $X_1, \dots, X_n$  entraîne alors :

$$\mathbb{P}(S = k) = \sum_{(x_1, \dots, x_n) \in A_k} \mathbb{P}(X_1 = x_1) \times \dots \times \mathbb{P}(X_n = x_n).$$

Or, on a  $\mathbb{P}(X_i = 1) = p$  et  $\mathbb{P}(X_i = 0) = 1 - p$  pour tout  $i \in \{1, \dots, n\}$ , donc

$$(x_1, \dots, x_n) \in A_k \implies \mathbb{P}(X_1 = x_1) \times \dots \times \mathbb{P}(X_n = x_n) = p^k (1 - p)^{n-k}$$

(puisque'il y a exactement  $k$  "1" dans la liste  $(x_1, \dots, x_n)$ ). Donc :

$$\mathbb{P}(S = k) = \sum_{(x_1, \dots, x_n) \in A_k} p^k (1 - p)^{n-k} = \#A_k \times p^k (1 - p)^{n-k} = \binom{n}{k} p^k (1 - p)^{n-k},$$

ce qui montre que  $S \sim \mathcal{B}(n, p)$ .

□

### III Covariance et coefficient de corrélation linéaire

On rappelle que si deux variables  $X$  et  $Y$  sont indépendantes, alors  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ . Mais qu'en est-il dans le cas général (si  $X$  et  $Y$  ne sont pas supposées indépendantes) ?

#### 1) Covariance de deux variables aléatoires

##### Définition 16 (Covariance).

Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ . On appelle **covariance** de  $X$  et  $Y$  le nombre réel

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

Si on a  $Cov(X, Y) = 0$ , alors on dit que  $X$  et  $Y$  sont **non corrélées**.

##### Remarque

- On a  $Cov(X, X) = \mathbb{E}((X - \mathbb{E}(X))^2) = V(X)$ .
- La covariance est faite pour être un indicateur de corrélation entre deux variables aléatoires.

**Proposition 17 (Autre écriture de la covariance).**

Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ . Alors, on a  $Cov(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ .

**Preuve :**  $Cov(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) = \mathbb{E}(XY - \mathbb{E}(X)Y - \mathbb{E}(Y)X + \mathbb{E}(X)\mathbb{E}(Y)) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$

□

**Remarque**

Ainsi,  $Cov(X, Y) = 0$  ssi  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ .

**Proposition 18 (Covariance = forme bilinéaire symétrique positive).**

Fixons un espace probabilisé fini  $(\Omega, \mathbb{P})$ , et notons  $E$  l'espace vectoriel des variables

aléatoires  $\Omega \rightarrow \mathbb{R}$ . La covariance  $Cov : \begin{matrix} E \times E & \longrightarrow & \mathbb{R} \\ (X, Y) & \longmapsto & Cov(X, Y) \end{matrix}$  est une

forme bilinéaire symétrique sur  $E$  qui vérifie la propriété de "positivité" :

$Cov(X, X) \geq 0$  pour tout  $X \in E$ .

**Preuve :**

- Symétrie : on a  $XY = YX$ , donc

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

- Linéarité à gauche : on utilise la linéarité de l'espérance :

$$\begin{aligned} Cov(\lambda X_1 + X_2, Y) &= \mathbb{E}((\lambda X_1 + X_2)Y) - \mathbb{E}(\lambda X_1 + X_2)\mathbb{E}(Y) \\ &= \lambda \mathbb{E}(X_1 Y) + \mathbb{E}(X_2 Y) - (\lambda \mathbb{E}(X_1) + \mathbb{E}(X_2)) \mathbb{E}(Y) \\ &= \lambda (\mathbb{E}(X_1 Y) - \mathbb{E}(X_1)\mathbb{E}(Y)) + (\mathbb{E}(X_2 Y) - \mathbb{E}(X_2)\mathbb{E}(Y)) \\ &= \lambda Cov(X_1, Y) + Cov(X_2, Y) \end{aligned}$$

- Positivité :

$$\text{Cov}(X, X) = V(X) \geq 0.$$

□

### ATTENTION !

La covariance n'est pas "définie positive", car

$$\text{Cov}(X, X) = 0 \iff V(X) = 0 \iff X = \mathbb{E}(X) \text{ presque sûrement}$$

(cela n'implique pas  $X = 0$ ). **La covariance n'est donc pas un produit scalaire.**

#### Proposition 19 (Propriétés de la covariance).

Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un univers fini  $(\Omega, \mathbb{P})$ .

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

On généralise :  $V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$  .

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

(iii) Si  $X$  et  $Y$  sont indépendantes, alors  $V(X + Y) = V(X) + V(Y)$  .

On généralise : Si  $X_1, \dots, X_n$  sont deux à deux indépendantes, alors

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) .$$

(iv) On a l'inégalité de Cauchy-Schwarz :  $|\text{Cov}(X, Y)| \leq \sigma(X)\sigma(Y)$  .

(où  $\sigma(X) = \sqrt{V(X)}$  désigne l'écart-type de  $X$ ).

**Preuve :**

(i)



(ii)

(iii)

- (iv) • Si  $V(X) \neq 0$ , examiner

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

C'est un trinôme du second degré en  $\lambda$  (car  $V(X) \neq 0$ ), dont le discriminant est  $4(\text{Cov}(X, Y)^2 - V(X)V(Y))$ .

Puisque  $T(\lambda) \geq 0$  pour tout  $\lambda \in \mathbb{R}$  (c'est une variance), on en déduit que le discriminant est négatif ou nul, ce qui donne l'inégalité voulue.

- Si  $V(X) = 0$ , alors  $X = \mathbb{E}(X)$  presque sûrement, et donc  $\text{Cov}(X, Y) = 0 \leq \sigma(X)\sigma(Y)$ .

□

### Remarque

On a donc, pour tous réels  $a$  et  $b$  et toutes variables aléatoires  $X, Y$  :

$$V(aX + bY) = a^2 V(X) + 2ab \text{Cov}(X, Y) + b^2 V(Y).$$

.....

### ATTENTION !

On a  $(X, Y \text{ indépendantes}) \implies (\text{Cov}(X, Y) = 0)$  mais **la réciproque est fausse !**

Contre-exemple : soit  $X \sim \mathcal{U}(\{-1, 0, 1\})$  et définissons la variable aléatoire  $Y$  prenant la valeur 0 quand  $X \neq 0$  et prenant la valeur 1 quand  $X = 0$ .

$X$  et  $Y$  ne sont pas indépendantes, car

$$\mathbb{P}((X = 0) \cap (Y = 0)) = 0 \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 0) = \frac{1}{3} \times \frac{2}{3}.$$

Mais on a  $\mathbb{E}(X) = 0$  et  $\mathbb{E}(XY) = \frac{1}{3}(-1 \times 0) + \frac{1}{3}(0 \times 1) + \frac{1}{3}(1 \times 0) = 0$ , d'où  $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$ .

**Remarque (Variance de la loi binomiale)**

On sait que si les  $X_i$  sont des variables aléatoires mutuellement indépendantes suivant toutes la loi  $\mathcal{B}(p)$ , alors la variable aléatoire  $X = \sum_{i=1}^n X_i$  suit la loi binomiale  $\mathcal{B}(n, p)$ .

On retrouve alors, par ce qui précède (vu que l'indépendance mutuelle implique l'indépendance deux à deux),  $V(X) = \sum_{i=1}^n V(X_i) = nV(X_1) = np(1-p)$ .

Au passage, on peut également retrouver l'espérance de la loi binomiale par cette méthode : par linéarité

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = p + \cdots + p = np,$$

mais ce calcul n'utilise pas l'indépendance des  $(X_i)_{1 \leq i \leq n}$ .

## 2) Coefficient de corrélation linéaire de deux variables aléatoires

**Définition 20 (Coefficient de corrélation linéaire).**

Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ , et de variance non nulle. On appelle **coefficient de corrélation linéaire**

du couple  $(X, Y)$  le nombre  $\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma(X)\sigma(Y)} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$ .

**Remarque**

- D'après l'inégalité de Cauchy-Schwarz, on a  $|\rho(X, Y)| \leq 1$ , c'est-à-dire  $-1 \leq \rho(X, Y) \leq 1$ .
- Analogie avec la formule du cosinus de l'angle non orienté entre deux vecteurs non nuls du plan euclidien  $\mathbb{R}^2$  :  $\cos(\vec{u}, \vec{v}) = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \|\vec{v}\|}$ .
- Si  $X$  et  $Y$  sont indépendantes, alors  $\rho(X, Y) = 0$  (réciproque fausse).

**Proposition 21 (Cas de corrélation maximale).**

Soient  $X$  et  $Y$  deux variables aléatoires réelles sur un même espace probabilisé fini  $(\Omega, \mathbb{P})$ , et de variance non nulle. On a :

$$\rho(X, Y) = \pm 1 \iff \text{il existe } a, b \in \mathbb{R} \text{ tels que } Y = aX + b \text{ presque sûrement}$$

(c'est-à-dire  $\mathbb{P}(Y = aX + b) = 1$ ).

**Preuve :** On raisonne encore avec le trinôme du second degré

$$T(\lambda) = V(\lambda X + Y)$$

de discriminant  $\Delta = 4(\text{Cov}(X, Y)^2 - V(X)V(Y))$  (on peut car  $V(X) \neq 0$ ). On a

$$\rho(X, Y) = \pm 1 \iff |\text{Cov}(X, Y)| = \sqrt{V(X)V(Y)} \iff \Delta = 0.$$

Cela revient donc à dire que  $T$  possède une unique racine réelle :

$$\exists! \lambda_0 \in \mathbb{R}, \quad T(\lambda_0) = 0,$$

c'est-à-dire  $V(\lambda_0 X + Y) = 0$ . Cela signifie que la variable aléatoire  $\lambda_0 X + Y$  est presque sûrement constante. En posant  $a = -\lambda_0$ , on obtient donc qu'il existe  $b \in \mathbb{R}$  tel que  $Y - aX = b$  presque sûrement.  $\square$