

# COUPLES ET VECTEURS DE VARIABLES ALÉATOIRES

« Ainsi, non autrement tout le monde travaille pour la petite espérance. »  
Charles Péguy

**Notations :** Dans ce chapitre,  $\Omega$  représente un ensemble fini, et  $\mathbf{P}$  une probabilité sur  $\Omega$ .

## 1 COUPLES DE VARIABLES ALÉATOIRES

### A Lois conjointes et marginales

**Définition 1.1** (*Loi conjointe*)

Soient  $X : \Omega \rightarrow E$  et  $Y : \Omega \rightarrow E'$  deux variables aléatoires.

$Z = (X, Y)$  est une nouvelle variable aléatoire de  $\Omega$  dans  $E \times E'$ .

La loi  $\mathbf{P}_Z$  du couple  $Z = (X, Y)$  est appelée **loi conjointe** de  $X$  et  $Y$ .

On définit alors

$$\mathbf{P}((X, Y) = (x, y)) = \mathbf{P}(X = x \text{ et } Y = y).$$

Cela revient à chercher la réalisation simultanée (conjointe) de  $X = x$  et de  $Y = y$ .

#### Exemple

On lance deux dés. On note  $X$  la valeur donnée par le premier dé et  $Y$  la valeur donnée par le second.

Si les dés sont équilibrés alors, pour tout  $(x, y) \in \llbracket 1, 6 \rrbracket^2$ ,

$$\mathbf{P}((X, Y) = (x, y)) = \frac{1}{36}.$$

#### Exemple

On tire successivement et sans remise deux boules dans une urne qui contient 5 boules rouges et 3 bleues. On note  $X_1, X_2$  les variables aléatoires qui valent 1 si la boule est rouge, et 0 sinon, respectivement au premier et au deuxième tirage.

On peut établir le tableau des probabilités conjointes :

		$X_2$	
		0	1
$X_1$	0	$\frac{6}{56}$	$\frac{15}{56}$
	1	$\frac{15}{56}$	$\frac{20}{56}$

On vérifie que la somme des probabilités est égale à 1.

**Exemple**

On lance deux dés équilibrés et on note  $X$  le minimum et  $Y$  le maximum.

		<b>Y</b>					
		1	2	3	4	5	6
<b>X</b>	1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	2		$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	3			$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	4				$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$
	5					$\frac{1}{36}$	$\frac{2}{36}$
	6						$\frac{1}{36}$

On vérifie que la somme des probabilités est égale à 1.

Dans cet exemple, on voit que l'univers image  $(X, Y)(\Omega)$  n'est pas à confondre avec l'ensemble  $X(\Omega) \times Y(\Omega)$  dont il n'est qu'une partie.

Ici, l'univers image est  $(X, Y)(\Omega) = \{(x, y) \in \llbracket 1, 6 \rrbracket^2, x \leq y\}$  (triangle supérieur dans le tableau).

Il est différent de l'univers  $X(\Omega) \times Y(\Omega) = \{(x, y) \in \llbracket 1, 6 \rrbracket^2\}$  (tableau entier).

Comme  $X(\Omega) \times Y(\Omega)$  est plus simple à calculer, on travaillera souvent sur cet univers par abus, quitte à donner la probabilité 0 aux événements « *en trop* ». Mais il faut garder en tête cette distinction.

**Définition 1.2 (Loi marginale)**

Soient  $X : \Omega \rightarrow E$  et  $Y : \Omega \rightarrow E'$  deux variables aléatoires.

La loi  $\mathbf{P}_X$  de la variable  $X$  est la première **loi marginale** de  $(X, Y)$ , et  $\mathbf{P}_Y$  est la seconde loi marginale.

**Explications**

La loi marginale suivant  $X$  revient à étudier la probabilité de réalisation de  $X$  quelle que soit la valeur prise par  $Y$ .

Dans le cas des probabilités discrètes on peut écrire les lois sous la forme de tableaux à deux entrées :

		<b>Y</b>						
		$y_1$	$y_2$	$\dots$	$y_j$	$\dots$	$y_q$	
<b>X</b>	$x_1$	$p_{1,1}$	$p_{1,2}$	$\dots$	$p_{1,j}$	$\dots$	$p_{1,q}$	$\mathbf{P}_X(x_1)$
	$x_2$	$p_{2,1}$	$p_{2,2}$	$\dots$	$p_{2,j}$	$\dots$	$p_{2,q}$	$\mathbf{P}_X(x_2)$
	$\vdots$				$\vdots$			$\vdots$
	$x_i$	$p_{i,1}$	$p_{i,2}$	$\dots$	$p_{i,j}$	$\dots$	$p_{i,q}$	$\mathbf{P}_X(x_i)$
	$\vdots$				$\vdots$			$\vdots$
	$x_p$	$p_{p,1}$	$p_{p,2}$	$\dots$	$p_{p,j}$	$\dots$	$p_{p,q}$	$\mathbf{P}_X(x_p)$
	$\Sigma$	$\mathbf{P}_Y(y_1)$	$\mathbf{P}_Y(y_2)$	$\dots$	$\mathbf{P}_Y(y_j)$	$\dots$	$\mathbf{P}_Y(y_q)$	$\rightarrow 1$

$\downarrow$   
 1

Pour obtenir  $p_i = \mathbf{P}(X = x_i) = \mathbf{P}_X(x_i)$ , on fait la somme de la ligne  $x_i$ .

Pour connaître  $p'_j = \mathbf{P}(Y = y_j) = \mathbf{P}_Y(y_j)$ , on somme suivant la colonne  $y_j$ .

Cela se résume par la formule qui suit.

**Propriété 1.3**

La loi conjointe détermine les lois marginales :

$$\mathbf{P}_X(x) = \sum_{y \in Y(\Omega)} \mathbf{P}(X = x \text{ et } Y = y).$$

$$\mathbf{P}_Y(y) = \sum_{x \in X(\Omega)} \mathbf{P}(X = x \text{ et } Y = y).$$

**Preuve**

C'est la formule des probabilités totales, simplement parce que

$$\begin{aligned} \mathbf{P}_X(x) &= \mathbf{P}\left([X = x] \cap \left(\bigcup_{y \in Y(\Omega)} [Y = y]\right)\right) \\ &= \mathbf{P}\left(\bigcup_{y \in Y(\Omega)} ([X = x] \cap [Y = y])\right) \\ &= \sum_{y \in Y(\Omega)} \mathbf{P}([X = x] \cap [Y = y]) \quad \text{événements incompatibles.} \end{aligned}$$

*Remarque :*  $\left\{ [X = x \text{ et } Y = y], (x, y) \in (X, Y)(\Omega) \right\}$  forme une famille complète d'événements. ■

**Exemple**

Pour le lancer de deux dés équilibrés :  $X$  traduit le résultat du premier lancer, et  $Y$  le résultat du second.

$$\mathbf{P}_X(4) = \mathbf{P}(X = 4) = \sum_{k=1}^6 \mathbf{P}(X = 4 \cap Y = k) = 6 \times \frac{1}{36} = \frac{1}{6}.$$

**Exemple**

Pour l'exemple du lancer de deux dés avec  $X$  le minimum et  $Y$  le maximum.

		Y						Σ
		1	2	3	4	5	6	
X	1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{11}{36}$
	2		$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{9}{36}$
	3			$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{7}{36}$
	4				$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{5}{36}$
	5					$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
	6						$\frac{1}{36}$	$\frac{1}{36}$
Σ		$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	→ 1

↓  
1

⚠ En général, les lois marginales ne permettent pas de définir la loi conjointe : il ne suffit pas de connaître la dernière ligne et la dernière colonne du tableau pour pouvoir le remplir tout entier !

**Exemple**

On tire successivement deux boules dans une urne avec  $R$  rouges et  $N - R$  noires.  $X$  vaut 1 si la première boule tirée est rouge,  $Y$  vaut 1 si la seconde boule tirée est rouge, 0 sinon.

Donner la loi de  $Y$ .

**Solution :**

$$\begin{aligned} \mathbf{P}_Y(1) &= \mathbf{P}(Y = 1) = \mathbf{P}(X = 1 \cap Y = 1) + \mathbf{P}(X = 0 \cap Y = 1) \\ &= \mathbf{P}_{X=1}(Y = 1)\mathbf{P}(X = 1) + \mathbf{P}_{X=0}(Y = 1)\mathbf{P}(X = 0) \\ &= \frac{R}{N} \times \frac{R-1}{N-1} + \frac{N-R}{N} \times \frac{R}{N-1} = \frac{R(R-1) + R(N-R)}{N(N-1)} \\ &= \frac{R}{N}. \end{aligned}$$

**Propriété 1.4** (*Loi image d'un couple*)

Soient  $X : \Omega \rightarrow E$  et  $Y : \Omega \rightarrow E'$  deux variables aléatoires.  
Soit  $u : (X, Y)(\Omega) \rightarrow F$ . La loi  $\mathbf{P}_u$  est définie par

$$\mathbf{P}_u(u(X, Y) = t) = \sum_{\substack{(x, y) \in (X, Y)(\Omega) \\ u(x, y) = t}} \mathbf{P}(X = x \text{ et } Y = y).$$

**Exemple** (*À retenir*)

Soit le tirage de deux dés.  $X$  est la variable aléatoire correspondant au résultat du premier dé, et  $Y$  au résultat du second.

La loi de la somme est donnée pour  $u(X, Y) = X + Y$ .

$$\mathbf{P}(X + Y = k) = \sum_{i=1}^{k-1} \mathbf{P}([X = i] \cap [Y = k - i]).$$

**B Espérance****Théorème 1.5** (*Théorème de transfert*)

Soient  $X$  et  $Y$  deux variables aléatoires finies sur  $\Omega$ , et  $u$  une application définie sur  $(X, Y)(\Omega)$  dans  $\mathbf{R}$ .

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

**Preuve**

Admis ■

**Exemple**

Calculer l'espérance de la somme de deux variables aléatoires  $X$  et  $Y$  en utilisant le théorème de transfert.

**Solution :**

$$\begin{aligned} \mathbf{E}(X + Y) &= \sum_{(x, y) \in (X, Y)(\Omega)} (x + y) \mathbf{P}(X = x, Y = y) \\ &= \sum_{(x, y) \in (X, Y)(\Omega)} x \mathbf{P}(X = x, Y = y) + \sum_{(x, y) \in (X, Y)(\Omega)} y \mathbf{P}(X = x, Y = y) \\ &= \sum_{x \in X(\Omega)} x \mathbf{P}(X = x) + \sum_{y \in Y(\Omega)} y \mathbf{P}(Y = y) = \mathbf{E}(X) + \mathbf{E}(Y) \end{aligned}$$

On retrouve la formule de la linéarité de l'espérance.

**C Covariance****Définition 1.6** (*Covariance*)

Soient  $X$  et  $Y$  deux variables aléatoires finies sur  $\Omega$ , on définit la covariance de  $X, Y$  par

$$\mathbf{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))].$$

Si on prend  $Y = X$ , on retrouve la formule de la variance de  $X$  :  $\mathbf{Cov}(X, X) = \mathbf{V}(X)$ . Comme pour la variance, on utilisera plutôt la formule de Kœnig-Huygens pour calculer la covariance.

**Propriété 1.7** (*Formule de Kœnig-Huygens*)

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

**Propriété 1.8**

Pour  $X, Y, X_1$  et  $X_2$  des variables aléatoires sur  $\Omega$ , et  $(a, b) \in \mathbf{R}^2$ ,

1.  $\mathbf{Cov}(X, Y) = \mathbf{Cov}(Y, X)$ . (symétrie)
2.  $\mathbf{Cov}(aX + b, Y) = a \mathbf{Cov}(X, Y)$ .
3.  $\mathbf{Cov}(aX_1 + X_2, Y) = a \mathbf{Cov}(X_1, Y) + \mathbf{Cov}(X_2, Y)$ . (linéarité)
4.  $\mathbf{Cov}(X, X) = \mathbf{V}(X) \geq 0$ . (positive)

**Propriété 1.9** (*Variance d'une somme*)

Soient  $X$  et  $Y$  deux variables aléatoires finies sur  $\Omega$ ,

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

**Preuve**

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

■

## D Coefficient de corrélation

### Définition 1.10 (Coefficient de corrélation)

Soient  $X$  et  $Y$  deux variables aléatoires, *non certaines*, sur  $\Omega$ .  
On définit le **coefficient de corrélation** du couple par

$$\rho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$

On peut aussi le noter  $\mathbf{Cor}(X, Y)$  ou  $r(X, Y)$ .

*Remarque* : on exige que les variables aléatoires soient *non certaines* pour que leur écart type soit non nul.

### Théorème 1.11 (Variables aléatoires affinement liées)

Soient  $X$  et  $Y$  deux variables aléatoires, *non certaines*, sur  $\Omega$ .

$$-1 \leq \rho(X, Y) \leq 1.$$

$\rho(X, Y) = -1$  ou  $\rho(X, Y) = 1$  si et seulement si les variables  $X$  et  $Y$  sont presque sûrement affinement liées, c'est-à-dire, s'il existe  $(a, b) \in \mathbf{R}^2$  tel que  $\mathbf{P}(X = aY + b) = 1$ .

### Explications

Dire que les variables sont *presque sûrement* affinement liées, veut dire que l'on peut écrire  $X = aY + b$  avec une quasi-certitude : si, pour certains  $\omega \in \Omega$ , l'égalité n'est pas vérifiée, alors, l'ensemble de ces  $\omega$  est inclus dans un événement de probabilité nulle. La probabilité de ne pas avoir l'égalité est donc nulle, ce qui n'est pas exactement la même chose que de dire que c'est impossible.

⚠ Ce n'est pas parce que le coefficient de corrélation est nul que les variables sont indépendantes (elles sont simplement décorréliées).

### Preuve (à retenir)

- Soit  $a \in \mathbf{R}$ ,  $\mathbf{V}(aX + Y) = a^2\mathbf{V}(X) + \mathbf{V}(Y) + 2a\mathbf{Cov}(X, Y)$ .  
Donc par positivité de la variance,  $\forall a \in \mathbf{R}$ ,  $a^2\mathbf{V}(X) + 2a\mathbf{Cov}(X, Y) + \mathbf{V}(Y) \geq 0$ .  
On reconnaît un trinôme du second degré en  $a$  qui reste donc positif sur  $\mathbf{R}$ .  
Son discriminant est négatif ou nul :  $\Delta = 4(\mathbf{Cov}(X, Y))^2 - 4\mathbf{V}(X)\mathbf{V}(Y) \leq 0$ .  
Donc  $(\mathbf{Cov}(X, Y))^2 \leq \mathbf{V}(X)\mathbf{V}(Y)$ , et, par croissance de la racine carrée sur  $\mathbf{R}_+$ ,  $|\mathbf{Cov}(X, Y)| \leq \sigma_X \sigma_Y$  : c'est l'**inégalité de Cauchy-Schwarz**.  
On obtient donc le résultat voulu,  $|\rho(X, Y)| \leq 1$ .
- $\rho(X, Y) \in \{-1, 1\}$  si et seulement si le cas d'égalité est vérifiée pour Cauchy-Schwarz. Cela correspond à un discriminant nul, c'est-à-dire une situation où l'expression polynomiale en  $a$  admet une racine.  
Il existe donc  $a \in \mathbf{R}$  tel que  $\mathbf{V}(aX + Y) = 0$  : la variable aléatoire  $aX + Y$  est donc presque certaine.  
Donc il existe  $b \in \mathbf{R}$  tel que  $aX + Y$  est presque sûrement égal à  $b$  :  $Y = -aX + b$ .  
Quitte à échanger les rôles de  $X$  et  $Y$  (rôles symétriques) et à changer  $a$  en  $-a$ , on trouve l'expression donnée dans le théorème. ■

### Propriété 1.12 (Invariance du coefficient de corrélation)

Le coefficient de corrélation est invariant par translations et changements d'échelle (strictement positifs).

$\forall (a, b) \in \mathbf{R}^2$ ,  $\forall (c, d) \in \mathbf{R}^2$ , avec  $a > 0$  et  $c > 0$

$$\rho(aX + b, cY + d) = \rho(X, Y).$$

### Explications

Un changement d'échelle positif correspond simplement à un changement d'unité entre deux unités proportionnelles (degrés et radians par exemple). On peut aussi y joindre une translation comme par exemple entre degrés Celsius, degrés Fahrenheit et Kelvin. On comprend bien que le changement d'unité ne doit pas modifier la corrélation entre les grandeurs physiques.

Par contre, si on multiplie une unité par un rapport strictement négatif, cela change le signe du coefficient de corrélation : comme nous le voyons dans la propriété qui suit, cela modifie le rapport de *croissance* ou *décroissance* entre  $X$  et  $Y$ .

### Preuve

$$\mathbf{Cov}(aX + b, cY + d) = a\mathbf{Cov}(X, cY + d) = ac\mathbf{Cov}(X, Y).$$

$$\sigma_{aX+b} = |a|\sigma_X = a\sigma_X \text{ car } a > 0. \text{ De même, } \sigma_{cY+d} = c\sigma_Y.$$

$$\text{Donc par quotient } \rho(aX + b, cY + d) = \rho(X, Y). \quad \blacksquare$$

### Propriété 1.13

Soient  $X$  et  $Y$  deux variables aléatoires *non certaines*, sur  $\Omega$ .

On note les **variables aléatoires centrées réduites**

$$X^* = \frac{X - \mathbf{E}(X)}{\sigma_X} \quad \text{et} \quad Y^* = \frac{Y - \mathbf{E}(Y)}{\sigma_Y}.$$

Si  $\rho(X, Y) = 1$ , alors  $\mathbf{P}(Y^* = X^*) = 1$ ,

Si  $\rho(X, Y) = -1$ , alors  $\mathbf{P}(Y^* = -X^*) = 1$ .

### Explications

Un coefficient de corrélation positif indique donc que  $Y$  est fonction *croissante* de  $X$ .

### Preuve

Si  $\rho(X, Y) \in \{-1, 1\}$ , alors d'après la propriété précédente :

$$\rho(X^*, Y^*) = \rho(X, Y) \in \{-1, 1\}.$$

Donc il existe  $(a, b) \in \mathbf{R}^2$  tel que  $X^* = aY^* + b$  presque sûrement.

Donc  $\mathbf{E}(X^*) = a\mathbf{E}(Y^*) + b$  par linéarité.

Or les variables sont centrées, donc d'espérance nulle. Ainsi  $b = 0$ .

$\sigma_{X^*} = |a|\sigma_{Y^*}$ , et comme les variables sont réduites, elles sont d'écart type égal à 1.

Donc  $|a| = 1$ . En remplaçant dans la formule du coefficient de corrélation, on trouve donc  $\rho(X^*, Y^*) = \frac{a}{|a|}$ . Ainsi,  $a = \rho(X, Y)$ . ■

## 2 INDÉPENDANCE DES VARIABLES ALÉATOIRES

### A Loi conditionnelle

#### Définition 2.1 (Loi conditionnelle)

Pour  $(X, Y)$  deux variables aléatoires finies sur  $\Omega$ , on définit la **loi conditionnelle** de  $X$  sachant  $(Y = y)$  par

$$f_{X|Y=y} : \begin{cases} X(\Omega) & \rightarrow [0, 1] \\ x & \mapsto \mathbf{P}_{[Y=y]}(X = x) = \mathbf{P}(X = x|Y = y) = \frac{\mathbf{P}(X = x, Y = y)}{\mathbf{P}(Y = y)}. \end{cases}$$

On parle aussi de probabilité conditionnelle de  $X$  sachant  $[Y = y]$ .

*Remarque :* Cette définition suppose que  $\mathbf{P}(Y = y) \neq 0$ .

Travailler sur une loi conditionnelle ne modifie pas les variables aléatoires, mais seulement la loi de probabilité. Cela illustre que les mêmes variables aléatoires peuvent être étudiées avec des lois de probabilités différentes.

Lorsque nous ne connaissons que les lois marginales, nous ne pouvons pas retrouver la loi conjointe. Par contre, si nous connaissons à la fois les lois marginales et les lois conditionnelles par rapport à une variable, alors il est possible de reconstruire la loi conjointe. En effet

$$\mathbf{P}((X, Y) = (x, y)) = \mathbf{P}([X = x] \cap [Y = y]) = \mathbf{P}_{[Y=y]}(X = x) \mathbf{P}(Y = y).$$

#### Exemple

Pour le lancer de deux dés, on peut considérer la probabilité que le plus grand dé soit  $y$  sachant que le plus petit est  $x$ . Si les dés sont équilibrés, on obtient la loi de probabilité conditionnelle.

Par exemple pour  $x = 2$  :

$Y$	1	2	3	4	5	6
$\mathbf{P}_{[X=2]}(Y = y)$	0	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$

Cette probabilité conditionnelle se construit facilement à partir du tableau des probabilités conjointes et marginales. Il suffit de sélectionner la bonne ligne  $X = 2$  et de diviser chaque probabilité conjointe par la probabilité marginale (somme des probabilités conjointes de la ligne).

### B Variables aléatoires indépendantes

#### Définition 2.2 (Variables aléatoires indépendantes)

Deux variables aléatoires  $X$  et  $Y$  définies sur  $\Omega$  sont dites **indépendantes**, si

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

#### Propriété 2.3

Si les événements  $[X = x]$  et  $[Y = y]$  sont tous de probabilité non nulle pour  $(x, y) \in (X, Y)(\Omega)$ , alors, l'indépendance de  $X$  et  $Y$  est aussi caractérisée par :

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

ou

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

#### Explications

Il est sous-entendu que les variables sont indépendantes pour la probabilité  $\mathbf{P}$ . L'indépendance dépend du choix de probabilité même si ce n'est pas dit explicitement.

*Remarque :* On remarque que pour des variables aléatoires indépendantes, on a nécessairement  $(X, Y)(\Omega) = X(\Omega) \times Y(\Omega)$ .

#### Exemple

Si  $X$  et  $Y$  désignent respectivement le minimum et le maximum lors du lancer de deux dés équilibrés, alors ces variables aléatoires ne sont pas indépendantes.

Par exemple  $\mathbf{P}(X = 6, Y = 5) = 0$ , mais  $\mathbf{P}(X = 6)\mathbf{P}(Y = 5) = \frac{1}{36} \frac{9}{39} \neq 0$ .

#### Exemple

Si  $X$  et  $Y$  désignent respectivement la valeur du premier dé et la valeur du second dé lors du lancer de deux dés équilibrés, alors ces variables aléatoires sont indépendantes.

**Propriété 2.4** (*Loi de Bernoulli indépendantes*)

Deux variables  $X_1$  et  $X_2$  qui suivent des lois de Bernoulli sont indépendantes si et seulement si les événements  $[X_1 = 1]$  et  $[X_2 = 1]$  sont indépendants.

**Preuve**

Si  $A$  et  $B$  sont indépendants, alors  $\bar{A}$  et  $B$  aussi,  $A$  et  $\bar{B}$  aussi et  $\bar{A}$  et  $\bar{B}$  aussi (en exercice).

Dans le cas des lois de Bernoulli, comme l'univers image se réduit à  $[X = 1]$  et  $[X = 0] = [\bar{X} = 1]$ , la vérification proposée est suffisante. ■

**Théorème 2.5**

Soient deux variables aléatoires **indépendantes**  $X$  et  $Y$  définies sur  $\Omega$ ,  
Pour tous  $A, B$  événements de  $(X, Y)(\Omega)$ ,

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

**Preuve**

On écrit  $A$  et  $B$  comme union d'événements élémentaires. ■

*Remarque* : Il existe aussi une réciproque :

Si pour *tout* couple d'événements  $(A, B)$  de  $(X, Y)(\Omega)$ , on a :

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

alors les variables aléatoires sont indépendantes pour  $\mathbf{P}$ .

**Théorème 2.6** (*L'indépendance passe à l'image*)

Soient deux variables aléatoires indépendantes  $X : \Omega \rightarrow E$  et  $Y : \Omega \rightarrow E'$ .  
Pour toutes les applications  $f : E \rightarrow F$  et  $g : E' \rightarrow F'$ ,

$$f(X) \text{ et } g(Y) \text{ sont indépendantes.}$$

**Explications**

Supposons que l'on associe un gain à chacune des variables aléatoires (selon la valeur qu'elles prennent), si les variables sont indépendantes, alors les gains le sont aussi.

**Preuve**

Admis ■

**Exemple** (*à retenir*)

Si  $X$  et  $Y$  sont indépendantes, alors  $X^2$  est indépendante avec  $Y$ .

**Théorème 2.7** (*Espérance du produit*)

Si deux variables aléatoires  $X$  et  $Y$  sont **indépendantes** pour  $\mathbf{P}$ , alors

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

⚠ La réciproque est **fausse**, il ne suffit pas d'avoir  $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$  pour que les variables aléatoires soient indépendantes. C'est une erreur fréquente.

**Preuve**

On note  $X(\Omega) = \{x_i\}_{i \in [1, p]}$  et  $Y(\Omega) = \{y_j\}_{j \in [1, q]}$ .

$$\begin{aligned} \mathbf{E}(XY) &= \sum_{i=1}^p \sum_{j=1}^q x_i y_j \mathbf{P}((X = x_i) \cap (Y = y_j)) \\ &= \sum_{i=1}^p \sum_{j=1}^q x_i y_j \mathbf{P}(X = x_i) \mathbf{P}(Y = y_j) \quad (\text{variables indépendantes}). \\ &= \sum_{i=1}^p x_i \mathbf{P}(X = x_i) \sum_{j=1}^q y_j \mathbf{P}(Y = y_j) = \mathbf{E}(X)\mathbf{E}(Y). \end{aligned}$$

En particulier, deux variables aléatoires indépendantes ont une covariance nulle. Par contre, ce n'est pas parce que la covariance est nulle que les variables sont indépendantes. ■

**Propriété 2.8**

Si  $X, Y$  sont deux variables aléatoires indépendantes, alors  $\mathbf{Cov}(X, Y) = 0$ .

**Preuve**

Découle de la formule de Koenig-Huygens. ■

Cela permet de réinterpréter la covariance comme une mesure de corrélation entre  $X$  et  $Y$ .

**Propriété 2.9** (*Égalité de Bienaymé*)

Si  $X, Y$  sont deux variables aléatoires indépendantes, alors

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

**Preuve**

Car la covariance est nulle. ■

**Exemple**

Cette égalité permet de retrouver la variance d'une variable suivant une loi binomiale très simplement.

En effet, une variable aléatoire qui suit une loi binomiale, peut être interprétée comme la somme de  $n$  variables indépendantes et de même loi de Bernoulli.

### 3 GÉNÉRALISATION À $n$ VARIABLES ALÉATOIRES

#### A Indépendance mutuelle

**Définition 3.1** (*Variables mutuellement indépendantes*)

Soient  $X_1, X_2, \dots, X_n$  des variables sur un même univers  $\Omega$ .  
 $(X_1, X_2, \dots, X_n)$  sont **mutuellement indépendantes** si

$$\forall (x_i)_{i \in [1, n]} \in (X_1, X_2, \dots, X_n)(\Omega), \quad \mathbf{P} \left( \bigcap_{i \in I} X_i = x_i \right) = \prod_{i \in I} \mathbf{P}(X_i = x_i).$$

**Propriété 3.2** (*Indépendance des sous-familles*)

Si  $(X_1, X_2, \dots, X_n)$  sont des variables mutuellement indépendantes, alors toute sous-famille l'est aussi.

**Propriété 3.3** (*Lien avec l'indépendance des événements*)

Soient  $n$  variables aléatoires mutuellement indépendantes  $(X_i)_{i \in [1, n]}$ ,  
 Pour  $(A_1, A_2, \dots, A_n)$  partie de  $X_1(\Omega) \times X_2(\Omega) \times \dots \times X_n(\Omega)$ , les événements  $(X_i \in A_i)$  sont mutuellement indépendants.

**Preuve**

Admis. ■

**Théorème 3.4** (*L'indépendance passe à l'image - lemme des coalitions.*)

Si  $(X_1, X_2, \dots, X_n)$  sont des variables mutuellement indépendantes,

- alors  $(f_1(X_1), f_2(X_2), \dots, f_n(X_n))$  sont aussi mutuellement indépendantes,
- alors  $g(X_1, X_2, \dots, X_p)$  et  $g(X_{p+1}, \dots, X_n)$  sont aussi indépendantes.

**Exemple**

Si  $X, Y, Z$  sont indépendantes, alors  $X + Y$  et  $Z$  sont indépendantes.

**Propriété 3.5** (*Variance d'une somme*)

Si  $(X_1, X_2, \dots, X_n)$  sont  $n$  variables aléatoires mutuellement indépendantes, alors

$$\mathbf{E}(X_1 X_2 \cdots X_n) = \mathbf{E}(X_1) \mathbf{E}(X_2) \cdots \mathbf{E}(X_n).$$

$$\mathbf{V}(X_1 + X_2 + \cdots + X_n) = \mathbf{V}(X_1) + \mathbf{V}(X_2) + \cdots + \mathbf{V}(X_n).$$

**Preuve**

Par récurrence. ■

#### B Application aux lois de Bernoulli et binomiale

**Propriété 3.6** (*Somme de lois de Bernoulli*)

Soient  $(X_i)_{i \in [1, n]}$  des variables aléatoires mutuellement indépendantes sur  $\Omega$ ,  
 Soit  $p \in [0, 1]$ ,  
 Si  $\forall i \in [1, n], X_i \hookrightarrow \mathcal{B}(p)$ , alors

$$\sum_{i=1}^n X_i \hookrightarrow \mathcal{B}(n, p).$$

**Propriété 3.7** (*Stabilité de la loi binomiale*)

Soient  $X, Y$  deux variables aléatoires **indépendantes** sur  $\Omega$ ,  
 Si  $X \hookrightarrow \mathcal{B}(n, p)$  et  $Y \hookrightarrow \mathcal{B}(m, p)$  avec le **même paramètre**  $p$ ,  
 alors  $X + Y \hookrightarrow \mathcal{B}(n + m, p)$ .

**Preuve** (*intuitive*)

$X$  représente la somme de  $n$  variables de Bernoulli indépendantes et  $Y$ , la somme de  $m$  lois également indépendantes.

Toutes ayant même paramètre,  $X + Y$  représente alors la somme de  $n + m$  loi de Bernoulli indépendantes et de même paramètre.

*Remarque sur cette preuve* : Cet argument n'est pas rigoureux en l'état, en effet, rien ne justifie qu'une variable aléatoire qui suit une loi binomiale peut bien s'écrire comme somme de variables aléatoires de Bernoulli mutuellement indépendantes (nous n'avons vu que la réciproque). Et que les variables pour  $X$  soient indépendantes avec celles pour  $Y$ . ■

**Exemple**

Un institut de recherche effectue des tests sur le comportement des souris dans un labyrinthe. Deux chercheurs se partagent la tâche : sur les 300 essais, l'un en effectue 100, et le second 200.

Le labyrinthe comporte un embranchement, et les souris ont la probabilité  $p$  d'aller à gauche et la probabilité  $1 - p$  d'aller à droite. Les épreuves sont indépendantes les unes des autres.

On compte pour chaque chercheur le nombre de souris qui sont allées à droite.

On note ce nombre  $X_1$  pour le premier chercheur et  $X_2$  pour le second.

Alors  $X_1 \hookrightarrow \mathcal{B}(100, p)$  et  $X_2 \hookrightarrow \mathcal{B}(200, p)$ .

On comprend bien que si on met ensemble les deux expériences, alors

$X_1 + X_2 \hookrightarrow \mathcal{B}(300, p)$ .



**Corollaire 3.8**

Soient  $(X_i)_{i \in [1, m]}$  des variables aléatoires mutuellement indépendantes sur  $\Omega$ ,  
 Soit  $p \in [0, 1]$  et  $(n_1, n_2, \dots, n_m) \in \mathbf{N}^m$ ,  
 Si  $\forall i \in [1, m], X_i \hookrightarrow \mathcal{B}(n_i, p)$ , alors

$$\sum_{i=1}^m X_i \hookrightarrow \mathcal{B}(n_1 + n_2 + \dots + n_m, p).$$

**C Loi faible des grands nombres**

Ce dernier résultat est hors programme.

**Théorème 3.9** (Loi faible des grands nombres  $(\star)$ )

Si  $(X_1, X_2, \dots, X_n, \dots)$  est une suite de variables aléatoires

1. mutuellement indépendantes,
2. de même loi  $\mathbf{P}_X$ .

Alors pour tout  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow +\infty} \mathbf{P} \left( \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mathbf{E}(X) \right| \geq \varepsilon \right) = 0.$$

**Explications**

Ce résultat permet de justifier que, quelque soit notre exigence de précision, lorsque l'on répète un grand nombre de fois une même expérience aléatoire, de façon indépendante, et que l'on calcule la moyenne, on est presque sûr d'obtenir l'espérance (à l'incertitude près fixée au préalable et arbitrairement petite).

**Preuve**

On note  $Y_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ . Alors  $\mathbf{E}(Y_n) = \mathbf{E}(X)$  par linéarité de l'espérance.  
 Or

$$\mathbf{V}(Y_n) = \frac{1}{n^2} \mathbf{V}(X_1 + X_2 + \dots + X_n) = \frac{\mathbf{V}(X_1) + \mathbf{V}(X_2) + \dots + \mathbf{V}(X_n)}{n^2} = \frac{\mathbf{V}(X)}{n}.$$

Donc d'après l'inégalité de Bienaymé-Tchebychev, pour tout  $\varepsilon > 0$

$$\mathbf{P} \left( \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mathbf{E}(X) \right| \geq \varepsilon \right) \leq \frac{\mathbf{V}(X)}{n\varepsilon^2}.$$

Donc pour  $n \rightarrow +\infty$ , on obtient la limite voulue. ■

**Exemple**

Donner un contre-exemple à la loi faible des grands nombres lorsque les variables ne sont pas supposées mutuellement indépendantes.

**Solution :**

On peut poser  $X_1 \hookrightarrow \mathcal{B}(\frac{1}{2})$ , et  $\forall n \geq 2, X_n = X_1$ .

Alors  $X_n \hookrightarrow \mathcal{B}(\frac{1}{2})$ . et  $\frac{1}{n} \sum_{k=1}^n X_k = X_1$ . C'est donc une variable aléatoire à valeurs dans  $\{0, 1\}$ .

En particulier pour  $\varepsilon = \frac{1}{4}$ ,

$$\mathbf{P} \left( \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mathbf{E}(X) \right| \geq \varepsilon \right) = \mathbf{P} \left( \left| X_1 - \frac{1}{2} \right| \geq \frac{1}{4} \right) = 1.$$

Ce qui ne tend pas vers 0 quand  $n \rightarrow +\infty$ .