

# Colles

semaine 14 : 13 décembre - 18 décembre

## I. Notion de couple de v.a.r.

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soient  $X$  et  $Y$  deux v.a.r. discrètes.

- On appelle couple de v.a.r.  $(X, Y)$  l'application définie par :

$$\begin{aligned} \Omega &\rightarrow \mathbb{R}^2 \\ \omega &\mapsto (X(\omega), Y(\omega)) \end{aligned}$$

## II. Loi d'un couple de v.a.r. discrètes, ou loi conjointe

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- On appelle **loi de probabilité du couple**  $(X, Y)$  ou **loi conjointe des v.a.r.  $X$  et  $Y$**  et on note  $\mathbb{P}_{(X,Y)}$  l'application :

$$\begin{aligned} \mathbb{P}_{(X,Y)} : X(\Omega) \times Y(\Omega) &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \mathbb{P}_{(X,Y)}(x, y) = \mathbb{P}([X = x] \cap [Y = y]) \end{aligned}$$

- Autrement dit, la loi du couple  $(X, Y)$  est la donnée de l'ensemble des valeurs  $\mathbb{P}([X = x] \cap [Y = y])$  pour  $x$  parcourant  $X(\Omega)$  et  $y$  parcourant  $Y(\Omega)$ .

*(dans la littérature, on pourra trouver la notation  $[(X, Y) = (x, y)]$  ou  $[X = x, Y = y]$  en lieu et place de  $[X = x] \cap [Y = y]$ )*

### MÉTHODO

### Déterminer la loi d'un couple $(X, Y)$

Afin de déterminer la loi du couple  $(X, Y)$ , on commence TOUJOURS par déterminer les ensembles image  $X(\Omega)$  et  $Y(\Omega)$  des v.a.r.  $X$  et  $Y$ .

### III. Système complet d'événements associé à un couple de v.a.r.

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- La famille  $([X = x] \cap [Y = y])_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}}$  est un système complet d'événements.

Il est appelé **système complet d'événements associé au couple**  $(X, Y)$ .

- On en déduit notamment :

$$\begin{aligned} \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \mathbb{P}([X = x] \cap [Y = y]) &= \sum_{x \in X(\Omega)} \left( \sum_{y \in Y(\Omega)} \mathbb{P}([X = x] \cap [Y = y]) \right) \\ &= \sum_{y \in Y(\Omega)} \left( \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [Y = y]) \right) = 1 \end{aligned}$$

- Démontrons que la famille  $([X = x] \cap [Y = y])_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}}$  est un sce.

1) Démontrons que cette famille est constituée d'événements 2 à 2 incompatibles.

Choisissons deux événements distincts de cette famille.

Soit  $(x_1, y_1) \in X(\Omega) \times Y(\Omega)$  et soit  $(x_2, y_2) \in X(\Omega) \times Y(\Omega)$  tels que  $(x_1, y_1) \neq (x_2, y_2)$ .

$$\begin{aligned} &([X = x_1] \cap [Y = y_1]) \cap ([X = x_2] \cap [Y = y_2]) \\ &= ([X = x_1] \cap [X = x_2]) \cap ([Y = y_1] \cap [Y = y_2]) \\ &= \emptyset \end{aligned}$$

En effet, on a forcément  $x_1 \neq x_2$  ou  $y_1 \neq y_2$  car  $(x_1, y_1) \neq (x_2, y_2)$ .

2) Démontrons que la réunion de ces événements est  $\Omega$ .

$$\begin{aligned} &\bigcup_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} [X = x] \cap [Y = y] \\ &= \bigcup_{x \in X(\Omega)} \left( \bigcup_{y \in Y(\Omega)} [X = x] \cap [Y = y] \right) \\ &= \bigcup_{x \in X(\Omega)} \left( [X = x] \cap \left( \bigcup_{y \in Y(\Omega)} [Y = y] \right) \right) \quad (\text{car } \cap \text{ est distributive} \\ &\quad \text{par rapport à } \cup) \\ &= \bigcup_{x \in X(\Omega)} ([X = x] \cap \Omega) \quad (\text{car } ([Y = y])_{y \in Y(\Omega)} \\ &\quad \text{est un sce}) \\ &= \bigcup_{x \in X(\Omega)} [X = x] = \Omega \quad (\text{car } ([X = x])_{x \in X(\Omega)} \\ &\quad \text{est un sce}) \end{aligned}$$

- Cette famille étant un sce, on en déduit :

$$\sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \mathbb{P}([X = x] \cap [Y = y]) = \mathbb{P} \left( \bigcup_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} [X = x] \cap [Y = y] \right) = \mathbb{P}(\Omega) = 1$$

## IV. Lois conditionnelles

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- 1) • Pour tout  $y \in Y(\Omega)$  tel que  $\mathbb{P}([Y = y]) \neq 0$ , on appelle **loi conditionnelle de  $X$  sachant (que l'événement)  $[Y = y]$**  (est réalisé) l'application :

$$\begin{aligned} X(\Omega) &\rightarrow \mathbb{R} \\ x &\mapsto \mathbb{P}_{[Y=y]}([X = x]) = \frac{\mathbb{P}([X = x] \cap [Y = y])}{\mathbb{P}([Y = y])} \end{aligned}$$

- Autrement dit, la loi conditionnelle de  $X$  sachant  $[Y = y]$  est la donnée des réels :

$$\mathbb{P}_{[Y=y]}([X = x]) \text{ pour } x \text{ décrivant } X(\Omega)$$

- 2) • Pour tout  $x \in X(\Omega)$  tel que  $\mathbb{P}([X = x]) \neq 0$ , on appelle **loi conditionnelle de  $Y$  sachant (que l'événement)  $[X = x]$**  (est réalisé) l'application :

$$\begin{aligned} Y(\Omega) &\rightarrow \mathbb{R} \\ y &\mapsto \mathbb{P}_{[X=x]}([Y = y]) = \frac{\mathbb{P}([X = x] \cap [Y = y])}{\mathbb{P}([X = x])} \end{aligned}$$

- Autrement dit, la loi conditionnelle de  $Y$  sachant  $[X = x]$  est la donnée des réels :

$$\mathbb{P}_{[X=x]}([Y = y]) \text{ pour } y \text{ décrivant } Y(\Omega)$$

### MÉTHODO

### Lien entre lois conditionnelles et loi conjointe via la loi d'une des v.a.r.

Supposons :  $\forall x \in X(\Omega), \mathbb{P}([X = x]) \neq 0$ .

- Tout d'abord, remarquons :

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

Ainsi, si on connaît :

- × la loi du couple  $(X, Y)$ ,
- × la loi de  $X$ ,

alors on obtient les lois conditionnelles de  $Y$  sachant  $[X = x]$  pour tout élément  $x \in X(\Omega)$ .

- On peut aussi lire cette égalité comme suit :

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

Ainsi, si on connaît :

- × la loi de  $X$ ,
- × les lois conditionnelles de  $Y$  sachant  $[X = x]$  pour tout élément  $x \in X(\Omega)$ ,

alors on obtient la loi du couple  $(X, Y)$ .

## V. Lois marginales

### V.1. Définition

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- On appelle **1<sup>ère</sup> loi marginale du couple**  $(X, Y)$  la loi de la v.a.r.  $X$ .
- On appelle **2<sup>ème</sup> loi marginale du couple**  $(X, Y)$  la loi de la v.a.r.  $Y$ .

### V.2. Expression d'une loi marginale via la loi du couple ou via une loi conditionnelle

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

#### 1) Loi de $X$ via la loi du couple $(X, Y)$

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

Loi de  $X$  via les lois conditionnelles de  $X$  sachant  $[Y = y]$  pour tout  $y \in Y(\Omega)$  :

On suppose :  $\forall y \in Y(\Omega), \mathbb{P}([Y = y]) \neq 0$ . Alors :

$$\mathbb{P}([X = x]) = \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y]) \times \mathbb{P}_{[Y=y]}([X = x])$$

#### 2) Loi de $Y$ via la loi du couple $(X, Y)$

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

Loi de  $Y$  via les lois conditionnelles de  $Y$  sachant  $[X = x]$  pour tout  $x \in X(\Omega)$  :

On suppose :  $\forall x \in X(\Omega), \mathbb{P}([X = x]) \neq 0$ . Alors :

$$\mathbb{P}([Y = y]) = \sum_{x \in X(\Omega)} \mathbb{P}([X = x]) \times \mathbb{P}_{[X=x]}([Y = y])$$

#### À RETENIR

- Si l'on connaît la loi du couple  $(X, Y)$ , la loi de  $X$  est déterminée à l'aide de la FPT appliquée sur le système complet d'événements  $([Y = y])_{y \in Y(\Omega)}$  (c'est le SCE associé à  $Y$ ).
- Si l'on connaît la loi du couple  $(X, Y)$ , la loi de  $Y$  est déterminée à l'aide de la FPT appliquée sur le système complet d'événements  $([X = x])_{x \in X(\Omega)}$  (c'est le SCE associé à  $X$ ).
- Le système complet d'événements  $([X = x] \cap [Y = y])_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}}$  n'est **JAMAIS** utilisé dans la formule des probabilités totales. Ce SCE permet essentiellement de conclure :

$$\begin{aligned} \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} \mathbb{P}([X = x] \cap [Y = y]) &= \sum_{x \in X(\Omega)} \left( \sum_{y \in Y(\Omega)} \mathbb{P}([X = x] \cap [Y = y]) \right) \\ &= \sum_{y \in Y(\Omega)} \left( \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [Y = y]) \right) = 1 \end{aligned}$$

## VI. Indépendance de variables aléatoires discrètes

### VI.1. Indépendance de deux variables aléatoires discrètes

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probablisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- Les v.a.r.  $X$  et  $Y$  sont **indépendantes** (pour la probabilité  $\mathbb{P}$ ) si :

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

- Autrement dit, les v.a.r.  $X$  et  $Y$  sont indépendantes si pour tout  $x \in X(\Omega)$  et tout  $y \in Y(\Omega)$ , les événements  $[X = x]$  et  $[Y = y]$  sont indépendants.

#### MÉTHODO

#### Démontrer que deux v.a.r. discrètes ne sont pas indépendantes

- Pour démontrer que  $X$  et  $Y$  ne sont pas indépendantes, il suffit de trouver  $x \in X(\Omega)$  et  $y \in Y(\Omega)$  tels que :

$$\mathbb{P}([X = x] \cap [Y = y]) \neq \mathbb{P}([X = x]) \times \mathbb{P}([Y = y])$$

- On essaiera de trouver  $x \in X(\Omega)$  et  $y \in Y(\Omega)$  tels que :

$$\mathbb{P}([X = x] \cap [Y = y]) = 0 \quad \text{et} \quad \mathbb{P}([X = x]) \neq 0, \mathbb{P}([Y = y]) \neq 0$$

#### Remarque

Soient  $X$  et  $Y$  sont deux v.a.r. discrètes indépendantes.

Alors, tout événement ne dépendant que de la variable  $X$  est indépendant de tout événement ne dépendant que de la variable  $Y$ . Plus précisément, pour tout  $(t_1, t_2) \in \mathbb{R} \times \mathbb{R}$  :

- $\mathbb{P}([X = t_1] \cap [Y \leq t_2]) = \mathbb{P}([X = t_1]) \times \mathbb{P}([Y \leq t_2])$ ,
- $\mathbb{P}([X \leq t_1] \cap [Y \leq t_2]) = \mathbb{P}([X \leq t_1]) \times \mathbb{P}([Y \leq t_2])$ ,
- $\mathbb{P}([X \leq t_1] \cap [Y > t_2]) = \mathbb{P}([X \leq t_1]) \times \mathbb{P}([Y > t_2])$ ,
- ...

Cela provient essentiellement de la propriété :  $[Y \leq t_2] = \bigcup_{\substack{y_j \in Y(\Omega) \\ y_j \leq t_2}} [Y = y_j]$

et que  $[X = t_1]$  est indépendant de tout événement  $[Y = y_j]$ .

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probablisé.

Soient  $(X, Y)$  un couple de v.a.r. discrètes.

- 1) Supposons :  $\forall x \in X(\Omega), \mathbb{P}([X = x]) \neq 0$ . Alors :

$X$  et  $Y$  sont indépendantes

$$\Leftrightarrow \forall x \in X(\Omega), \forall y \in Y(\Omega), \mathbb{P}_{[X=x]}([Y = y]) = \mathbb{P}([Y = y])$$

- 2) Supposons :  $\forall y \in Y(\Omega), \mathbb{P}([Y = y]) \neq 0$ . Alors :

$X$  et  $Y$  sont indépendantes

$$\Leftrightarrow \forall x \in X(\Omega), \forall y \in Y(\Omega), \mathbb{P}_{[Y=y]}([X = x]) = \mathbb{P}([X = x])$$

## VI.2. Indépendance mutuelle de variables aléatoires discrètes

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probablisé.

Soient  $X_1, X_2, \dots, X_n$  (avec  $n \geq 2$ ) des v.a.r. discrètes.

- (CULTURE) Les v.a.r.  $X_1, X_2, \dots, X_n$  sont **(mutuellement) indépendantes** (pour la probabilité  $\mathbb{P}$ ) lorsque :

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), \forall J \subset \llbracket 1, n \rrbracket,$$

$$\mathbb{P} \left( \bigcap_{j \in J} [X_j = x_j] \right) = \prod_{j \in J} \mathbb{P}([X_j = x_j])$$

- Cette propriété est équivalente à (c'est la définition du programme officiel) :

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega),$$

$$\mathbb{P} \left( \bigcap_{i=1}^n [X_i = x_i] \right) = \prod_{i=1}^n \mathbb{P}([X_i = x_i])$$

- On dit que  $(X_n)_{n \in \mathbb{N}^*}$  est une suite de variables (mutuellement) indépendantes (pour la probabilité  $\mathbb{P}$ ) lorsque :

pour tout  $n \geq 2$ , les variables  $X_1, X_2, \dots, X_n$  sont (mutuellement) indépendantes

### Remarque (CULTURE)

- Des deux définitions, c'est celle du programme officiel (évidemment!) qu'il faut retenir.
- L'intérêt de la première définition est qu'elle permet d'insister sur un point fondamental de la notion d'indépendance mutuelle : si les v.a.r.  $X_1, \dots, X_n$  sont (mutuellement) indépendantes, alors, pour tout  $J \subset \llbracket 1, n \rrbracket$ , les variables de la famille  $(X_j)_{j \in J}$  sont indépendantes.
- L'équivalence entre les deux définitions est un peu technique.

Il est évident que la première définition donnée implique celle du programme officiel (il suffit de prendre  $J = \llbracket 1, n \rrbracket$ ).

La réciproque est plus difficile à obtenir. Pour plus de lisibilité, on prend  $n = 3$ . Supposons :

$$\forall (x_1, x_2, x_3) \in X_1(\Omega) \times X_2(\Omega) \times X_3(\Omega), \mathbb{P} \left( \bigcap_{i=1}^3 [X_i = x_i] \right) = \prod_{i=1}^3 \mathbb{P}([X_i = x_i])$$

Notons  $([X_2 = u_2])_{u_2 \in X_2(\Omega)}$  le sce associé à  $X_2$ . Soit  $(x_1, x_3) \in X_1(\Omega) \times X_3(\Omega)$ .

Alors, d'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}([X_1 = x_1] \cap [X_3 = x_3]) &= \sum_{u_2 \in X_2(\Omega)} \mathbb{P}([X_1 = x_1] \cap [X_2 = u_2] \cap [X_3 = x_3]) \\ &= \sum_{u_2 \in X_2(\Omega)} \mathbb{P}([X_1 = x_1]) \times \mathbb{P}([X_2 = u_2]) \times \mathbb{P}([X_3 = x_3]) \\ &= \mathbb{P}([X_1 = x_1]) \times \mathbb{P}([X_3 = x_3]) \sum_{u_2 \in X_2(\Omega)} \mathbb{P}([X_2 = u_2]) \\ &= \mathbb{P}([X_1 = x_1]) \times \mathbb{P}([X_3 = x_3]) \end{aligned}$$

On démontre ainsi que si  $(X_1, X_2, X_3)$  sont indépendantes, il en est de même des variables de la famille  $(X_j)_{j \in J}$  avec  $J = \{1, 3\}$  (il faudrait le faire pour tout  $J \subset \llbracket 1, 3 \rrbracket$ !).

### VI.3. Lemme des coalitions

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

Soient  $f : X(\Omega) \rightarrow \mathbb{R}$  et  $g : Y(\Omega) \rightarrow \mathbb{R}$  deux fonctions.

- Cas de 2 v.a.r.

$X$ et $Y$ indépendantes $\Rightarrow f(X)$ et $g(Y)$ sont indépendantes
--

- Généralisation à  $n$  v.a.r.

Soient  $X_1, \dots, X_n$  des v.a.r. discrètes.

Soient  $f_1 : X_1(\Omega) \rightarrow \mathbb{R}, \dots, f_n : X_n(\Omega) \rightarrow \mathbb{R}$  des fonctions.

$X_1, \dots, X_n$ v.a.r. discrètes mutuellement indépendantes	$\Rightarrow$	$f_1(X_1), \dots, f_n(X_n)$ v.a.r. discrètes mutuellement indépendantes
---	---------------	---

$X_1, \dots, X_n$ v.a.r. discrètes mutuellement indépendantes	$\Rightarrow$	Toute v.a.r. s'exprimant en fonction des v.a.r. $X_1, \dots, X_p$ est indépendante de toute v.a.r. s'exprimant en fonction des v.a.r. $X_{p+1}, \dots, X_n$ (pour $p \in \llbracket 2, n-1 \rrbracket$ )
---	---------------	--

#### Remarque

Soient  $X_1, X_2, \dots, X_n$  des v.a.r. discrètes mutuellement indépendantes.

Alors toute famille de  $n$  événements dont chacun est construit à l'aide d'une v.a.r.  $X_i$  est une famille d'événements mutuellement indépendants.

Plus précisément, pour tout  $(t_1, \dots, t_n) \in \mathbb{R}^n$  :

- × les événements  $[X_1 = t_1], [X_2 \leq t_2], \dots, [X_n \leq t_n]$  sont mutuellement indépendants,
- × les événements  $[X_1 \leq t_1], [X_2 > t_2], \dots, [X_n \leq t_n]$  sont mutuellement indépendants,
- × ...

#### Exemple

- Soient  $X_1, \dots, X_5$  des v.a.r. discrètes mutuellement indépendantes. Alors :

- × les v.a.r.  $X_1, X_2^2, 2X_3, e^{X_4} - 1$  et  $|X_5|$  sont mutuellement indépendantes.
- × les v.a.r.  $2X_1X_3 - X_5$  et  $X_2^2$  sont indépendantes.
- × les v.a.r.  $\min(X_1, X_2)$  et  $\max(X_3, X_4, X_5)$  sont indépendantes.

- Soit  $(X_n)_{n \in \mathbb{N}^*}$  une suite de v.a.r. mutuellement indépendantes.

Alors, pour tout  $n \in \mathbb{N}^*$ ,  $S_n = \sum_{k=1}^n X_k$  et  $X_{n+1}$  sont indépendantes.

- Si  $X^2$  et  $Y^2$  ne sont pas indépendantes, alors, en procédant par l'absurde, on démontre que  $X$  et  $Y$  ne le sont pas non plus.

## VII. Opérations sur les v.a.r. discrètes

### VII.1. Cas général : loi de $g(X, Y)$

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

Soit  $g : X(\Omega) \times Y(\Omega) \rightarrow \mathbb{R}$ .

- La v.a.r.  $Z = g(X, Y)$  est une v.a.r. discrète.
- L'ensemble des valeurs prises par  $Z = g(X, Y)$  est donnée par :

$$\begin{aligned} Z(\Omega) &= \{g(X(\omega), Y(\omega)) \mid \omega \in \Omega\} \\ &\subseteq \{g(x, y) \mid x \in X(\Omega), y \in Y(\Omega)\} \end{aligned}$$

- La loi de  $Z = g(X, Y)$  est donnée par (aspect théorique) :

$$\forall z \in Z(\Omega), \mathbb{P}([Z = z]) = \sum_{\substack{(x, y) \in X(\Omega) \times Y(\Omega) \\ z = g(x, y)}} \mathbb{P}([X = x] \cap [Y = y])$$

- En pratique, on optera pour l'une des deux rédactions suivantes.

1) La famille  $([X = x])_{x \in X(\Omega)}$  est un sce.

Soit  $z \in Z(\Omega)$ . D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}([Z = z]) &= \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [g(X, Y) = z]) \\ &= \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [g(x, Y) = z]) \end{aligned}$$

2) La famille  $([Y = y])_{y \in Y(\Omega)}$  est un sce.

Soit  $z \in Z(\Omega)$ . D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}([Z = z]) &= \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y] \cap [g(X, Y) = z]) \\ &= \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y] \cap [g(X, y) = z]) \end{aligned}$$

Le choix de l'introduction du sce  $([Y = y])_{y \in Y(\Omega)}$  ou  $([X = x])_{x \in X(\Omega)}$  est guidé par les lois de  $X$  et  $Y$ . On optera toujours pour le sce le plus simple.



## VII.2. Cas particuliers (opérations classiques)

### a) Loi de la somme de deux v.a.r. discrètes

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- La v.a.r.  $X + Y$  est une v.a.r. discrète.
- On détermine la loi de  $X + Y$  à l'aide d'une des deux rédactions suivantes.

1) La famille  $([X = x])_{x \in X(\Omega)}$  est un sce.

Soit  $z \in (X + Y)(\Omega)$ . D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}([X + Y = z]) &= \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [X + Y = z]) \\ &= \sum_{\substack{x \in X(\Omega) \\ \text{tq } z - x \in Y(\Omega)}} \mathbb{P}([X = x] \cap [Y = z - x]) \end{aligned}$$

2) La famille  $([Y = y])_{y \in Y(\Omega)}$  est un sce.

Soit  $z \in (X + Y)(\Omega)$ . D'après la formule des probabilités totales :

$$\begin{aligned} \mathbb{P}([X + Y = z]) &= \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y] \cap [X + Y = z]) \\ &= \sum_{\substack{y \in Y(\Omega) \\ \text{tq } z - y \in X(\Omega)}} \mathbb{P}([Y = y] \cap [X = z - y]) \end{aligned}$$

Le choix de l'introduction du sce  $([Y = y])_{y \in Y(\Omega)}$  ou  $([X = x])_{x \in X(\Omega)}$  est guidé par les lois de  $X$  et  $Y$ . On optera toujours pour le sce le plus simple.

#### À RETENIR

- On note que la probabilité  $\mathbb{P}([X = x] \cap [Y = z - x])$  est nulle dès que  $z - x$  n'appartient pas à  $Y(\Omega)$  puisqu'alors  $[Y = z - x] = \emptyset$ .
- Ceci a pour conséquence de restreindre les indices de sommation puisque :

$$\mathbb{P}([X = x] \cap [Y = z - x]) = \mathbb{P}(\emptyset) = 0 \quad \text{dès que } z - x \notin Y(\Omega)$$

#### MÉTHODO

**Trouver les indices convenables lors de la détermination de la loi de  $X + Y$**

Il est fréquent, lors de l'étape de restriction des indices de la somme de tomber sur une contrainte s'exprimant par deux inégalités :

$$a \leq i \leq b \quad \text{ET} \quad c \leq i \leq d \quad \text{où } (a, b, c, d) \in (\mathbb{R})^4$$

On pourra alors utiliser le résultat suivant :

$$(a \leq i \leq b \quad \text{ET} \quad c \leq i \leq d) \quad \Leftrightarrow \quad \max(a, c) \leq i \leq \min(b, d)$$

(à gauche c'est le plus grand élément qui contraint le plus, à droite c'est le plus petit élément qui contraint le plus)

**1) Stabilité des lois binomiales**

Soit  $p \in ]0, 1[$  et soient  $m \in \mathbb{N}^*$  et  $n \in \mathbb{N}^*$ .

- |  |               |   |
|--|---------------|---|
| <ul style="list-style-type: none"> <li>• <math>X \hookrightarrow \mathcal{B}(m, p)</math> et <math>Y \hookrightarrow \mathcal{B}(n, p)</math></li> <li>• <math>X</math> et <math>Y</math> indépendantes</li> </ul> | $\Rightarrow$ | $X + Y \hookrightarrow \mathcal{B}(m + n, p)$ |
|--|---------------|---|

**2) Stabilité des lois de Poisson**

Soient  $\lambda \in \mathbb{R}^{+*}$  et  $\mu \in \mathbb{R}^{+*}$ .

- |  |               |  |
|--|---------------|--|
| <ul style="list-style-type: none"> <li>• <math>X \hookrightarrow \mathcal{P}(\lambda)</math> et <math>Y \hookrightarrow \mathcal{P}(\mu)</math></li> <li>• <math>X</math> et <math>Y</math> indépendantes</li> </ul> | $\Rightarrow$ | $X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$ |
|--|---------------|--|

*Démonstration.*

- Comme  $X(\Omega) = \mathbb{N}$  et  $Y(\Omega) = \mathbb{N}$ , on a :  $(X + Y)(\Omega) \subset \mathbb{N}$ .
- La famille  $([X = i])_{i \in \mathbb{N}}$  est un système complet d'événements.  
Soit  $k \in \mathbb{N}$ . D'après la formule des probabilités totales :

$$\begin{aligned}
 & \mathbb{P}([X + Y = k]) \\
 = & \sum_{i=0}^{+\infty} \mathbb{P}([X = i] \cap [X + Y = k]) \\
 = & \sum_{i=0}^{+\infty} \mathbb{P}([X = i] \cap [Y = k - i]) \\
 = & \sum_{i=0}^{+\infty} \mathbb{P}([X = i]) \mathbb{P}([Y = k - i]) && \text{(par indépendance de } X \text{ et } Y) \\
 = & \sum_{\substack{i=0 \\ k-i \in Y(\Omega)}}^{+\infty} \mathbb{P}([X = i]) \mathbb{P}([Y = k - i]) + \sum_{\substack{i=0 \\ k-i \notin Y(\Omega)}}^{+\infty} \mathbb{P}([X = i]) \mathbb{P}([Y = k - i]) && \text{(car } [Y = k - i] = \emptyset \text{ si } k - i \notin Y(\Omega)) \\
 = & \sum_{i=0}^k \mathbb{P}([X = i]) \mathbb{P}([Y = k - i])
 \end{aligned}$$

- La dernière ligne est obtenue en constatant :

$$\left\{ \begin{array}{l} k - i \in Y(\Omega) = \mathbb{N} \\ i \in \llbracket 0, +\infty \llbracket \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} 0 \leq k - i \\ 0 \leq i \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} i \leq k \\ 0 \leq i \end{array} \right\} \Leftrightarrow \{ 0 \leq i \leq k \}$$

- Ainsi, en reprenant les égalités précédentes :

$$\begin{aligned}
 \mathbb{P}([X + Y = k]) &= \sum_{i=0}^k \mathbb{P}([X = i]) \mathbb{P}([Y = k - i]) \\
 &= \sum_{i=0}^k \frac{\lambda^i}{i!} e^{-\lambda} \frac{\mu^{k-i}}{(k-i)!} e^{-\mu} && \text{(car } X \hookrightarrow \mathcal{P}(\lambda) \text{ et } Y \hookrightarrow \mathcal{P}(\mu)) \\
 &= e^{-\lambda} e^{-\mu} \sum_{i=0}^k \frac{1}{i! (k-i)!} \lambda^i \mu^{k-i} = e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{i=0}^k \frac{k!}{i! (k-i)!} \lambda^i \mu^{k-i} \\
 &= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{i=0}^k \binom{k}{i} \lambda^i \mu^{k-i} = e^{-(\lambda+\mu)} \frac{1}{k!} (\lambda + \mu)^k && \text{(d'après la formule du binôme de Newton)}
 \end{aligned}$$

Finalement :  $X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$ . □

### b) Loi du produit de deux v.a.r. discrètes

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

- $XY$  est une v.a.r. discrète.
- On détermine la loi de  $XY$  à l'aide d'une des deux rédactions suivantes.

1) La famille  $([X = x])_{x \in X(\Omega)}$  est un sce.

Ainsi, d'après la formule des probabilités totales :

$$\begin{aligned} \forall z \in (XY)(\Omega), \mathbb{P}([XY = z]) &= \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [XY = z]) \\ &= \sum_{x \in X(\Omega)} \mathbb{P}([X = x] \cap [xY = z]) \end{aligned}$$

2) La famille  $([Y = y])_{y \in Y(\Omega)}$  est un sce.

Ainsi, d'après la formule des probabilités totales :

$$\begin{aligned} \forall z \in (XY)(\Omega), \mathbb{P}([XY = z]) &= \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y] \cap [XY = z]) \\ &= \sum_{y \in Y(\Omega)} \mathbb{P}([Y = y] \cap [yX = z]) \end{aligned}$$

Le choix de l'introduction du sce  $([Y = y])_{y \in Y(\Omega)}$  ou  $([X = x])_{x \in X(\Omega)}$  est guidé par les lois de  $X$  et  $Y$ . On optera toujours pour le sce le plus simple.

### c) Loi du maximum / minimum de deux v.a.r. discrètes

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soient  $X$  et  $Y$  deux v.a.r. discrètes **indépendantes**.

On note  $F_X$  et  $F_Y$  les fonctions de répartition de ces v.a.r.

On note  $U = \min(X, Y)$  et  $V = \max(X, Y)$ .

Remarquons tout d'abord :

$$\forall t \in \mathbb{R}, [\min(X, Y) > t] = [X > t] \cap [Y > t]$$

$$\forall t \in \mathbb{R}, [\max(X, Y) \leq t] = [X \leq t] \cap [Y \leq t]$$

- 1) • La v.a.r.  $U = \min(X, Y)$  est une v.a.r. discrète.
- La fonction de répartition de  $U = \min(X, Y)$  vérifie :

$$\forall t \in \mathbb{R}, F_U(t) = 1 - (1 - F_X(t)) (1 - F_Y(t))$$

- 2) • La v.a.r.  $V = \max(X, Y)$  est une v.a.r. discrète.
- La fonction de répartition de  $V = \max(X, Y)$  vérifie :

$$\forall t \in \mathbb{R}, F_V(t) = F_X(t) F_Y(t)$$

*Démonstration.*

• Soit  $t \in \mathbb{R}$ .

$$\begin{aligned}
 & F_V(t) \\
 &= \mathbb{P}([Z \leq t]) \\
 &= \mathbb{P}([\max(X, Y) \leq t]) \\
 &= \mathbb{P}([X \leq t] \cap [Y \leq t]) \\
 &= \mathbb{P}([X \leq t]) \mathbb{P}([Y \leq t]) \quad (\text{car } X \text{ et } Y \text{ sont} \\
 &= F_X(t) F_Y(t) \quad \text{indépendantes})
 \end{aligned}$$

• Soit  $t \in \mathbb{R}$ .

$$\begin{aligned}
 & F_U(t) \\
 &= \mathbb{P}(U \leq t) \\
 &= 1 - \mathbb{P}([U > t]) \\
 &= 1 - \mathbb{P}(\min(X, Y) > t) \\
 &= 1 - \mathbb{P}([X > t] \cap [Y > t]) \\
 &= 1 - \mathbb{P}([X > t]) \mathbb{P}([Y > t]) \quad (\text{car } X \text{ et } Y \text{ sont} \\
 &= 1 - (1 - F_X(t)) (1 - F_Y(t)) \quad \text{indépendantes}) \quad \square
 \end{aligned}$$

## VIII. Calculs d'espérance

### VIII.1. Espérance de $Z = g(X, Y)$

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

Alors, sous réserve de convergence absolue :

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

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

#### 1) Si $X$ et $Y$ sont finies

Notons alors  $X(\Omega) = \{x_1, \dots, x_m\}$  et  $Y(\Omega) = \{y_1, \dots, y_n\}$ .

La v.a.r.  $g(X, Y)$  admet (sans autre hypothèse!) une espérance donnée par :

$$\begin{aligned}
 \mathbb{E}(g(X, Y)) &= \sum_{i=1}^m \left( \sum_{j=1}^n g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right) \\
 &= \sum_{j=1}^n \left( \sum_{i=1}^m g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right)
 \end{aligned}$$

#### 2) Si $X$ est finie et $Y$ infinie

• Notons alors  $X(\Omega) = \{x_1, \dots, x_m\}$  et  $Y(\Omega) = \{y_j \mid j \in \mathbb{N}\}$ .

La v.a.r.  $g(X, Y)$  admet une espérance ssi pour tout  $i \in \llbracket 1, m \rrbracket$ , la série

$\sum_{j \geq 0} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j])$  est absolument convergente. Dans ce cas :

$$\begin{aligned}
 \mathbb{E}(g(X, Y)) &= \sum_{i=1}^m \left( \sum_{j=0}^{+\infty} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right) \\
 &= \sum_{j=0}^{+\infty} \left( \sum_{i=1}^m g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right)
 \end{aligned}$$

2) Si  $X$  est finie et  $Y$  infinie (suite)

- Le cas  $X$  infinie et  $Y$  finie se traite de manière similaire.

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

La v.a.r.  $g(X, Y)$  admet une espérance ssi pour tout  $j \in \llbracket 1, n \rrbracket$ , la série

$\sum_{i \geq 0} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j])$  est absolument convergente. Dans ce cas :

$$\begin{aligned} \mathbb{E}(g(X, Y)) &= \sum_{i=0}^{+\infty} \left( \sum_{j=1}^n g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right) \\ &= \sum_{j=1}^n \left( \sum_{i=0}^{+\infty} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right) \end{aligned}$$

3) Si  $X$  est infinie et  $Y$  infinie

Notons alors  $X(\Omega) = \{x_i \mid i \in \mathbb{N}\}$  et  $Y(\Omega) = \{y_j \mid j \in \mathbb{N}\}$ .

La v.a.r.  $g(X, Y)$  admet une espérance ssi :

- a) pour tout  $i \in \mathbb{N}$ , la série  $\sum_{j \geq 0} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j])$  est absolument convergente.

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

Dans ce cas :

$$\begin{aligned} \mathbb{E}(g(X, Y)) &= \sum_{i=0}^{+\infty} \left( \sum_{j=0}^{+\infty} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right) \\ &= \sum_{j=0}^{+\infty} \left( \sum_{i=0}^{+\infty} g(x_i, y_j) \mathbb{P}([X = x_i] \cap [Y = y_j]) \right) \end{aligned}$$

## VIII.2. Espérance d'une somme

Soit  $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ .

Soient  $X_1, \dots, X_n$  des v.a.r. discrètes admettant une espérance.

- 1) Alors  $\lambda_1 \cdot X_1 + \dots + \lambda_n \cdot X_n$  admet une espérance.

- 2) De plus :

$$\mathbb{E} \left( \sum_{i=1}^n \lambda_i \cdot X_i \right) = \sum_{i=1}^n \lambda_i \mathbb{E}(X_i)$$

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

$$\mathbb{E} \left( \sum_{i=1}^n X_i \right) = \sum_{i=1}^n \mathbb{E}(X_i)$$

### MÉTHODO

### Démontrer qu'une v.a.r. admet une espérance

On retiendra que si une v.a.r. s'écrit comme combinaison linéaire (resp. somme) de v.a.r. qui admettent une espérance, alors cette v.a.r. admet une espérance.

### VIII.3. Espérance d'un produit

#### a) Formulation générale

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose que  $X$  et  $Y$  admettent chacune un moment d'ordre 2.

1) Alors  $XY$  admet une espérance.

2) De plus :

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

#### b) Cas des v.a.r. indépendantes

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé.

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose :

- $X$  et  $Y$  admettent une espérance.
- $X$  et  $Y$  sont indépendantes.

Alors  $XY$  admet une espérance et :

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



Ce théorème n'énonce pas un équivalence.

Autrement dit, il existe des variables aléatoires  $X$  et  $Y$  :

- × qui ne sont pas indépendantes,
- × qui vérifient  $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$ .

## IX. Calculs de variance et covariance

### IX.1. Covariance

#### a) Définition

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose que  $X$  et  $Y$  admettent chacune un moment d'ordre 2.

La **covariance** de  $X$  et  $Y$  est définie par :

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

En particulier :

$$\text{Cov}(X, X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{V}(X)$$

## b) Formule de Kœnig-Huygens

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose que  $X$  et  $Y$  admettent chacune un moment d'ordre 2.

Alors :

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

## c) Propriétés de l'opérateur Cov

Soient  $X, Y, X_i, Y_i$  des v.a.r. discrètes.

Supposons que ces v.a.r. admettent chacune un moment d'ordre 2.

L'opérateur de covariance vérifie les propriétés suivantes.

1)  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

(propriété de symétrie)

2)  $\forall (\lambda_1, \lambda_2) \in \mathbb{R}^2, \text{Cov}(\lambda_1 X_1 + \lambda_2 X_2, Y) = \lambda_1 \text{Cov}(X_1, Y) + \lambda_2 \text{Cov}(X_2, Y)$

(linéarité à gauche)

3)  $\forall (\mu_1, \mu_2) \in \mathbb{R}^2, \text{Cov}(X, \mu_1 Y_1 + \mu_2 Y_2) = \mu_1 \text{Cov}(X, Y_1) + \mu_2 \text{Cov}(X, Y_2)$

(linéarité à droite)

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

Démonstration.

1) D'après la formule de Kœnig-Huygens :

$$\begin{aligned} \text{Cov}(Y, X) &= \mathbb{E}(YX) - \mathbb{E}(Y) \mathbb{E}(X) \\ &= \mathbb{E}(XY) - \mathbb{E}(X) \mathbb{E}(Y) \\ &= \text{Cov}(X, Y) \end{aligned}$$

2) Soit  $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ . Comme  $X_1$  et  $X_2$  admettent un moment d'ordre 2, c'est aussi le cas de  $\lambda_1 X_1 + \lambda_2 X_2$ . De plus :

$$\begin{aligned} \text{Cov}(\lambda_1 X_1 + \lambda_2 X_2, Y) &= \mathbb{E}((\lambda_1 X_1 + \lambda_2 X_2) Y) - \mathbb{E}(\lambda_1 X_1 + \lambda_2 X_2) \mathbb{E}(Y) && \text{(par la formule de Kœnig-Huygens)} \\ &= \mathbb{E}(\lambda_1 X_1 Y + \lambda_2 X_2 Y) - (\lambda_1 \mathbb{E}(X_1) + \lambda_2 \mathbb{E}(X_2)) \mathbb{E}(Y) && \text{(par linéarité de l'espérance)} \\ &= \lambda_1 \mathbb{E}(X_1 Y) + \lambda_2 \mathbb{E}(X_2 Y) - \lambda_1 \mathbb{E}(X_1) \mathbb{E}(Y) - \lambda_2 \mathbb{E}(X_2) \mathbb{E}(Y) && \text{(par linéarité de l'espérance)} \\ &= \lambda_1 (\mathbb{E}(X_1 Y) - \mathbb{E}(X_1) \mathbb{E}(Y)) + \lambda_2 (\mathbb{E}(X_2 Y) - \mathbb{E}(X_2) \mathbb{E}(Y)) \\ &= \lambda_1 \text{Cov}(X_1, Y) + \lambda_2 \text{Cov}(X_2, Y) && \text{(par la formule de Kœnig-Huygens)} \end{aligned}$$

3) Soit  $(\mu_1, \mu_2) \in \mathbb{R}^2$ .

$$\begin{aligned} \text{Cov}(X, \mu_1 Y_1 + \mu_2 Y_2) &= \text{Cov}(\mu_1 Y_1 + \mu_2 Y_2, X) && \text{(par symétrie)} \\ &= \mu_1 \text{Cov}(Y_1, X) + \mu_2 \text{Cov}(Y_2, X) && \text{(par le point 2))} \end{aligned}$$

4) Soit  $a \in \mathbb{R}$ .  $\text{Cov}(X, a) = \mathbb{E}(aX) - \mathbb{E}(a) \mathbb{E}(X) = a \mathbb{E}(X) - a \mathbb{E}(X) = 0$ . □

## d) Covariance et indépendance

Soient  $X$  et  $Y$  deux v.a.r. (discrètes).

On suppose que  $X$  et  $Y$  admettent chacune un moment d'ordre 2.

$$X \text{ et } Y \text{ indépendantes} \Rightarrow \text{Cov}(X, Y) = 0$$

**Remarque**

- Ce résultat n'est que la reformulation, avec le vocabulaire de la covariance de la propriété énonçant que, dans le cas où les v.a.r.  $X$  et  $Y$  sont indépendantes,  $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ .
- Généralement c'est la contraposée de cet énoncé qui est utilisée. Elle permet de démontrer que deux v.a.r.  $X$  et  $Y$  ne sont pas indépendantes.

$$\text{Cov}(X, Y) \neq 0 \Rightarrow X \text{ et } Y \text{ ne sont pas indépendantes}$$



Ce résultat **N'EST PAS** une équivalence.

Autrement dit, il existe des variables aléatoires  $X$  et  $Y$  :

- × qui ne sont pas indépendantes,
- × qui vérifient  $\text{Cov}(X, Y) = 0$ .

**IX.2. Variance d'une somme**

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose que  $X$  et  $Y$  admettent chacune un moment d'ordre 2.

(i.e.  $X$  et  $Y$  admettent une variance)

1) Alors  $X + Y$  admet une variance.

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

3) Enfin :

$$X \text{ et } Y \text{ indépendantes} \Rightarrow \mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$$

4) **Généralisation**

- On suppose :
- $X_1, \dots, X_n$  admettent une variance.
  - $X_1, \dots, X_n$  mutuellement indépendantes.

Alors  $X_1 + \dots + X_n$  admet une variance et :

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



### IX.3. Coefficient de corrélation linéaire

#### a) Définition

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose que  $X$  et  $Y$  admettent des variances non nulles.

Le **coefficient de corrélation linéaire de  $X$  et de  $Y$**  est :

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

#### b) Propriétés

Soit  $(X, Y)$  un couple de v.a.r. discrètes.

On suppose que  $X$  et  $Y$  admettent des variances non nulles.

1) Inégalité de Cauchy-Schwarz :

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

On en déduit (reformulation) :  $|\rho(X, Y)| \leq 1$

2) •  $|\rho(X, Y)| = 1 \Leftrightarrow$  une des v.a.r. est, presque sûrement, une fonction affine de l'autre

• Plus précisément :

×  $\rho(X, Y) = 1$  ssi une des v.a.r. est presque sûrement une fonction affine strictement croissante de l'autre v.a.r. .

Cela signifie qu'il existe  $a > 0$  et  $b \in \mathbb{R}$  tels que :

$$\mathbb{P}([X = aY + b]) = 1 \quad \text{OU} \quad \mathbb{P}([Y = aX + b]) = 1$$

×  $\rho(X, Y) = -1$  ssi une des v.a.r. est presque sûrement une fonction affine strictement décroissante de l'autre v.a.r. .

Cela signifie qu'il existe  $a > 0$  et  $b \in \mathbb{R}$  tels que :

$$\mathbb{P}([X = -aY + b]) = 1 \quad \text{OU} \quad \mathbb{P}([Y = -aX + b]) = 1$$

#### Remarque

- Les valeurs intermédiaires entre  $-1$  et  $1$  renseignent sur le degré de dépendance linéaire entre les deux v.a.r. Plus le coefficient est proche des valeurs extrêmes  $-1$  et  $1$ , plus la corrélation linéaire entre les v.a.r. est forte.
- Deux v.a.r. dont la covariance est nulle (et donc le coefficient de corrélation linéaire est nul) sont dites non corrélées.

*Démonstration.*

1) • Considérons la fonction  $f : t \mapsto \mathbb{V}(X + tY)$ .

Remarquons tout d'abord que cette fonction est bien définie. En effet, la v.a.r.  $X + tY$  admet une variance en tant que somme de v.a.r. qui admettent une variance.

• Par ailleurs :

$$\begin{aligned}
 f(t) &= \mathbb{V}(X + tY) = \text{Cov}(X + tY, X + tY) \\
 &= \text{Cov}(X, X + tY) + \text{Cov}(tY, X + tY) && \text{(par linéarité à gauche)} \\
 &= \text{Cov}(X, X) + \text{Cov}(X, tY) + \text{Cov}(tY, tY) && \text{(par linéarité à droite)} \\
 &\quad + \text{Cov}(tY, X) \\
 &= \text{Cov}(X, X) + t \text{Cov}(X, Y) + t \text{Cov}(Y, tY) \\
 &\quad + t \text{Cov}(Y, X) \\
 &= \text{Cov}(X, X) + 2t \text{Cov}(X, Y) + t^2 \text{Cov}(Y, Y) \\
 &= \mathbb{V}(X) + 2\text{Cov}(X, Y) t + \mathbb{V}(Y)t^2
 \end{aligned}$$

Ainsi, la fonction  $f$  est polynomiale de degré 2.

• Or, pour tout  $t \in \mathbb{R}$  :

$$f(t) = \mathbb{V}(X + tY) \geq 0$$

Cette fonction polynomiale étant de signe constant, on en déduit que le discriminant du polynôme associé est de signe négatif. Or :

$$\Delta = (2\text{Cov}(X, Y))^2 - 4\mathbb{V}(X)\mathbb{V}(Y) = 4(\text{Cov}(X, Y))^2 - 4\mathbb{V}(X)\mathbb{V}(Y)$$

Et enfin :

$$\begin{aligned}
 \Delta \leq 0 &\Leftrightarrow (\text{Cov}(X, Y))^2 \leq \mathbb{V}(X)\mathbb{V}(Y) \\
 &\Leftrightarrow \sqrt{(\text{Cov}(X, Y))^2} \leq \sqrt{\mathbb{V}(X)\mathbb{V}(Y)} && \text{(par stricte croissance de } \sqrt{\cdot} \text{)} \\
 &\Leftrightarrow |\text{Cov}(X, Y)| \leq \sqrt{\mathbb{V}(X)}\sqrt{\mathbb{V}(Y)} \\
 &\Leftrightarrow |\text{Cov}(X, Y)| \leq \sigma(X) \sigma(Y) \\
 &\Leftrightarrow \frac{|\text{Cov}(X, Y)|}{\sigma(X) \sigma(Y)} \leq 1 \Leftrightarrow |\rho(X, Y)| \leq 1
 \end{aligned}$$

2) Déterminons maintenant sous quelle condition a lieu l'égalité de l'énoncé.

$$\begin{aligned}
 |\rho(X, Y)| &= 1 \\
 &\Leftrightarrow (\text{Cov}(X, Y))^2 = \mathbb{V}(X)\mathbb{V}(Y) \\
 &\Leftrightarrow \Delta = 0 \\
 &\Leftrightarrow \text{Le polynôme } P \text{ admet une unique racine } \alpha \in \mathbb{R} \\
 &\Leftrightarrow \exists! \alpha \in \mathbb{R}, P(\alpha) = 0 \\
 &\Leftrightarrow \exists! \alpha \in \mathbb{R}, \mathbb{V}(\alpha X + Y) = 0 \\
 &\Leftrightarrow \exists! \alpha \in \mathbb{R}, \alpha X + Y \text{ est constante presque sûrement} \\
 &\Leftrightarrow \exists! \alpha \in \mathbb{R}, \exists \beta \in \mathbb{R}, Y = -\alpha X + \beta \text{ presque sûrement}
 \end{aligned}$$

$(\text{Cov}(X, Y))^2 = \mathbb{V}(X)\mathbb{V}(Y) \Leftrightarrow$ La v.a.r. $Y$ est une transformée affine de $X$ , presque sûrement
--

3) Rappelons que d'après le point précédent :

$$\begin{aligned} & \rho(X, Y) \in \{-1, 1\} \\ \Leftrightarrow & |\rho(X, Y)| = 1 \\ \Leftrightarrow & \text{Le polynôme } P \text{ admet une unique racine } \alpha \in \mathbb{R} \\ \Leftrightarrow & \exists! \alpha \in \mathbb{R}, \exists \beta \in \mathbb{R}, Y = -\alpha X + \beta \text{ presque sûrement} \end{aligned}$$

Par la formule des racines des polynômes de second degré, on obtient que l'unique racine  $\alpha$  de  $P$  s'écrit sous la forme :

$$\begin{aligned} \alpha &= \frac{-2 \operatorname{Cov}(X, Y)}{2 \mathbb{V}(X)} \\ &= -\frac{\operatorname{Cov}(X, Y)}{\mathbb{V}(X)} \frac{\sigma(X) \sigma(Y)}{\sigma(X) \sigma(Y)} \\ &= -\frac{\operatorname{Cov}(X, Y)}{\sigma(X) \sigma(Y)} \frac{\sigma(X) \sigma(Y)}{\mathbb{V}(X)} \\ &= -\rho(X, Y) \frac{\sigma(Y)}{\sigma(X)} \quad (\text{car } \mathbb{V}(X) = (\sigma(X))^2) \end{aligned}$$

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

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

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

□

## Informations concernant cette semaine de colles

### Exercices types

Les compétences attendues sur le chapitre couple de v.a.r. sont les suivantes :

- savoir déterminer la loi d'un couple (et rédaction associée).
- savoir déterminer une loi marginale connaissant la loi du couple.
- savoir déterminer une loi marginale connaissant une loi conditionnelle.
- savoir passer de la loi du couple à la loi conditionnelle (et inversement) à l'aide de la connaissance de la loi de la v.a.r. adéquate.
- savoir démontrer que deux v.a.r.  $X$  et  $Y$  sont non indépendantes en exhibant un couple  $(x, y) \in X(\Omega) \times Y(\Omega)$  qui met en défaut la propriété d'indépendance.  
(on exhibe **un** couple particulier, on ne cherche en aucun cas tous les couples qui mettent en défaut la propriété)
- savoir déterminer la loi d'une v.a.r.  $Z = g(X, Y)$  dans les cas particuliers classiques.  
Plus précisément, savoir déterminer la loi d'une somme  $X + Y$  / d'une différence  $X - Y$  / d'une distance  $|X - Y|$  / d'un produit  $XY$  de deux v.a.r.  
(on pourra se reporter à l'énoncé HEC 2010 - exercice 7 des annales passées, thème Probabilités discrètes)  
Pour ce faire, on testera l'événement  $[Z = z]$  en fonction du sce  $([X = x])_{x \in X(\Omega)}$  ou du sce  $([Y = y])_{y \in Y(\Omega)}$  (on choisira le sce le plus simple) et on déterminera la probabilité de  $\mathbb{P}([Z = z])$  à l'aide de la FPT.
- savoir déterminer la loi du max ou du min de deux v.a.r. discrètes indépendantes (en particulier la loi du min de deux v.a.r. indépendantes suivant la même loi géométrique  $\mathcal{G}(p)$  où  $p \in ]0, 1[$ )
- 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  $\text{Cov}$ .
- 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).

## MÉTHODO

## Calcul de probabilités (rappel / bilan des chapitres précédents)

Afin de résoudre un exercice de calcul de probabilités, il faudra penser au schéma suivant.

0) Introduction des événements basiques (le fait d'avoir tiré une boule blanche au  $i^{\text{ème}}$  tirage, le fait d'avoir obtenu pile au  $i^{\text{ème}}$  tirage, le fait d'avoir obtenu un 6 au  $i^{\text{ème}}$  tirage ...) liés à l'expérience considérée.

Nommage de l'événement  $A$  dont on cherche à déterminer la probabilité.

(ces deux étapes sont parfois directement données dans l'énoncé)

1) Décomposition de l'événement  $A$  à l'aide d'événements basiques.

2) Deux cas se présentent alors :

(i) si cette décomposition fait apparaître une union, il faut retenir le triptyque :

union / incompatibilité / somme
---------------------------------

Dans le cas d'une union finie d'événements

- Si cela est possible, on simplifie cette union (cas d'une union croissante d'événements par exemple).
- Sinon, on vérifie si les événements sont 2 à 2 incompatibles.
  - × si c'est le cas, on utilise l'additivité de  $\mathbb{P}$ .
  - × si ce n'est pas le cas, on peut penser à utiliser la formule du crible.

Dans le cas d'une union infinie d'événements

- On vérifie si les événements sont 2 à 2 incompatibles :
  - × si c'est le cas, on utilise la  $\sigma$ -additivité de  $\mathbb{P}$ .
  - × si ce n'est pas le cas, on se ramène au cas d'une union finie d'événements en utilisant le théorème de la limite monotone.

Si toutes ces tentatives échouent, on peut se ramener au cas d'une intersection d'événements en considérant la formule liant probabilité d'un événement à la probabilité de l'événement contraire.

(ii) si cette décomposition fait apparaître une intersection, il faut retenir le triptyque :

intersection / indépendance / produit
---------------------------------------

Dans le cas d'une intersection finie d'événements

- Si cela est possible, on simplifie cette intersection (cas d'une intersection décroissante d'événements par exemple).
- Sinon, on vérifie si les événements sont mutuellement indépendants.
  - × si c'est le cas, on utilise la formule associée.
  - × si ce n'est pas le cas, on peut penser à utiliser la Formule des Probabilités Composées (FPC).

Dans le cas d'une intersection infinie d'événements

- On se ramène au cas d'une intersection finie d'événements en utilisant le théorème de la limite monotone.

Si toutes ces tentatives échouent, on peut se ramener au cas d'une union d'événements en considérant la formule liant probabilité d'un événement à la probabilité de l'événement contraire.

### Remarque

- Il est à noter que la Formule des Probabilités Totales (FPT) rentre dans ce schéma. En effet, si la famille  $(A_i)_{i \in I}$  est un système complet d'événements, alors tout événement  $B$  s'écrit comme une réunion d'événements 2 à 2 incompatibles.

$$B = \left( \bigcup_{i \in I} A_i \right) \cap B = \bigcup_{i \in I} (A_i \cap B)$$

- L'étape de décomposition des événements est **primordiale**.  
On raisonne TOUJOURS sur les événements et JAMAIS directement sur les probabilités.

~~$\mathbb{P}(A) = 0$  car c'est la probabilité d'obtenir ...~~

(cf démarche de l'exercice sur la limite monotone)

- Lorsqu'il s'agit de raisonner sur les événements, on adopte la rédaction suivante :

~~L'événement  $A$  signifie que ...~~

L'événement  $A$  est réalisé si et seulement si ... ✓

- Afin de déterminer une probabilité conditionnelle  $\mathbb{P}_A(B)$  on pourra rédiger comme suit :

Si l'événement  $A$  est réalisé, c'est que ...

Dans ce cas, l'événement  $B$  est réalisé si et seulement si ...

- Dans un énoncé de probabilités discrètes, on manipule différents niveaux d'objets.

1) Au premier niveau, on trouve l'expérience aléatoire considérée.

On note  $\Omega$  l'univers des possibles : c'est l'**ensemble** des résultats possibles (appelés aussi issues) de l'expérience. Si on considère l'expérience consistant à effectuer trois lancers successifs d'une même pièce, alors :  $\Omega = \{P, F\}^3$ .

Autrement dit,  $\Omega$  est l'ensemble des triplets à coefficients dans l'ensemble  $\{P, F\}$ .

Ces triplets pourront être nommés des 3-lancers (on s'adapte ainsi au vocabulaire des probabilités).

Par exemple,  $\omega = (F, F, P)$  est un 3-lancer qui est un résultat possible de l'expérience. Ce résultat est obtenu si le 1<sup>er</sup> lancer fournit *Face*, le 2<sup>ème</sup> fournit *Face*, le 3<sup>ème</sup> fournit *Pile*.

2) Au deuxième niveau, on trouve les événements : un événement  $A$  n'est rien d'autre qu'un **ensemble** qui regroupe certaines issues de l'expérience. Ainsi :  $A \subset \Omega$  (un événement est un sous-ensemble de l'univers). Par exemple, l'événement  $P_1$  : « obtenir *Pile* au premier lancer » regroupe tous les 3-lancers dont le premier coefficient vaut  $P$ .

$$P_1 = \{ (P, F, F), (P, F, P), (P, P, F), (P, P, P) \}$$

Par exemple,  $\omega = (P, F, F) \in P_1$ . Lorsque  $\omega \in P_1$ , on dit que  $\omega$  **réalise** l'événement  $P_1$ .

3) Au troisième niveau, on trouve les v.a.r. . Ce sont des **applications** particulières :

– elles prennent comme argument un **résultat possible de l'expérience** et renvoient une **valeur réelle**. Considérons la v.a.r.  $X$  qui donne le nombre de *Pile* obtenus au cours de l'expérience. Avec le 3-lancer  $\omega$  précédent, on obtient :  $X(\omega) = X((P, F, F)) = 1$ .

Cela démontre que la v.a.r.  $X$  peut prendre la valeur 1 (on a exhibé un 3-lancer  $\omega$  tel que  $X(\omega) = 1$ ).

– elles sont des machines à créer des événements. Par exemple,  $[X = 2]$  est un événement.

Il regroupe **tous** les 3-lancers  $\omega$  tels que :  $X(\omega) = 2$ .

Autrement dit :  $[X = 2] = \{\omega \in \Omega \mid X(\omega) = 2\} = \{ (P, P, F), (P, F, P), (F, F, P) \}$ .

Ce deuxième point nous replonge au deuxième niveau. Ainsi, pour comprendre le chapitre sur les v.a.r. , il est donc essentiel de maîtriser celui sur les probabilités générales.

**À RETENIR** (détermination de la loi d'un couple de v.a.r. discrètes)

- La loi du couple  $(X, Y)$  n'est JAMAIS obtenue à l'aide de la formule des probabilités totales.
- Pour obtenir la loi du couple  $(X, Y)$  on opère TOUJOURS comme suit.
  - a) Détermination de  $X(\Omega)$  et de  $Y(\Omega)$ .  
On pourra se contenter dans cet ensemble de déterminer un sur-ensemble de ces deux ensembles image ( $X(\Omega) \subset I$  et  $Y(\Omega) \subset J$ ).

b) On rédige alors comme suit.

Soit  $(x, y) \in X(\Omega) \times Y(\Omega)$  (ou plutôt  $(x, y) \in I \times J$ ).

L'événement  $[X = x] \cap [Y = y]$  est réalisé

$\Leftrightarrow$  L'événement  $[X = x]$  est réalisé et l'événement  $[Y = y]$  est réalisé

$\Leftrightarrow$  ... et ...

Cette rédaction ne permet pas de conclure quant à la loi du couple  $(X, Y)$ .

Elle sert simplement à s'assurer de la bonne compréhension de l'expérience aléatoire décrite dans l'énoncé.

c) Il y a alors deux grandes manières de procéder.

(i) Décomposition de l'événement  $[X = x] \cap [Y = y]$

On utilise en priorité cette méthode. Pour ce faire, on pensera à introduire des événements basiques liés à l'expérience :

- ▶ pour tout  $k \in \mathbb{N}^*$ ,  $P_k$  : « on obtient Pile lors du  $k^{\text{ème}}$  lancer »
- ▶ pour tout  $k \in \mathbb{N}^*$ ,  $B_k$  : « on obtient une boule blanche lors du  $k^{\text{ème}}$  tirage »
- ▶ ...

Il est à noter que l'événement  $[X = x] \cap [Y = y]$  peut aussi être décomposé à l'aide d'événements dépendant d'autres v.a.r.

(ii) Méthode par dénombrement

Cette méthode peut être utilisée lorsque :

- ▶ l'expérience consiste à effectuer un nombre fini  $n$  d'épreuves.  
*(en réalité, on peut aussi étendre cette méthode aux expériences qui procèdent à un nombre infinie d'épreuves pour peu que la réalisation de l'événement  $[X = x] \cap [Y = y]$  ne dépende que d'un nombre fini (notons-le  $n$ ) d'épreuves de l'expérience, les résultats des épreuves ultérieures étant libres.)*
- ▶ il y a équiprobabilité de toutes les issues de l'expérience (tous les  $n$ -tirages ou  $n$ -lancers ont même probabilité d'apparaître).

On rédige alors comme suit.

L'événement  $[X = x] \cap [Y = y]$  est réalisé par tous les  $n$ -tirages (ou  $n$ -lancers suivant l'expérience) qui contiennent ...

Un tel  $n$ -lancer est entièrement déterminé par :

- le(s) numéro(s) apparaissant en position ... : ... possibilités.
- le(s) numéro(s) apparaissant en position ... : ... possibilités.
- ...

Le contexte oriente naturellement vers l'une ou l'autre de ces deux méthodes.

- Il est enfin à noter que la loi d'un couple  $(X, Y)$  peut aussi être obtenue à l'aide :
  - × de la loi de  $X$  et des lois conditionnelles de  $Y$  sachant l'événement  $[X = x]$  pour tout  $x \in X(\Omega)$ .  
Il suffit alors simplement d'écrire que pour tout  $(x, y) \in X(\Omega) \times Y(\Omega)$  :

$$\mathbb{P}([X = x] \cap [Y = y]) = \mathbb{P}([X = x]) \times \mathbb{P}_{[X=x]}([Y = y]) \quad (*)$$

- × de la loi de  $Y$  et des lois conditionnelles de  $X$  sachant l'événement  $[Y = y]$  pour tout  $y \in Y(\Omega)$ .  
Il suffit alors simplement d'écrire que pour tout  $(x, y) \in X(\Omega) \times Y(\Omega)$  :

$$\mathbb{P}([X = x] \cap [Y = y]) = \mathbb{P}([Y = y] \cap [X = x]) = \mathbb{P}([Y = y]) \times \mathbb{P}_{[Y=y]}([X = x]) \quad (**)$$

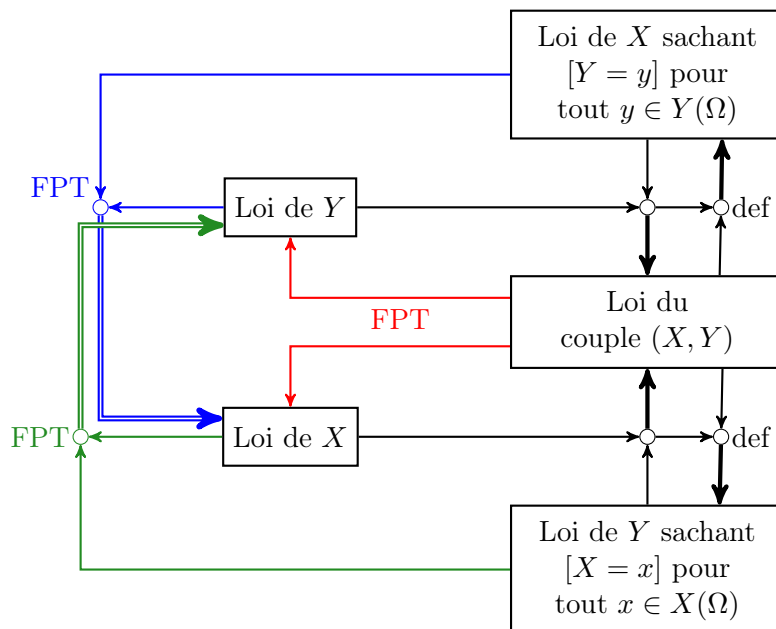
Il y a peu de chance que cette méthode permette de déterminer la loi du couple  $(X, Y)$  si l'énoncé n'est pas écrit dans un contexte de probabilités conditionnelles. Les égalités (\*) et (\*\*) ont un intérêt limité : déterminer une loi conditionnelle est aussi difficile que déterminer une loi de couple. Écrire (\*) ou (\*\*) n'a donc d'intérêt que si l'énoncé amène naturellement à travailler sur les lois conditionnelles.

## Bilan : liens entre les différentes lois

Il faut savoir :

- × faire le lien entre la loi de couples et les lois conditionnelles.
- × déterminer les lois marginales si on connaît la loi du couple,
- × déterminer les lois marginales si on connaît les lois conditionnelles.

Les liens en ces différentes notions sont rappelés dans le schéma suivant.



On retiendra que la formule des probabilités totales est la clé pour déterminer les lois marginales. Si la détermination de la loi du couple / les lois conditionnelles constitue généralement la plus grande difficulté d'un exercice sur les couples, la détermination des lois marginales se résume simplement à une application de la formule des probabilités totales. Des difficultés calculatoires peuvent apparaître mais il n'y a aucune difficulté méthodologique.