

# Chapitre 1: Variable aléatoire discrète

ESILV A2 (2017/2018) – Probabilités 2

*Mohamad Ghassany*

## Table des matières

<b>1</b>	<b>Notion de variable aléatoire réelle (v.a.r.)</b>	<b>1</b>
<b>2</b>	<b>Variables aléatoires discrètes</b>	<b>3</b>
2.1	Définition, loi de probabilité . . . . .	3
2.2	Fonction de répartition d'une variable aléatoire discrète . . . . .	3
<b>3</b>	<b>Moments d'une variable aléatoire discrète</b>	<b>5</b>
3.1	Espérance mathématique . . . . .	5
3.2	Espérance d'une fonction d'une variable aléatoire . . . . .	7
3.3	Variance . . . . .	7
3.4	Ecart-type . . . . .	8
3.5	Inégalité de Bienaymé-Tchebychev . . . . .	9
3.6	Moments non centrés et centrés . . . . .	9
<b>4</b>	<b>Couple de variables aléatoires discrètes</b>	<b>10</b>
4.1	Table de probabilité conjointe . . . . .	10
4.2	Lois marginales . . . . .	11
4.3	Loi conditionnelles . . . . .	11
4.4	Indépendance de variables aléatoires . . . . .	12
4.5	Covariance . . . . .	12
4.6	Coefficient de corrélation linéaire . . . . .	12

## 1 Notion de variable aléatoire réelle (v.a.r.)

Après avoir réalisé une expérience aléatoire, il arrive bien souvent qu'on s'intéresse plus à une fonction du résultat qu'au résultat lui-même. Expliquons ceci au moyen des exemples suivants: lorsqu'on joue au dés, certains jeux accordent de l'importance à la somme obtenue sur deux dés, 7 par exemple, plutôt qu'à la question de savoir si c'est la paire (1,6) qui est apparue, ou (2,5), (3,4), (4,3), (5,2) ou plutôt (6,1). Dans le cas du jet d'une pièce, il peut être plus intéressant de connaître le nombre de fois où le côté pile est apparue plutôt que la séquence détaillée des jets pile et face. Ces grandeurs auxquelles on s'intéresse sont en fait des fonctions réelles définies sur l'ensemble fondamental et sont appelées *variables aléatoires*.

Du fait que la valeur d'une variable aléatoire est déterminée par le résultat de l'expérience, il est possible d'attribuer une probabilité aux différentes valeurs que la variable aléatoire peut prendre.

Soient  $\varepsilon$  une expérience aléatoire et  $(\Omega, \mathcal{A}, P)$  un espace probabilisé lié à cette expérience. Dans de nombreuses situations, on associe à chaque résultat  $\omega \in \Omega$  un nombre réel noté  $X(\omega)$ ; on construit ainsi une application  $X : \Omega \rightarrow \mathbb{R}$ . Historiquement,  $\varepsilon$  était un jeu et  $X$  représentait le gain du joueur.

**Exemple 1.1** Un joueur lance un dé équilibré à 6 faces numérotées de 1 à 6, et on observe le numéro obtenu.

- Si le joueur obtient 1, 3 ou 5, il gagne 1 euro.
- S'il obtient 2 ou 4, il gagne 5 euros.
- S'il obtient 6, il perd 10 euros.

Selon l'expérience aléatoire (lancer d'un dé équilibré) l'ensemble fondamental est  $\Omega = \{1, 2, 3, 4, 5, 6\}$ ,  $\mathcal{A} = \mathcal{P}(\Omega)$  et  $P$  l'équiprobabilité sur  $(\Omega, \mathcal{A})$ . Soit  $X$  l'application de  $\Omega$  dans  $\mathbb{R}$  qui à tout  $\omega \in \Omega$  associe le gain correspondant. On a donc

- $X(1) = X(3) = X(5) = 1$
- $X(2) = X(4) = 5$
- $X(6) = -10$

On dit que  $X$  est une **variable aléatoire** sur  $\Omega$ .

On peut s'intéresser à la probabilité de gagner 1 euro, c'est-à-dire d'avoir  $X(\omega) = 1$ , ce qui se réalise si et seulement si  $\omega \in \{1, 3, 5\}$ . La probabilité cherchée est donc égale à  $P(\{1, 3, 5\}) = 1/2$ . On écrira aussi  $P(X = 1) = 1/2$ .

On pourra donc considérer l'événement :

$$\{X = 1\} = \{\omega \in \Omega / X(\omega) = 1\} = \{\omega \in \Omega / X(\omega) \in \{1\}\} = X^{-1}(\{1\}) = \{1, 3, 5\}.$$

On aura du même  $P(X = 5) = 1/3$  et  $P(X = -10) = 1/6$ . Ce que l'on peut présenter dans un tableau

$x_i$	-10	1	5
$p_i = P(X = x_i)$	1/6	1/2	1/3

Cela revient à considérer un nouvel ensemble d'événements élémentaires:

$$\Omega_X = X(\Omega) = \{-10, 1, 5\}$$

et à munir cet ensemble de la probabilité  $P_X$  définie par le tableau des  $P(X = x_i)$  ci dessus. Cette nouvelle probabilité s'appelle **loi de la variable aléatoire**  $X$ .

Remarquer que

$$P\left(\bigcup_{x_i \in \Omega_X} \{X = x_i\}\right) = \sum_{x_i \in \Omega_X} P(X = x_i) = 1$$

□

Dans ce chapitre, nous traitons le cas où  $X(\Omega)$  est dénombrable. La variable aléatoire est alors dite **discrète**. Sa loi de probabilité, qui peut être toujours définie par sa fonction de répartition, le

sera plutôt par les probabilités individuelles. Nous définirons les deux caractéristiques numériques principales d'une variable aléatoire discrète, l'espérance caractéristique de valeur centrale, et la variance, caractéristique de dispersion. Nous définirons aussi les couples de variables aléatoires.

## 2 Variables aléatoires discrètes

### 2.1 Définition, loi de probabilité

**Définition 2.1** On dit qu'une variable aléatoire réelle (v.a.r.)  $X$  est **discrète** (v.a.r.d.) si l'ensemble des valeurs que prend  $X$  est fini ou infini dénombrable.

Si on suppose  $X(\Omega)$  l'ensemble des valeurs de  $X$  qui admet un plus petit élément  $x_1$ . Alors la v.a.r.d.  $X$  est entièrement définie par:

- L'ensemble  $X(\Omega)$  des valeurs prises par  $X$ , rangées par ordre croissant:  $X(\Omega) = \{x_1, x_2, \dots, x_i, \dots\}$  avec  $x_1 \leq x_2 \leq \dots \leq x_i \leq \dots$
- La **loi de probabilité** définie sur  $X(\Omega)$  par

$$p_i = P(X = x_i) \quad \forall i = 1, 2, \dots$$

**Remarques:**

- Soit  $B$  un ensemble de  $\mathbb{R}$ ,

$$P(X \in B) = \sum_{i/x_i \in B} p(x_i)$$

- En particulier

$$P(a < X \leq b) = \sum_{i/a < x_i \leq b} p(x_i)$$

- Bien sûr tous les  $p(x_i)$  sont positives et  $\sum_{i=1}^{\infty} p(x_i) = 1$ .
- Si  $X$  ne prend qu'un petit nombre de valeurs, cette loi est généralement présentée dans un tableau (cf. l'exemple 1.1).

### 2.2 Fonction de répartition d'une variable aléatoire discrète

**Définition 2.2** On appelle fonction de répartition de la v.a.  $X$ , qu'on note  $F(a)$  de la v.a.r.d.  $X$ , ou  $F_X(a)$ , la fonction définie pour tout réel  $a$ ,  $-\infty < a < \infty$ , par

$$F(a) = P(X \leq a) = \sum_{i/x_i \leq a} P(X = x_i)$$

Cette valeur représente la probabilité de toutes les réalisations inférieures ou égales au réel  $a$ .

**Propriétés:** Voici quelques propriétés de cette fonction:

1. C'est une fonction en escalier (constante par morceaux).
2.  $F(a) \leq 1$  car c'est une probabilité.
3.  $F(a)$  est continue à droite.

$$4. \lim_{a \rightarrow -\infty} F(a) = 0 \text{ et } \lim_{a \rightarrow \infty} F(a) = 1$$

La fonction de répartition caractérise la loi de  $X$ , autrement dit:  $F_X = F_Y$  si et seulement si les variables aléatoires  $X$  et  $Y$  ont la même loi de probabilité.

### Fonction de répartition et probabilités sur $X$

Tous les calculs de probabilité concernant  $X$  peuvent être traités en termes de fonction de répartition. Par exemple,

$$P(a < X \leq b) = F(b) - F(a) \quad \text{pour tout } a < b$$

On peut mieux s'en rendre compte en écrivant  $\{X \leq b\}$  comme union des deux événements incompatibles  $\{X \leq a\}$  et  $\{a < X \leq b\}$ , soit

$$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\}$$

et ainsi

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

ce qui établit l'égalité ci dessus.

**Remarque** On peut déduire de  $F$  les probabilités individuelles par:

$$p_i = F(x_i) - F(x_{i-1}) \quad \text{pour } 1 \leq i \leq n$$

**Exemple 2.1** On joue trois fois à pile ou face. Soit  $X$  la variable aléatoire "nombre de pile obtenus". Ici  $\Omega = \{P, F\}^3$ , et donc

$$X(\Omega) = \{0, 1, 2, 3\}.$$

On a  $\text{card}(\Omega) = 2^3 = 8$ . Calculons par exemple  $P(X = 1)$ , c'est à dire la probabilité d'avoir exactement une pile.

$$X^{-1}(1) = \{(P, F, F), (F, P, F), (F, F, P)\}$$

D'où  $P(X = 1) = \frac{3}{8}$ .

En procédant de la même façon, on obtient la loi de probabilité de  $X$ :

$k$	0	1	2	3
$P(X = k)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

La fonction de répartition de  $X$  est donc donnée par:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/8 & \text{si } 0 \leq x < 1 \\ 1/2 & \text{si } 1 \leq x < 2 \\ 7/8 & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}$$

Le graphe de cette dernière est représentée dans la figure 1.

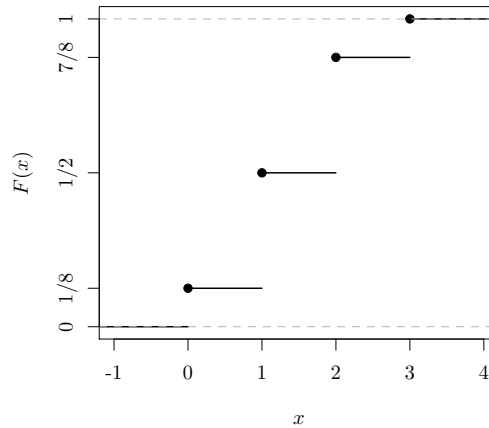


Figure 1: Fonction de répartition

**Exemple 2.2** Soit  $A$  un événement quelconque. On appelle variable aléatoire *indicatrice* de cet événement  $A$ , la variable aléatoire définie par:

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \in \bar{A} \end{cases}$$

et notée  $X = \mathbb{1}_A$ . Ainsi:

$$\begin{aligned} P(X = 1) &= P(A) = p \\ P(X = 0) &= P(\bar{A}) = 1 - p \end{aligned}$$

La fonction de répartition de  $X$  est donc donnée par:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

On peut prendre par exemple le cas d'un tirage d'une boule dans une urne contenant 2 boules blanches et 3 boules noires. Soit  $A$  : "obtenir une boule blanche" et  $X$  la variable indicatrice de  $A$ . La loi de probabilité de  $X$  est alors

$k$	0	1
$P(X = k)$	$\frac{3}{5}$	$\frac{2}{5}$

et sa fonction de répartition est:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 3/5 & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

### 3 Moments d'une variable aléatoire discrète

#### 3.1 Espérance mathématique

**Définition 3.1** Pour une variable aléatoire discrète  $X$  de loi de probabilité  $p(\cdot)$ , on définit l'espérance de  $X$ , notée  $E(X)$ , par l'expression

$$E(X) = \sum_{i \in \mathbb{N}} x_i p(x_i)$$

En termes concrets, l'espérance de  $X$  est la moyenne pondérée des valeurs que  $X$  peut prendre, les poids étant les probabilités que ces valeurs soient prises.

**Exemple 3.1** Reprenons l'exemple 2.1 où on joue 3 fois à pile ou face. L'espérance de  $X$  = "nombre de pile obtenus" est égal à:

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = 1.5$$

**Exemple 3.2** Pour la variable aléatoire indicatrice de l'exemple 2.2:

$$E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = P(A) = p$$

Ceci signifie que l'espérance de la variable indicatrice pour l'événement  $A$  est égale à la probabilité que  $A$  se produise.

**Exemple 3.3** Dans le cas de la loi uniforme sur  $X(\Omega) = \{x_1, \dots, x_k\}$ , c'est à dire avec équiprobabilité de toutes les valeurs  $p_i = 1/k$ , on obtient:

$$E(X) = \frac{1}{k} \sum_{i=1}^k x_i$$

et dans ce cas  $E(X)$  se confond avec la moyenne arithmétique simple  $\bar{x}$  des valeurs possibles de  $X$ .

Pour le jet d'un dé équilibré par exemple:

$$E(X) = \frac{1}{6} \sum_{i=1}^6 i = \frac{7}{2} = 3.5$$

## Linéarité de l'espérance

Propriétés de l'espérance:

1.  $E(X + a) = E(X) + a$ ,  $a \in \mathbb{R}$   
résultat qui se déduit de:

$$\sum_i p_i(x_i + a) = \sum_i p_i x_i + \sum_i a p_i = \sum_i p_i x_i + a \sum_i p_i = \sum_i p_i x_i + a$$

2.  $E(aX) = aE(X)$ ,  $a \in \mathbb{R}$   
il suffit d'écrire:

$$\sum_i p_i a x_i = a \sum_i p_i x_i$$

3.  $E(X + Y) = E(X) + E(Y)$ ,  $X$  et  $Y$  étant deux variables aléatoire.

On peut résumer ces trois propriétés en disant que l'espérance mathématique est linéaire:

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y), \quad \forall \lambda \in \mathbb{R}, \forall \mu \in \mathbb{R}.$$

### 3.2 Espérance d'une fonction d'une variable aléatoire

**Théorème 1** Si  $X$  est une variable aléatoire discrète pouvant prendre ses valeurs parmi les valeurs  $x_i$ ,  $i \geq 1$ , avec des probabilités respectives  $p(x_i)$ , alors pour toute fonction réelle  $g$  on a

$$E(g(X)) = \sum_i g(x_i)p(x_i)$$

**Exemple 3.4** Soit  $X$  une variable aléatoire qui prend une des trois valeurs  $\{-1, 0, 1\}$  avec les probabilités respectives

$$P(X = -1) = 0.2 \quad P(X = 0) = 0.5 \quad P(X = 1) = 0.3$$

Calculer  $E(X^2)$ .

#### Solution

Première approche: Soit  $Y = X^2$ . La distribution de  $Y$  est donnée par

$$P(Y = 1) = P(X = -1) + P(X = 1) = 0.5$$

$$P(Y = 0) = P(X = 0) = 0.5$$

Donc

$$E(X^2) = E(Y) = 1(0.5) + 0(0.5) = 0.5$$

Deuxième approche: En utilisant le théorème

$$\begin{aligned} E(X^2) &= (-1)^2(0.2) + 0^2(0.5) + 1^2(0.3) \\ &= 1(0.2 + 0.3) + 0(0.5) = 0.5 \end{aligned}$$

**Remarque** Remarquer que

$$0.5 = E(X^2) \neq (E(X))^2 = 0.01$$

### 3.3 Variance

**Définition 3.2** La variance est un indicateur mesurant la dispersion des valeurs  $x_i$  que peut prendre la v.a.  $X$  et son espérance  $E(X)$ . On appelle **variance** de  $X$ , que l'on note  $V(X)$ , la quantité

$$V(X) = E[(X - E(X))^2]$$

lorsque cette quantité existe.

C'est l'espérance mathématique du carré de la v.a. centrée  $X - E(X)$ .

On peut établir une autre formule pour le calcul de  $V(X)$ :

$$V(X) = E(X^2) - E^2(X)$$

Or:

$$\begin{aligned} V(X) &= E[X^2 - 2XE(X) + E^2(X)] \\ &= E(X^2) - E[2XE(X)] + E[E^2(X)] \\ &= E(X^2) - 2E^2(X) + E^2(X) \\ &= E(X^2) - E^2(X) \end{aligned}$$

**Exemple 3.5** On cherche  $V(X)$  où  $X$  est le nombre obtenu lors du jet d'un dé équilibré. On a vu dans l'exemple 3.3 que  $E(X) = \frac{7}{2}$ . De plus,

$$\begin{aligned} E(X^2) &= 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right) \\ &= \left(\frac{1}{6}\right)(91) = \frac{91}{6}. \end{aligned}$$

Et donc

$$V(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

### Propriétés de la variance

1.  $V(X) \geq 0$
2.  $V(X + a) = V(X)$   
en effet:

$$\begin{aligned} V(X + a) &= E[(X + a - E(X + a))^2] \\ &= E[(X + a - E(X) - a)^2] \\ &= E[(X - E(X))^2] \\ &= V(X). \end{aligned}$$

3.  $V(aX) = a^2V(X)$   
en effet:

$$\begin{aligned} V(aX) &= E[(aX - E(aX))^2] \\ &= E[(aX - aE(X))^2] \\ &= E[a^2(X - E(X))^2] \\ &= a^2[E(X - E(X))^2] \\ &= a^2V(X). \end{aligned}$$

### 3.4 Ecart-type

**Définition 3.3** La racine carrée de  $V(X)$  est appelée l'**écart-type** de  $X$ , qui se note  $\sigma_X$ . On a

$$\sigma_X = \sqrt{V(X)}$$



$\sigma_X$  s'exprime dans les mêmes unités de mesure que la variable aléatoire  $X$ .

<sup>a</sup>ou  $\sigma(X)$ , ou tout simplement  $\sigma$ .

A noter:

- L'écart type sert à mesurer la dispersion d'un ensemble de données.
- Plus il est faible, plus les valeurs sont regroupées autour de la moyenne.
- Exemple: La répartition des notes d'une classe. Plus l'écart type est faible, plus la classe est homogène.
- L'espérance et l'écart-type sont reliés par l'*inégalité de Bienaymé-Tchebychev*.

### 3.5 Inégalité de Bienaymé-Tchebychev

**Théorème 2** Soit  $X$  une variable aléatoire d'espérance  $\mu$  et de variance  $\sigma^2$ . Pour tout  $\varepsilon > 0$ , on a l'inégalité suivante:

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

**Remarque** On peut l'écrire autrement. Soit  $k = \varepsilon/\sigma$ .

$$P(|X - E(X)| \geq k\sigma) \leq \frac{1}{k^2}$$

**Importance:** Cette inégalité relie la probabilité pour  $X$  de s'écarter de sa moyenne  $E(X)$ , à sa variance qui est justement un indicateur de dispersion autour de la moyenne de la loi. Elle montre quantitativement que "plus l'écart type est faible, plus la probabilité de s'écarter de la moyenne est faible".

**Théorème 3** Inégalité de Markov Soit  $X$  une variable aléatoire à valeur non négatives. Pour tout réel  $a > 0$

$$P(X > a) \leq \frac{E(X)}{a}$$

### 3.6 Moments non centrés et centrés

On appelle moment non centré d'ordre  $r \in \mathbb{N}^*$  de  $X$  la quantité, lorsqu'elle existe:

$$m_r(X) = \sum_{i \in \mathbb{N}} x_i^r p(x_i) = E(X^r).$$

Le moment centré d'ordre  $r \in \mathbb{N}^*$  est la quantité, lorsqu'elle existe:

$$\mu_r(X) = \sum_{i \in \mathbb{N}} p_i [x_i - E(X)]^r = E[X - E(X)]^r.$$

Les premiers moments sont:

$$\begin{aligned} m_1(X) &= E(X), & \mu_1(X) &= 0 \\ \mu_2(X) &= V(X) = m_2(X) - m_1^2(X) \end{aligned}$$

## 4 Couple de variables aléatoires discrètes

Considérons deux variables aléatoires discrètes  $X$  et  $Y$ . Il nous faut pour modéliser le problème une fonction qui nous donne la probabilité que  $(X = x_i)$  en même temps que  $(Y = y_j)$ . C'est la loi de probabilité conjointe.

**Définition 4.1** Soit  $X$  et  $Y$  deux variables aléatoires réelles discrètes, définies sur un espace probabilisé  $(\Omega, \mathcal{A}, P)$  et que

$$\begin{aligned} X(\Omega) &= \{x_1, x_2, \dots, x_l\} \\ Y(\Omega) &= \{y_1, y_2, \dots, y_k\} \\ &\quad (l \text{ et } k \in \mathbb{N}) \end{aligned}$$

La loi du couple  $(X, Y)$ , dite **loi de probabilité conjointe ou simultanée**, est entièrement définie par les probabilités:

$$p_{ij} = P(X = x_i; Y = y_j) = P(\{X = x_i\} \cap \{Y = y_j\})$$

On a

$$p_{ij} \geq 0 \quad \text{et} \quad \sum_{i=1}^l \sum_{j=1}^k p_{ij} = 1$$

Le couple  $(X, Y)$  s'appelle variable aléatoire à deux dimensions et peut prendre  $l \times k$  valeurs.

### 4.1 Table de probabilité conjointe

Les probabilités  $p_{ij}$  peuvent être présentées dans un tableau à deux dimensions qu'on appelle table de probabilité conjointe:

Table 1: Table de probabilité conjointe

$X \backslash Y$	$y_1$	$y_2$	$\dots$	$y_j$	$\dots$	$y_k$
$x_1$	$p_{11}$	$p_{12}$		$p_{1j}$		$p_{1k}$
$x_2$	$p_{21}$	$p_{22}$		$p_{2j}$		$p_{2k}$
$\vdots$						
$x_i$	$p_{i1}$	$p_{i2}$		$p_{ij}$		$p_{ik}$
$\vdots$						
$x_l$	$p_{l1}$	$p_{l2}$		$p_{lj}$		$p_{lk}$

A la première ligne figure l'ensemble des valeurs de  $Y$  et à la première colonne figure l'ensemble des valeurs de  $X$ . La probabilité  $p_{ij} = P(X = x_i; Y = y_j)$  est à l'intersection de la  $i^e$  et de la  $j^e$  colonne.

## 4.2 Lois marginales

Lorsqu'on connaît la loi conjointe des variables aléatoires  $X$  et  $Y$ , on peut aussi s'intéresser à la loi de probabilité de  $X$  seule et de  $Y$  seule. Ce sont les lois de probabilité marginales.

- Loi marginale de  $X$ :

$$p_{i.} = P(X = x_i) = P[\{X = x_i\} \cap \Omega] = \sum_{j=1}^k p_{ij} \quad \forall i = 1, 2, \dots, l$$

- Loi marginale de  $Y$ :

$$p_{.j} = P(Y = y_j) = P[\Omega \cap \{Y = y_j\}] = \sum_{i=1}^l p_{ij} \quad \forall j = 1, 2, \dots, k$$

Table 2: Table de probabilité conjointe avec les lois marginales

$X \backslash Y$	$y_1$	$y_2$	$\dots$	$y_j$	$\dots$	$y_k$	Marginale de $X$
$x_1$	$p_{11}$	$p_{12}$		$p_{1j}$		$p_{1k}$	$p_{1.}$
$x_2$	$p_{21}$	$p_{22}$		$p_{2j}$		$p_{2k}$	$p_{2.}$
$\vdots$							
$x_i$	$p_{i1}$	$p_{i2}$		$p_{ij}$		$p_{ik}$	$p_{i.}$
$\vdots$							
$x_l$	$p_{l1}$	$p_{l2}$		$p_{lj}$		$p_{lk}$	$p_{l.}$
Marginale de $Y$	$p_{.1}$	$p_{.2}$		$p_{.j}$		$p_{.k}$	1

On peut calculer les lois marginales directement depuis la table de la loi conjointe (cf Table 2). La loi marginale de  $X$  est calculée en faisant les totaux par ligne, tandis que celle de  $Y$  l'est en faisant les totaux par colonne.

**Remarque** C'est le fait que les lois de  $X$  et  $Y$  individuellement puissent être lues dans les marges du tableau qui leur vaut leur nom de lois marginales.

**Exemple 4.1** On tire au hasard 3 boules d'une urne contenant 3 boules rouges, 4 blanches et 5 noires.  $X$  et  $Y$  désignent respectivement le nombre de boules rouges et celui de boules blanches tirées. Déterminer la loi de probabilité conjointe du couple  $(X, Y)$  ainsi que les lois marginales de  $X$  et de  $Y$ .

## 4.3 Loi conditionnelles

Pour chaque valeur  $y_j$  de  $Y$  telle que  $p_{.j} = P(Y = y_j) \neq 0$  on peut définir la loi conditionnelle de  $X$  sachant  $Y = y_j$  par

$$p_{i/j} = P(X = x_i / Y = y_j) = \frac{P(X = x_i; Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{.j}} \quad \forall i = 1, 2, \dots, l$$

De même on définit la loi de  $Y$  sachant  $X = x_i$  par

$$p_{j/i} = P(Y = y_j / X = x_i) = \frac{P(X = x_i; Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i.}} \quad \forall j = 1, 2, \dots, k$$

## 4.4 Indépendance de variables aléatoires

On dit que deux v.a.r.d sont indépendantes si et seulement si

$$P(X = x_i; Y = y_j) = P(X = x_i)P(Y = y_j) \quad \forall i = 1, 2, \dots, l \text{ et } j = 1, 2, \dots, k$$

On montre que

$$P(\{X \in A\} \cap \{Y \in B\}) = P(\{X \in A\})P(\{Y \in B\}) \quad \forall A \text{ et } B \in \mathcal{A}$$

### Propriétés

Soit deux v.a.r.d.  $X$  et  $Y$ ,

1.  $E(X + Y) = E(X) + E(Y)$
2. Si  $X$  et  $Y$  sont indépendantes alors  $E(XY) = E(X)E(Y)$ . Mais la réciproque n'est pas toujours vraie.

## 4.5 Covariance

**Définition 4.2** Soit  $X$  et  $Y$  deux v.a.r.d. On appelle **covariance** de  $X$  et de  $Y$  la valeur si elle existe de

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))] = \sum_i \sum_j (x_i - E(X))(y_j - E(Y))p_{ij}$$

qu'on peut calculer en utilisant la formule suivante

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

### Propriétés

- $Cov(X, Y) = Cov(Y, X)$
- $Cov(aX_1 + bX_2, Y) = aCov(X_1, Y) + bCov(X_2, Y)$
- $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$
- Si  $X$  et  $Y$  sont indépendantes alors
  - $Cov(X, Y) = 0$  (la réciproque n'est pas vraie)
  - $V(X + Y) = V(X) + V(Y)$  (la réciproque n'est pas vraie)

## 4.6 Coefficient de corrélation linéaire

**Définition 4.3** On appelle coefficient de corrélation linéaire de  $X$  et de  $Y$  la valeur définie par

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

On peut montrer que

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

**Remarque** Pour le montrer on peut partir du fait que la variance est toujours positive ou nulle. Donc  $V(\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}) \geq 0$  et  $V(\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}) \geq 0$ .

### Interprétation de $\rho$

- Le coefficient de corrélation est une mesure du degré de linéarité entre  $X$  et  $Y$ .
- Les valeurs de  $\rho$  proches de 1 ou  $-1$  indiquent une linéarité quasiment rigoureuse entre  $X$  et  $Y$ .
- Les valeurs de  $\rho$  proche de 0 indiquent une absence de toute relation linéaire.
- Lorsque  $\rho(X, Y)$  est positif,  $Y$  a tendance à augmenter si  $X$  en fait autant.
- Lorsque  $\rho(X, Y) < 0$ ,  $Y$  a tendance à diminuer si  $X$  augmente.
- Si  $\rho(X, Y) = 0$ , on dit que ces deux statistiques sont non corrélées.

■