# Cours 6 : Variables aléatoires

# Essentiel du cours 5

## Définitions et algorithmes

- Dichotomie :
  On coupe et on garde un demi-intervalle qui contient une racine, et on itère.
- Méthode du point fixe :
  On itère $x^{(h)} = g(x^{(h-1)})$
- Méthode de Newton
- Extremum d'une fonction (Newton) :

  Initialisation : Choisir $x^{(0)}$ et une précision $\epsilon$.
  Itération $h$ : Calculer

  $$x^{(h)} = x^{(h-1)} - \frac{f'(x^{(h-1)})}{f''(x^{(h-1)})}$$

  Arrêt : Quand $|x^{(h)} - x^{(h-1)}| < \epsilon$.

  Si $f''(x^*) < 0$ : maximum.
  Si $f''(x^*) > 0$ : minimum.
- Méthode de Newton par coordonnée.

## Définitions et algorithmes

- La dérivée partielle $\partial_{u_j} f$ de la fonction $f$ de $\mathbb{R}^d$ dans $\mathbb{R}$ au point $\mathbf{u} = (u_1, u_2, \dots u_d)$ :
  - on fixe les $u_k$ $(k \neq j)$ ;
  - on dérive par rapport à $u_j$.
- Le vecteur gradient de $f$ en $\mathbf{u}$ :

  $$\nabla f(\mathbf{u}) = [\partial_{u_1} f(\mathbf{u})\ \partial_{u_2} f(\mathbf{u})\ \dots\ \partial_{u_2} f(\mathbf{u})]^\intercal .$$

- Minimisation par descente de gradient

  Initialisation : Choisir un point $\mathbf{u}^{(0)}$, un pas $\gamma$ et une précision $\epsilon$.
  Itération $h$ : Calculer le vecteur gradient au point $\mathbf{u}^{(h-1)}$ : $\nabla f(\mathbf{u}^{(h-1)})$ et itérer

  $$\mathbf{u}^{(h)} = \mathbf{u}^{(h-1)} - \gamma \nabla f(\mathbf{u}^{(h-1)})$$

  Arrêt : Quand $\|\mathbf{u}^{(h)} - \mathbf{u}^{(h-1)}\| < \epsilon$.

# Notations

**Objectif.** *Inférence* = généralisation à d'autres données que celles qui sont observées
**Cadre classique.** Considérer les observations $x_1, x_2, \dots x_n$ comme des réalisations de variables aléatoires $X_1, X_2, \dots X_n$.

## Notations

- $X$ : variable aléatoire (majuscule)
- $x =$ réalisation de $X$ (minuscule)
- $\mathcal{X} =$ ensemble des valeurs possibles de $X$ :
- $p =$ loi de probabilité de $X$ : $p(x) = \mathbb{P}\{X = x\}$

## Exemple : pile ou face

$$X = \text{« pile ou face »}, \quad x = \text{« pile »}, \quad \mathcal{X} = \{\text{pile}, \text{face}\}, \quad p(\text{pile}) = p(\text{face}) = 1/2.$$

**Menu du jour.**

- Définition de notions analogue à celles vues dans le cadre descriptif (espérance, variance, covariance, loi jointe, marginale, conditionnelle, ...)
- Introduction de la notion de dépendance

On se concentre pour l'instant aux variables discrètes, i.e. $\mathcal{X}$ est fini ($\mathcal{X} = \{0, 1\}$) ou dénombrable ($\mathcal{X} = \mathbb{N}$)

**Définition 1 (Loi de probabilité discrète)**

La loi de probabilité d'une variable aléatoire $X$ à valeur dans $\mathcal{X}$ fini ou dénombrable est définie par la fonction

$$\begin{array}{rccl} p: & \mathcal{X} & \mapsto & [0,1] \\ & x & \to & p(x) = \mathbb{P}\{X = x\}. \end{array}$$

Par définition, $p$ satisfait l'équation

$$\sum_{x \in \mathcal{X}} p(x) = 1.$$

**Définition 2 (Loi de Bernoulli)**

La loi de Bernoulli (de paramètre $\pi$) est la loi de la variable aléatoire $X$ binaire à valeur dans $\mathcal{X} = \{0, 1\}$ telle que

$$p(0) = \mathbb{P}\{X = 0\} = 1 - \pi, \qquad p(1) = \mathbb{P}\{X = 1\} = \pi.$$

On note alors

$$X \sim \mathcal{B}(\pi).$$

**Exemple : pile ou face**

Le tirage d'une pièce (non faussée) à pile face suit une loi de Bernoulli

$$\mathcal{B}(1/2),$$

en comptant 0 pour *pile* et 1 pour *face*.

# Loi binomiale

## Définition 3 (Loi binomiale)

La loi binomiale (de paramètres $n$ et $\pi$) est la loi de la variable aléatoire $X$ à valeur dans $\mathcal{X} = \{0, \ldots, n\}$ telle que

$$p(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}.$$

On note alors

$$X \sim \mathcal{B}(n, \pi).$$

## Exemple : nombre de « pile »

Le nombre de *face* obtenu au cours 10 tirages *indépendants* d'une pièce (non faussée) à *pile* ou *face* suit une loi binomiale

$$\mathcal{B}(10, 1/2).$$

(définition de l'indépendance à suivre)

**Définition 4 (Fonction de répartition)**

Si la variable $X$ est à valeurs entières (ou à défaut, si l'ensemble $\mathcal{X}$ est ordonnée), la fonction de répartition de la variable $X$ est la fonction

$$\begin{array}{rccl} F: & \mathcal{X} & \mapsto & [0,1] \\ & x & \to & F(x) = \mathbb{P}\{X \leqslant x\}. \end{array}$$

**Proposition 1 (Fonction de répartition d'une variable discrète)**

*La fonction de répartition de répartition de la variable aléatoire $X$ de probabilité $p$ vaut*

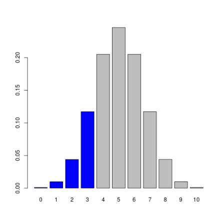
$$F(x) = \mathbb{P}\{X \leqslant x\} = \sum_{y \leqslant x} p(y).$$

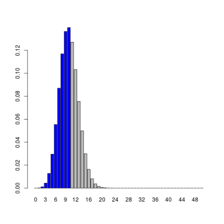
## Fonction de répartition

## Loi binomiale

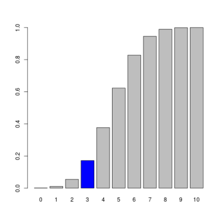
$X \sim \mathcal{B}(10, 1/2)$ $\quad$ $X \sim \mathcal{B}(50, 1/5)$

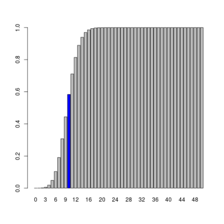
$p(x) = \mathbb{P}\{X = x\}$




$F(x) = \mathbb{P}\{X \leqslant x\}$




$F(3) \simeq 0.1719$ $\quad$ $F(10) \simeq 0.5836$

# Espérance

**Définition 5 (Espérance)**

L'espérance (notée $\mathbb{E}$) d'une variable $X$ est définie par $\mathbb{E}(X) = \sum_{x \in \mathcal{X}} p(x)x$.

**Remarques.**

1. On peut généraliser la notion d'espérance à toute fonction (à valeur réelle) de $X$ :

$$\mathbb{E}\left(f(X)\right) = \sum_{x \in \mathcal{X}} p(x)f(x).$$

   On a ainsi

$$\begin{aligned}
\mathbb{E}\left(X^2\right) &= \sum_{x \in \mathcal{X}} p(x)x^2, && \\
\mathbb{E}\left(\log(X)\right) &= \sum_{x \in \mathcal{X}} p(x)\log(x) && \text{pour } \mathcal{X} \subset \mathbb{R}^{*+}, \\
\mathbb{E}(1/X) &= \sum_{x \in \mathcal{X}} p(x)/x && \text{pour } \mathcal{X} \subset \mathbb{R}^{*}.
\end{aligned}$$

2. Quand $\mathcal{X}$ est infini dénombrable ($\mathcal{X} = \mathbb{N}$), il faut s'assurer que les séries correspondantes sont bien définies (hors programme du cours).

### Définition 6 (Variance)

La variance (notée $\mathbb{V}$) de $X$ est l'espérance du carré de l'écart à son espérance :

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \sum_{x \in \mathcal{X}} p(x)\,(x - \mathbb{E}(X))^2 \,.$$

L'écart-type de $X$ est la racine carré de sa variance $\sqrt{\mathbb{V}(X)}$.

### Proposition 2 (Formule alternative de la variance)

*La variance d'une variable aléatoire $X$ entière vérifie*

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

### Proposition 3 (Variable de Bernoulli)

*Si $X \sim \mathcal{B}(\pi)$, alors*

$$\mathbb{E}(X) = \pi, \qquad \mathbb{E}(X^2) = \pi, \qquad \mathbb{V}(X) = \pi(1 - \pi).$$

# Variance

# Variance

**Proposition 4 (Transformation linéaire)**

$$\mathbb{E}(a + bX) = a + b\,\mathbb{E}(X), \qquad \mathbb{V}(a + bX) = b^2\,\mathbb{V}(X).$$

**Remarque importante.** La propriété précédente ne traite que des transformations *linéaires*. En général :

$$\mathbb{E}\left(\frac{1}{X}\right) \neq \frac{1}{\mathbb{E}(X)}, \qquad \mathbb{E}(X^2) \neq [\mathbb{E}(X)]^2\,.$$

# Propriétés de l’espérance et de la variance

# Couple de variables aléatoires discrètes

**Objectifs.** Etudier deux variables aléatoires $X$ et $Y$ à valeurs dans $\mathcal{X}$ et $\mathcal{Y}$ : $(X, Y)$ prend ses valeurs dans $\mathcal{X} \times \mathcal{Y}$.
**Notation.** On indice la loi de probabilité $p$ par la variable à laquelle elle se réfère : $p_X$ pour la loi de $X$, $p_Y$ pour $Y$, $p_{XY}$ pour $(X, Y)$.

## Définition 7 (Loi jointe)

La loi jointe de deux variables aléatoires discrètes $X$ et $Y$ à valeur respectivement dans $\mathcal{X}$ et $\mathcal{Y}$ est définie par la fonction

$$\begin{array}{rccl} p_{XY} : & \mathcal{X} \times \mathcal{Y} & \mapsto & [0, 1] \\ & (x, y) & \rightarrow & p_{XY}(x, y) = \mathbb{P}\{X = x, Y = y\}. \end{array}$$

## Exemple de loi jointe de deux variables discrètes

$X$ à valeur dans $\mathcal{X} = \{0, 1\}$, $Y$ à valeur dans $\mathcal{Y} = \{1, 2, 3\}$.

$$\mathcal{X} \times \mathcal{Y} = \{(0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3)\}$$

| $p_{XY}(x, y)$ | 1 | 2 | 3 |
|---|---|---|---|
| 0 | 1/4 | 1/4 | 0 |
| 1 | 1/6 | 1/6 | 1/6 |

**Loi marginale** de $X$ = loi de probabilité de $X$, *sans tenir compte* de la valeur de $Y$.

**Proposition 5 (Loi marginale)**

*Les lois marginales de $X$ et de $Y$ sont données par*

$$p_X(x) = \mathbb{P}\{X = x\} = \sum_{y \in \mathcal{Y}} p_{XY}(x, y)$$

$$p_Y(y) = \mathbb{P}\{Y = y\} = \sum_{x \in \mathcal{X}} p_{XY}(x, y).$$

# Loi marginale : démonstration

# Loi marginale : exemple

## Exemple de loi jointe

**Loi jointe :**

| $p_{XY}(x,y)$ | 1 | 2 | 3 | $\sum_{y\in\mathcal{Y}} p_{XY}(x,y)$ |
|---|---|---|---|---|
| 0 | 1/4 | 1/4 | 0 | 1/2 |
| 1 | 1/6 | 1/6 | 1/6 | 1/2 |
| $\sum_{x\in\mathcal{X}} p_{XY}(x,y)$ | 5/12 | 5/12 | 1/6 | 1 |

**Lois marginales :**

| $x$ | 0 | 1 |
|---|---|---|
| $p_X(x)$ | 1/2 | 1/2 |

| $y$ | 1 | 2 | 3 |
|---|---|---|---|
| $p_Y(y)$ | 5/12 | 5/12 | 1/6 |

**Espérances :**

$$\mathbb{E}(X) = 1/2, \qquad \mathbb{E}(Y) = 7/4 = 1.75,$$

**Espérances des carrés :**

$$\mathbb{E}(X^2) = 1/2, \qquad \mathbb{E}(Y^2) = 43/12 \simeq 3.583,$$

**Variances :**

$$\mathbb{V}(X) = 1/4, \qquad \mathbb{V}(Y) = 43/12 - (7/4)^2 = 25/48 \simeq 0.521,$$

**Définition 8 (Covariance)**

La covariance entre les variables $X$ et $Y$ est l'espérance du produit des écarts à leurs espérances respectives, soit, en notant $\mu_X = \mathbb{E}(X)$ et $\mu_Y = \mathbb{E}(Y)$ :

$$\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}\left((X - \mu_X)(Y - \mu_Y)\right).$$

**Remarque.**

$$\mathbb{C}\mathrm{ov}(X, X) = \mathbb{V}(X).$$

**Proposition 6 (Formule alternative de la covariance)**

*La covariance des variables $X$ et $Y$ vérifie :*

$$\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y).$$

# Covariance

# Corrélation

**Définition 9 (Corrélation)**

La corrélation entre les variables $X$ et $Y$ est le rapport entre leur covariance et le produit de leur écarts-types :

$$\mathbb{C}\mathrm{or}(X, Y) = \frac{\mathbb{C}\mathrm{ov}(X, Y)}{\sqrt{\mathbb{V}(X)\,\mathbb{V}(Y)}}.$$

Comme pour sa version descriptive $\mathrm{cor}(\mathbf{x}, \mathbf{y})$, on peut montrer que

$$-1 \leqslant \mathbb{C}\mathrm{or}(X, Y) \leqslant +1.$$

# Loi conditionnelle (1/2)

**Loi conditionnelle** de $Y$ = loi de $Y$ *connaissant* la valeur de $X$.
Sachant que $X = x$, l'ensemble des possibles pour le couple $(X, Y)$ n'est plus $\mathcal{X} \times \mathcal{Y}$, mais seulement $\{x\} \times \mathcal{Y}$.

**Définition 10 (Loi conditionnelle)**

La loi conditionnelle de $Y$ *sachant* $\{X = x\}$ est définie par

$$p_{Y|X=x}(y) = \mathbb{P}\{Y = y \mid X = x\} = \frac{p_{XY}(x, y)}{p_X(x)}.$$

**Remarques.**

1. Cas particulier de probabilité conditionnelle de l'événement $A$ sachant l'événement $B$

$$\mathbb{P}\{A \mid B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}$$

avec $A = \{X = x, Y = y\}$ et $B = \{X = x\}$.

2. Symétriquement, la loi conditionnelle de $X$ *sachant* $\{Y = y\}$ vaut

$$p_{X|Y=y}(x) = \mathbb{P}\{X = x \mid Y = y\} = \frac{p_{XY}(x, y)}{p_Y(y)}.$$

# Loi conditionnelle (2/2)

**Proposition 7 (Probabilité conditionnelle totale)**

$$\sum_{y \in \mathcal{Y}} p_{Y|X=x}(y) = 1.$$

**Exemple de loi jointe**

**Loi jointe :**

| $p_{XY}(x,y)$ | 1 | 2 | 3 |
|---|---|---|---|
| 0 | $1/4$ | $1/4$ | 0 |
| 1 | $1/6$ | $1/6$ | $1/6$ |

**Lois marginales :**

| $x$ | 0 | 1 |
|---|---|---|
| $p_X(x)$ | $1/2$ | $1/2$ |

| $y$ | 1 | 2 | 3 |
|---|---|---|---|
| $p_Y(y)$ | $5/12$ | $5/12$ | $1/6$ |

**Loi conditionnelles :**

| $p_{Y\|X=x}(y)$ | 1 | 2 | 3 |
|---|---|---|---|
| $x=0$ | $1/2$ | $1/2$ | 0 |
| $x=1$ | $1/3$ | $1/3$ | $1/3$ |

| $p_{X\|Y=y}(x)$ | 0 | 1 |
|---|---|---|
| $y=1$ | $3/5$ | $2/5$ |
| $y=2$ | $3/5$ | $2/5$ |
| $y=3$ | 0 | 1 |

# Probabilité conditionnelle totale : démonstration

# Tirage dans une loi jointe

**Proposition 8 (Tirage d'une réalisation de $(X, Y)$)**

*On peut tirer une réalisation du couple $(X, Y)$ selon la loi $p_{XY}$ en*

1. *tirant une réalisation $x$ de $X$ selon sa loi marginale $p_X$,*
2. *puis en tirant ensuite une réalisation $y$ de $Y$ selon sa loi conditionnelle $p_{Y|X=x}$,*

**Exemple de loi jointe**

On tire $X$ selon sa loi marginale :

$$\mathbb{P}\{X = 0\} = 1/2, \qquad \mathbb{P}\{X = 1\} = 1/2,$$

puis

- si $X = 0$, on tire $Y$ selon la loi conditionnelle $p_{Y|X=0}$ :

$$\mathbb{P}\{Y = 1 \mid X = 0\} = 1/2, \qquad \mathbb{P}\{Y = 2 \mid X = 0\} = 1/2;$$

- si $X = 1$, on tire $Y$ selon la loi conditionnelle $p_{Y|X=1}$ :

$$\mathbb{P}\{Y = 1 \mid X = 1\} = 1/3, \quad \mathbb{P}\{Y = 2 \mid X = 1\} = 1/3, \quad \mathbb{P}\{Y = 3 \mid X = 1\} = 1/3.$$

# Tirage d'une réalisation de $(X, Y)$ : démonstration

La loi conditionnelle est une loi de probabilité, donc on peut définir son espérance.

**Définition 11 (Espérance conditionnelle)**

L'espérance conditionnelle de $Y$ sachant $X = x$ vaut

$$\mathbb{E}(Y \mid X = x) = \sum_{y \in \mathcal{Y}} p_{Y|X=x}(y) y.$$

## Exemple de loi jointe

**Espérances conditionnelles de $Y$ sachant $X = x$ :**

| $p_{Y\mid X=x}(y)$ | 1 | 2 | 3 | $\mathbb{E}(Y \mid X = x)$ |
|---|---|---|---|---|
| $x = 0$ | $1/2$ | $1/2$ | $0$ | $3/2$ |
| $x = 1$ | $1/3$ | $1/3$ | $1/3$ | $2$ |

**Espérances conditionnelles de $X$ sachant $Y = y$ :**

| $p_{X\mid Y=y}(x)$ | 0 | 1 | $\mathbb{E}(Y \mid X = x)$ |
|---|---|---|---|
| $y = 1$ | $3/5$ | $2/5$ | $2/5$ |
| $y = 2$ | $3/5$ | $2/5$ | $2/5$ |
| $y = 3$ | $0$ | $1$ | $1$ |

**Espérances marginales :**

$$\mathbb{E}(X) = 1/2, \qquad \mathbb{E}(Y) = 7/4 = 1.75.$$

On peut définir de façon analogue l'espérance conditionnelle de toute fonction $g$ de $Y$ :

$$\mathbb{E}(g(Y) \mid X = x) = \sum_{y \in \mathcal{Y}} p_{Y|X=x}(y) g(y).$$

**Définition 12 (Variance conditionnelle)**

En notant $\mu_{Y|X=x} = \mathbb{E}(Y \mid X = x)$, la variance conditionnelle de $Y$ sachant $X = x$ vaut :

$$\mathbb{V}(Y \mid X = x) = \mathbb{E}\left((Y - \mu_{Y|X=x})^2 \mid X = x\right)$$

$$= \mathbb{E}\left(Y^2 \mid X = x\right) - (\mu_{Y|X=x})^2.$$

# Indépendance

**Sens littéral :** les variables $X$ et $Y$ sont indépendantes si la distribution de la valeur de l'une ne dépend pas de la valeur de l'autre.

**Définition 13 (Indépendance)**

Les variables $X$ et $Y$ sont indépendantes si et seulement si pour tout couple $(x, y)$ dans $\mathcal{X} \times \mathcal{Y}$, on a

$$p_{XY}(x, y) = p_X(x)p_Y(y).$$

**Exemple à deux variables $(X, Y)$**

loi jointe $\neq$ loi produit

| $p_{XY}(x,y)$ | 1 | 2 | 3 |
|---|---|---|---|
| 0 | 1/4 | 1/4 | 0 |
| 1 | 1/6 | 1/6 | 1/6 |

| $p_X(x)p_Y(y)$ | 1 | 2 | 3 |
|---|---|---|---|
| 0 | 5/24 | 5/24 | 1/12 |
| 1 | 5/24 | 5/24 | 1/12 |

(notamment : $p_{XY}(0, 3) = 0$ et $p_X(x)p_Y(y) \neq 0$)

$\Rightarrow$ $X$ et $Y$ ne sont pas indépendantes.

**Proposition 9 (Propriétés des variables indépendantes)**

*Si les variables $X$ et $Y$ sont indépendantes, alors elles vérifient :*

1. *pour tout couple $(x, y)$ dans $\mathcal{X} \times \mathcal{Y}$*

$$p_{Y|X=x}(y) = p_Y(y) \qquad et \qquad p_{X|Y=y}(x) = p_X(x);$$

2. *l'espérance de leur produit est égale au produit de leurs espérances :*

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

3. *leur covariance est nulle : $\mathbb{C}\mathrm{ov}(X, Y) = 0$.*

**Remarques :**

1. Intuition : distribution conditionnelle $p_{Y|X=x}$ = distribution marginale $p_Y$.
2. Généralisable à toutes fonctions de $X$ et de $Y$ : $\mathbb{E}\left(f(X)g(Y)\right) = \mathbb{E}\left(f(X)\right)\mathbb{E}\left(g(Y)\right)$.
3. Implication et non équivalence :

$$\begin{aligned} \{X \text{ et } Y \text{ indépendantes}\} \quad &\Rightarrow \quad \mathbb{C}\mathrm{ov}(X, Y) = 0, \\ \mathbb{C}\mathrm{ov}(X, Y) = 0 \quad &\not\Rightarrow \quad \{X \text{ et } Y \text{ indépendantes}\}. \end{aligned}$$

**Proposition 10 (Loi binomiale)**

*La somme $Y$ du $n$ variables de Bernoulli indépendantes $X_1, X_2, \ldots X_n$ de même paramètre $\pi$ suit une loi binomiale de paramètre $n$ et $\pi$.*

**Étapes de la démonstration.**

1. Loi de chaque $X_i$ : $p_X(x) = \pi^x(1-\pi)^{1-x}$.
2. Loi jointe des $X_1, X_2, \ldots X_n$.
3. Loi de leur somme.

**Objectif.** Introduire la notion d'indépendance conditionnelle (hypothèse implcite des classificateurs « bayésiens naïfs »).

### Définition 14 (Loi d'un triplet de variables aléatoires)

La loi jointe du triplet de variables aléatoires $(X, Y, Z)$ à valeurs dans $\mathcal{X}$, $\mathcal{Y}$ et $\mathcal{Z}$ est la fonction

$$\begin{array}{rrcl} p_{XYZ}: & \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} & \mapsto & [0,1] \\ & (x,y,z) & \to & p_{XYZ}(x,y,z) = \mathbb{P}\{X = x, Y = y, Z = z\}. \end{array}$$

### Exemple à trois variables $(X, Y, Z)$

On considère $X$, $Y$ et $Z$ à valeurs dans $\mathcal{X} = \{0,1\}$, $\mathcal{Y} = \{1,2,3\}$ et $\mathcal{Z} = \{1,2\}$ de loi jointe

$z = 1$

| $p_{XYZ}(x,y,1)$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|
| $x=0$ | $1/18$ | $1/9$ | $1/6$ |
| $x=1$ | $1/36$ | $1/18$ | $1/12$ |

$z = 2$

| $p_{XYZ}(x,y,2)$ | $y=1$ | $y=2$ | $y=3$ |
|---|---|---|---|
| $x=0$ | $1/24$ | $1/24$ | $1/24$ |
| $x=1$ | $1/8$ | $1/8$ | $1/8$ |

**Proposition 11 (Lois marginales)**

*Les lois marginales de chaque variable sont*

$$p_X(x) = \sum_{y\in\mathcal{Y}}\sum_{z\in\mathcal{Z}} p_{XYZ}(x,y,z), \quad p_Y(y) = \sum_{x\in\mathcal{X}}\sum_{z\in\mathcal{Z}} p_{XYZ}(x,y,z), \quad p_Z(z) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p_{XYZ}(x,y,z).$$

*Les lois marginales de chaque couple de variables sont*

$$p_{XY}(x,y) = \sum_{z\in\mathcal{Z}} p_{XYZ}(x,y,z), \quad p_{XZ}(x,z) = \sum_{y\in\mathcal{Y}} p_{XYZ}(x,y,z), \quad p_{YZ}(y,z) = \sum_{x\in\mathcal{X}} p_{XYZ}(x,y,z).$$

# Lois marginales : démonstration

## Exemple à trois variables $(X, Y, Z)$

**Lois marginales $p_X$ et $p_Y$ :**

| $x$ | 0 | 1 |
|---|---|---|
| $p_X(x)$ | 11/24 | 13/24 |

| $y$ | 1 | 2 | 3 |
|---|---|---|---|
| $p_Y(y)$ | 1/4 | 1/3 | 5/12 |

**Comparaison entre la loi marginale jointe $p_{XY}$ et loi produit $p_X p_Y$ :**

| $p_{XY}(x, y)$ | $y = 1$ | $y = 2$ | $y = 3$ |
|---|---|---|---|
| $x = 0$ | 7/72 | 11/72 | 5/24 |
| $x = 1$ | 11/72 | 13/72 | 5/24 |

| $p_X(x)p_Y(y)$ | $y = 1$ | $y = 2$ | $y = 3$ |
|---|---|---|---|
| $x = 0$ | 11/96 | 11/72 | 55/288 |
| $x = 1$ | 13/96 | 13/72 | 65/288 |

$\Rightarrow$ les variables $X$ et $Y$ ne sont pas indépendantes.

**Espérances conditionnelles de $Y$ sachant $X = x$ :**

| $p_{XY}(x, y)$ | $y = 1$ | $y = 2$ | $y = 3$ | $\mathbb{E}(Y \mid X = x)$ |
|---|---|---|---|---|
| $x = 0$ | 7/72 | 11/72 | 5/24 | 37/36 |
| $x = 1$ | 11/72 | 13/72 | 5/24 | 41/36 |

**Conclusion :** $Z$ étant inconnue, la connaissance de $X$ apporte une information sur $Y$.

La notion d'indépendance se généralise également à plus de deux variables.

**Définition 15 (Indépendance de trois variables discrètes)**

Les variables $X$, $Y$ et $Z$ sont indépendantes si et seulement si, pour tout triplet $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ :

$$p_{XYZ}(x, y, z) = p_X(x)p_Y(y)p_Z(z).$$

De même, on peut définir la loi jointe des variables $X$ et $Y$ sachant $Z = z$.

**Définition 16 (Loi jointe conditionnelle)**

La loi jointe des variables $X$ et $Y$ sachant $Z = z$ vaut

$$p_{X,Y|Z=z}(x, y) = \mathbb{P}\{X = x, Y = y \mid Z = z\} = \frac{p_{XYZ}(x, y, z)}{p_Z(z)}.$$

## Exemple à trois variables $(X, Y, Z)$

**Lois conditionnelles de $X$ et $Y$ sachant $Z = z$ :**

| $p_{X\|Z=z}(x)$ | $x = 0$ | $x = 1$ |
|---|---|---|
| $z = 1$ | $2/3$ | $1/3$ |
| $z = 2$ | $1/4$ | $3/4$ |

| $p_{Y\|Z=z}(y)$ | $y = 1$ | $y = 2$ | $y = 3$ |
|---|---|---|---|
| $z = 1$ | $1/6$ | $1/3$ | $1/2$ |
| $z = 2$ | $1/3$ | $1/3$ | $1/3$ |

**Lois jointes conditionnelles de $(X, Y)$ sachant $Z = z$ :**

$z = 1$

| $p_{XY\|Z=1}(x, y)$ | $y = 1$ | $y = 2$ | $y = 3$ |
|---|---|---|---|
| $x = 0$ | $1/9$ | $2/9$ | $1/3$ |
| $x = 1$ | $1/18$ | $1/9$ | $1/6$ |

$z = 2$

| $p_{XY\|Z=2}(x, y)$ | $y = 1$ | $y = 2$ | $y = 3$ |
|---|---|---|---|
| $x = 0$ | $1/12$ | $1/12$ | $1/12$ |
| $x = 1$ | $1/4$ | $1/4$ | $1/4$ |

On peut appliquer la notion d'indépendance à une loi jointe conditionnelle.

**Définition 17 (Indépendance conditionnelle)**

1. Les variables $X$ et $Y$ sont indépendante conditionnellement à $Z = z$ si et seulement si, pour tout couple $(x, y)$ de $\mathcal{X} \times \mathcal{Y}$,

$$p_{X,Y|Z=z}(x, y) = p_{X|Z=z}(x)p_{Y|Z=z}(y).$$

2. Les variables $X$ et $Y$ sont indépendante conditionnellement à $Z$ si et seulement si la propriété ci-dessous est vraie pour tout $z$ de $\mathcal{Z}$ :

$$\forall z \in \mathcal{Z}, \qquad p_{X,Y|Z=z}(x, y) = p_{X|Z=z}(x)p_{Y|Z=z}(y).$$

# Indépendance conditionnelle : exemple

## Exemple à trois variables $(X, Y, Z)$

**Comparaison de la loi jointe conditionnelle avec le produit des loi conditionnelles :** On peut vérifier que pour tout $(x, y) \in \mathcal{X} \times \mathcal{Y}$ et pour chaque $z \in \mathcal{Z}$ :

$$p_{X,Y|Z=z}(x, y) = p_{X|Z=z}(x) \times p_{Y|Z=z}(y)$$

$\Rightarrow$ $X$ et $Y$ sont indépendantes conditionnellement à $Z = 1$ et $Z = 2$
$\Rightarrow$ $X$ et $Y$ sont indépendantes conditionnellement à $Z$

**Espérance conditionnelles :**

| | | $p_{XYZ}(x, y, z)$ | | | $p_{Y\|X=x,Z=z}(y)$ | | | |
|---|---|---|---|---|---|---|---|---|
| | | $y = 1$ | $y = 2$ | $y = 3$ | $y = 1$ | $y = 2$ | $y = 3$ | $\mathbb{E}(Y \mid X = x, Z = z)$ |
| $z = 1$ | $x = 0$ | $1/18$ | $1/9$ | $1/6$ | $1/6$ | $1/3$ | $1/2$ | $7/3$ |
| $z = 1$ | $x = 1$ | $1/36$ | $1/18$ | $1/12$ | $1/6$ | $1/3$ | $1/2$ | $7/3$ |
| $z = 2$ | $x = 0$ | $1/24$ | $1/24$ | $1/24$ | $1/3$ | $1/3$ | $1/3$ | $2$ |
| $z = 2$ | $x = 1$ | $1/8$ | $1/8$ | $1/8$ | $1/3$ | $1/3$ | $1/3$ | $2$ |

**Conclusion :** La distribution (et l'espérance) conditionnelle de $Y$ sachant $(X = x, Z = z)$ varient avec $z$ mais pas avec $x$ : si la valeur de $Z$ est connue, la valeur de $X$ n'apporte pas d'information supplémentaire sur $Y$.

# Essentiel du cours 6

## Définitions

- Loi de probabilité : $\sum_{x \in \mathcal{X}} p(x) = 1$
- Espérance : $\mathbb{E}(X) = \sum_{x \in \mathcal{X}} p(x) x$
- Variance :
  $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \sum_{x \in \mathcal{X}} p(x)\left(x - \mathbb{E}(X)\right)^2$
- Covariance :
  $\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right)$
- Corrélation : $\mathbb{C}\mathrm{or}(X, Y) = \dfrac{\mathbb{C}\mathrm{ov}(X, Y)}{\sqrt{\mathbb{V}(X)\,\mathbb{V}(Y)}}$
- Loi conditionnelle :
  $p_{Y|X=x}(y) = \mathbb{P}\{Y = y \mid X = x\} = \dfrac{p_{XY}(x, y)}{p_X(x)}$
- Espérance conditionnelle :
  $\mathbb{E}(Y \mid X = x) = \sum_{y \in \mathcal{Y}} p_{Y|X=x}(y) y$
- Variance conditionnelle : $\mathbb{V}(Y \mid X = x)$
  $= \mathbb{E}\left((Y - \mathbb{E}(Y \mid X = x))^2 \mid X = x\right)$
  $= \mathbb{E}\left(Y^2 \mid X = x\right) - \left(\mathbb{E}(Y \mid X = x)\right)^2$
- $X$ et $Y$ indépendantes si et seulement si :
  $\forall (x, y), p_{XY}(x, y) = p_X(x) p_Y(y)$

## Propositions

- Fonction de répartition :
  $F(x) = \mathbb{P}\{X \leqslant x\} = \sum_{y \leqslant x} p(y)$
- Formule alternative de la variance :
  $\mathbb{V}(X) = \mathbb{E}(X^2) - \left(\mathbb{E}(X)\right)^2$
- Transformation linéaire :
  $\mathbb{E}(a + bX) = a + b\,\mathbb{E}(X)$
  $\mathbb{V}(a + bX) = b^2\,\mathbb{V}(X)$
- Loi marginale : $p_X(x) = \mathbb{P}\{X = x\} = \sum_{y \in \mathcal{Y}} p_{XY}(x, y)$
- Formule alternative de la covariance :
  $\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$
- Probabilité conditionnelle totale :
  $\sum_{y \in \mathcal{Y}} p_{Y|X=x}(y) = 1$
- Propriétés des variables $X$, $Y$ indépendantes :
  $p_{Y|X=x}(y) = p_Y(y)$
  $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$
  $\mathbb{C}\mathrm{ov}(X, Y) = 0$