**Université des Sciences et Technologies de Lille**
**U.F.R. de Mathématiques Pures et Appliquées**

# M206ICP : Introduction au calcul des probabilités

Notes de cours par Clément Boulonne

**L2 Mathématiques** **2007 - 2008**

# Table des matières

# Références

Certaines parties du cours ont été recopiées du polycopié de cours suivant :

1) Ch. Suquet, *Introduction au Calcul des Probabilités*, 2007-2008

Les cours sont téléchargeables sur le site IPEIS (Intégration, Probabilités Elémentaires et Initiation à la Statistique) de l'Université Lille 1.

# Chapitre 1

# Espaces probabilisés

## 1.1 Evenements

**Définition 1.1.1.** On note $\Omega$ un ensemble non vide dont les éléments représentent tous les résultats possibles ou événements élémentaire d'une expérience alétoire. $\Omega = \{\omega\}$, on dira que $\omega$, élément de $\Omega$ est un événement élémentaire.

**Exemple 1.1.1.** On lance une pièce symétrique et on regarde à sa tombée sur le sol, la pièce tombe sur pile ou face. On aura alors :

$$\Omega = \{P, F\}$$

On pioche sur un jeu de 52 cartes, une carte. On aura alors :

$$\Omega = \{A\clubsuit, A\spadesuit, A\diamondsuit, A\heartsuit, ..., K\clubsuit, K\spadesuit, K\diamondsuit, K\heartsuit\}$$

Dans un intervalle $[0, 1]$, on choisit un point au hasard sur le segment :

$$\Omega = \{x \in [0, 1]\}$$

On peut aussi mesurer une température à n'importe quel point du globe :

$$\Omega = \{-273°C, \infty\}$$

**Définition 1.1.2.** $\mathcal{F} = \{A\}$, $A \subset \Omega$ et est appelé événement et $\mathcal{F}$ est la famille des événements possibles.

**Exemple 1.1.2.** Si on jette un dé, l'évément $A$ peut être “obtention d'un chiffre paire” et donc $A = \{2, 4, 6\}$ composé de trois événements élémentaires $\omega_1 = 2, \omega_2 = 4, \omega_3 = 6$.

**Définition 1.1.3.** On note $A^C$, un événement qui se réalise si $A$ ne se réalise pas. On l'appelle événement contraie ou complémentaire de $A$.

**Définition 1.1.4.** $A \cap B$ est la réalisation de $A$ et $B$ en même temps.

**Définition 1.1.5.** $A \cup B$ est la réalisation de $A$ ou la réalisation de $B$.

## 1.2 La probabilité comme fonction d'ensembles

**Propriété 1.2.1.** *$\mathcal{F}$ a les propriétés suivantes :*

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^C \in \mathcal{F}$
3. *$\mathcal{F}$ est stable par opérations de réunion et d'intersection sur les suites d'événements. C'est-à-dire, si $(A_k)$ est un ensemble d'évenement fini et dénombrable alors :*

$$\bigcup_k A_k \in \mathcal{F} \text{ et } \bigcap_k A_k \in \mathcal{F}$$

*Démonstration.* On démontre la propriété 3 sur les opérations d'intersection.

$$\left(\bigcap_k A_k\right)^C = \bigcup_K A_k^C \in \mathcal{F}$$

□

**Définition 1.2.1.** La famille qui vérifie les trois propriétés s'appelle une tribu (ou $\sigma$-algébre).

**Exemple 1.2.1** (Plus grande et plus petite tribu possible)**.** Soit $\Omega$ fixé. $\mathcal{F} = \{\emptyset, \Omega\}$ est la plus petite tribu possible. $\mathcal{F} = \mathcal{P}(\Omega) = \{A | A \subset \Omega\}$ est la plus grande tribu possible.

**Définition 1.2.2** (Parties de $\Omega$)**.** On appelle parties de $\Omega$ et on note $\mathcal{P}(\Omega)$ :

$$\mathcal{P}(\Omega) = \{A | A \subset \Omega\}$$

**Exemple 1.2.2.** Si $\Omega = \{P, F\}$ alors $\mathcal{P}(\Omega) = \{\Omega, \emptyset, \{P\}, \{F\}\}$.

**Définition 1.2.3.** Si $\text{card}(\Omega) = n$ alors $\text{card}(\mathcal{P}(\Omega)) = 2^n$.

**Définition 1.2.4.** Soit $\Omega$ un ensemble et $\mathcal{F}$ une famille d'événement observables sur $\Omega$. On appelle $P(A)$ la probabilité d'un événement, c'est-à-dire : $P : \mathcal{F} \to \mathbb{R}$ tel que $\frac{n_A}{n} \to P(A)$ la fréquence où l'événement est réalisé. Concrétement, $P$ doit vérifié :

1. $P(\Omega) = 1$
2. $0 \leq P(\Omega) \leq 1, \forall A \in \mathcal{F}$
3. $(A_k)$, $(A_k \in \mathcal{F})$, si $A_k \cap A_j = \emptyset$ (avec $k \neq j$) alors :

$$P\left(\bigcup_k A_k\right) = \sum_k P(A_k)$$

La troisième propriété s'appelle $\sigma$-additivité.

**Définition 1.2.5.** Soit $\Omega \neq 0$, $\mathcal{F}$ tribu et $P$ une probabilité alors $(\Omega, \mathcal{F}, P)$ est un espace probabilisé.

**Exemple 1.2.3.** $\Omega = \{P, F\}$, $\mathcal{F} = \mathcal{P}(\Omega) = \{\emptyset, \Omega, \{F\}, \{P\}\}$

1. cas symétrique :

$$P(\emptyset) = 0, P(\Omega) = 1, P(P) = \frac{1}{2}, P(F) = \frac{1}{2}$$

2. cas non-symétrique :

$$P(\emptyset) = 0, P(\Omega) = 1, P(P) = p, P(F) = 1 - p$$

## 1.3 Probabilités classiques

**Proposition 1.3.1.** *Soit* $\mathrm{card}(\Omega) < \infty$*,* $P(\{\omega_i\}) = P(\{\omega_j\})$*,* $\forall \omega_i, \omega_j \in \Omega$ *et* $\mathcal{F} = \mathcal{P}(\Omega)$*. Alors :*

$$P(A) = \frac{\mathrm{card}\, A}{\mathrm{card}\, \Omega} \textit{ (Définition classique des probabilités)}$$

*Démonstration.* $\exists p$ tel que $p = P(\{\omega_i\} = P(\{\omega_j\})$. On a : $A = \bigcup_{\omega \in A} \{\omega\}$. D'après la $\sigma$-additivité.

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) = p(\mathrm{card}\, A)$$

et si $A = \Omega$ alors $P(\Omega) = 1 = p(\mathrm{card}(\Omega))$. □

**Exemple 1.3.1.** Les probabilités classiques peuvent être considéré par le lancer d'une pièce dans le cas symétrique.

**Exemple 1.3.2.** On effectue une partie de pile ou face en trois coups. Quelle est la probabilité d'obtenir pile aux premier et troisième lancers ? On peut modéliser cette expérience en prenant $\Omega = \{P, F\}^3$ et pour famille d'événement observables $\mathcal{F} = \mathcal{P}(\Omega)$ l'ensemble de toutes les parties de $\Omega$. La pièce étant supposée symétrique, nous n'avous a priori pas de raison de supposer que l'un des 8 triplets de résultats possibles soit favorisé ou dévaforisé par rapport aux autres. On choisit donc $P$ de sorte que tous les événements élémentaires aient même probabilité (hypothèse d'équiprobabilité ou probabilité classique) soit :

$$\forall \omega \in \Omega,\ P(\{\omega\}) = \frac{1}{\mathrm{card}\, \Omega} = \frac{1}{2^3}$$

L'événement $B$ dnt on veut calculer la probabilité s'écrit :

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

D'où :

$$P(B) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

**Exemple 1.3.3.** Dans une urne, on a mls $n$ boules blanches et $n$ boules noirs indiscernable au toucher. On a donc $n + m$ boules. On choisit $k$ boules et on veut calculer $P(A_l)$ tel que :

$$A_l = \{\text{on a pioché } l \text{ boules blanches}\}$$

Donc :

$$P(A_l) = \frac{\mathrm{card}(A_l)}{\mathrm{card}\, \Omega}$$

On a alors :

$$\mathrm{card}(\Omega) = C_{n+m}^k$$

et :

$$\mathrm{card}(A_l) = C_m^l \times C_n^{k-l}$$

Donc :

$$P(A_l) = \frac{C_m^l \times C_n^{k-l}}{C_{n+m}^k}$$

Pour un exemple concret, on peut prendre celui du Loto. $n = 6$, $n + m = 49$, $l = 6$ et $k = 6$.

$$P(A_6) = \frac{1}{C_{49}^6} = 7.15112384202e - 08$$

Pour $n = m$ et $k = n$, on a :

$$P(A_l) = \frac{C_n^l C_n^{k-l}}{C_n^{2n}} = \frac{\left(C_n^l\right)^2}{C_{2n}^n}$$

On a : $l = 0, ..., n$. Soit $A_0, ..., A_n$ alors :

$$A_k \cap A_l = \emptyset,\ k \neq l$$

$$\bigcup_k A_k \supseteqq \Omega$$

On a aussi :

$$1 = P(\Omega) = P\left(\bigcup_l A_l\right) = \sum_{l=0}^{n} P(A_l)$$

Donc :

$$\sum_{l=0}^{n} \frac{\left(C_n^l\right)^2}{C_{2n}^n} = 1 \Rightarrow C_{2n}^n = \sum_{l=0}^{n} \left(C_n^l\right)^2$$

**Définition 1.3.1.** On dit que $\Omega$ est un espace discret si $\mathrm{card}(\Omega) < \infty$.

**Proposition 1.3.2.** *Soit $\mathcal{F} = \mathcal{P}(\Omega)$ et soit $P$ une probabilité sur $\mathcal{F}$. Il existe une famille $(p_k)_{k\geq 0}$ tel que :*

1. $p_k \geq 0$
2. $\sum_k p_k = 1$
3. $\forall A,\ P(A) = \sum_{k|\omega_k \in A} p_k$

*Démonstration.* On a $P(\{w_k\}) = p_k$. La propriété 1 est évidente. On démontre la propriété 3 avant la 2.

$$A = \bigcup_{k|\omega_k \in A} \{\omega_k\}$$

$$P(A) = \sum_{k|\omega_k \in A} P(\{\omega_k\}) = \sum_{k|\omega_k \in A} p_k$$

$\square$

**Exemple 1.3.4.** Soit $\Omega = \mathbb{N}$, $p_k = \frac{\lambda^k e^{-\lambda}}{k!}$ pour $\lambda > 0$, $k = 0...$ On appelle loi de Poisson $\mathcal{P}(\lambda)$ la probabilité qui vérifie $p_k = \frac{\lambda^k e^{-\lambda}}{k!}$.

## 1.4 Propriétés générales des probabilités

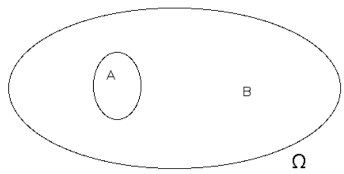
Soit $P$ une probabilité :

**Propriété 1.4.1.** *$P(A^C) = 1 - P(A)$, $P(A) = 1 - P(A^C)$.*

**Propriété 1.4.2.** *$A \subset B \Rightarrow P(A) \leq P(B)$*

*Démonstration.* On a : $B = A \cup (B \backslash A)$. Or $A$ et $B \backslash A$ sont disjoints alors :

$$P(B) = P(A) + P(B \backslash B) \geq P(A)$$




□

**Propriété 1.4.3.** *Soit $A$ et $B$ deux ensembles alors :*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

*Démonstration.* $A \cup B = A \cup (B \backslash A)$ alors :

$$P(A \cup B) = P(A) + P(B \backslash A)$$

et $B = (A \cap B) \cup B \backslash A$

$$P(B) = P(A \cap B) + P(B \backslash A)$$

□

**Propriété 1.4.4.** $P(A \cup B) \leq P(A) + P(B)$

*Démonstration.* Conséquence de la **Propriété 1.4.3.** □

**Propriété 1.4.5.** *Soient $(A_1, ..., A_n)$ alors :*

$$P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P(A_k)$$

*Démonstration.* Par réccurence. Initialisation évidente car c'est la **Propriété 1.4.3.**

$$P\left(\bigcup_{k=1}^{n+1} A_k\right) = P\left(\bigcup_{k+1}^{n} A_k \cup A_{k+1}\right) \leq P\left(\bigcup_{k=1}^{n} A_k\right) + P(A_{n+1})$$

□

*Remarque.* Si $(A_k)$ infini alors la **Propriété 1.4.5.** est vérifiée.

**Propriété 1.4.6** (Propriété de continuité de $P$)**.** *a) $(A_n)$ croissante $(\forall n, A_n \subset A_{n+1})$ alors $P(A_n)$ est croissante. On a alors :*

$$\lim_{n \to +\infty} P(A_n) = P\left(\bigcup_{k} A_k\right)$$

*b) $(A_n)$ décroissate $(\forall n, A_{n+1} \subset A_n)$ alors $P(A_n)$ converge vars* $P\left(\bigcap_{k} A_k\right)$.

*Démonstration.* $a) \Rightarrow b)$ : soit $(A_n)$ décroissante. Si on prend $B_n = A_n^C$, on a $(B_n)$ croissante. Alors $P(B_n) \xrightarrow[n\to\infty]{} P\left(\bigcup_k B_k\right)$. Mais $P(B_n) = 1 - P(A_n)$. Donc : $P(A_n) \to 1 - P\left(\bigcup_k B_k\right)$ Or : $1 - P\left(\left(\bigcup_k B_k\right)^C\right) = P\left(\bigcap_k (B_k)^C\right) = P\left(\bigcap_k (A_K)\right)$. □

**Propriété 1.4.7.** *La Remarque précédente peut constitué une propriété. On rappelle : soit* $(A_k)$ *infini alors :*

$$P\left(\bigcup_{k=0}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} P(A_k)$$

*Démonstration.* $B_n = \bigcup_{k=1}^{n} A_k$ avec $(B_n)$ croissante. On a :

$$P(B_n) \xrightarrow[n\to\infty]{} P\left(\bigcup_{k=1}^{\infty} B_n\right)$$

et :

$$\bigcup_{k=1}^{\infty} B_k = \bigcup_{k=1}^{\infty} A_k$$

alors :

$$P(B_n)\left(\to P\left(\bigcup_{k=1}^{\infty} A_k\right)\right) = P\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} P(A_k) \leq P(A_n)$$

□

# Chapitre 2

# Conditionnement et indépendance

## 2.1 Probabilités conditionnelles

### 2.1.1 Introduction

Comment doit-on modifier la probabilité que l'on attribue à un évenement lorsque l'on dispose d'une information supplémentaire ? Le concept de probabilité conditionnelle permet de répondre à cette question.

**Exemple 2.1.1.** Soit $\mathrm{card}(\Omega) < \infty$ alors :

$$P(A) = \frac{\mathrm{card}(A)}{\mathrm{card}(\Omega)}$$

et $B$ une partie de $\Omega$ avec $B \cap A \neq \emptyset$. Alors :

$$P_A(B) = \frac{\mathrm{card}(A \cap B)}{\mathrm{card}(A)} = \frac{P(A \cap B)}{P(A)}$$

**Définition 2.1.1.** $P_A : \mathcal{F} \to \mathbb{R}^+$ alors :

$$P_A(B) = \frac{P(A \cap B)}{P(A)},\ B \in \mathcal{F},\ P(A) > 0$$

$P_A(B)$ s'appelle la probabilité conditionelle sachant $A$.

*Démonstration.* On vérifie que la **Définition 2.1.1.** est une relation de probabilité.

1) $P_A(\Omega) = \dfrac{P(A \cap \Omega)}{P(A)} = \dfrac{P(A)}{P(A)} = 1$

2) $0 \leq P_A(B) \leq 1$ car $P(A \cap B) \leq P(A)$.

3) $(B_k)$, $B_k \cap B_l = \emptyset$, $k \neq l$ :

$$P_A\left(\bigcup_k (B_k)\right) = \frac{P\left(A \cap \left(\bigcup_k B_k\right)\right)}{P(A)} = \frac{P\left(\bigcup_k A \cap B_k\right)}{P(A)} = \frac{\sum_k P(A \cap B_k)}{P(A)} = \sum_k P_A(B_k)$$

$\square$

**Exemple 2.1.2.** Supposons qu'on a deux urnes. Dans l'urne 1, on a : $m$ boules blanches et $n$ noires et dans l'urne 2, $m$ boules blanches et $n$ boules noires. On transpose une boule au hasard (sans regarder sa couleur) de la urne 1 vers la urne 2. Ensuite, on tire au hasard une boule dans l'urne 2 et on regarde sa couleur.

Soit $B = \{$la deuxième boule tirée est blanche$\}$. On simplifie l'expérience en utilisant :

- $A_1 = \{$Première blanche$\}$
- $A_2 = \{$Première noire$\}$

alors on a :

$$P_{A_1}(B) = \frac{m+1}{m+n+1} \qquad P_{A_2}(B) = \frac{m}{m+n+1}$$

### 2.1.2 Propriétés

**Proposition 2.1.1** (Conditionnement successifs)**.** *Soit* $B = A_1 \cap A_2 \cap ... \cap A_n$ *avec* $P(A_i) > 0$ *alors :*

$$P(B) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_3)...P_{A_1\cap A_2\cap...\cap A_{n-1}}(A_n)$$

*Démonstration.*

$$\begin{aligned} P(A_1)\frac{P(A_1\cap A_2}{P(A_1)}\frac{P(A_3\cap(A_2\cap A_1))}{P(A_1\cap A_2)} \times ... \times \frac{P(A_n\cap(A_{n-1}\cap ... \cap A_1)}{P(A_{n-1}\cap ... \cap A_1)} \\ = P(A_1\cap A_2)\frac{P(A_3\cap(A_2\cap A_1))}{P(A_1\cap A_2)} \times ... \times \frac{P(A_n\cap(A_{n-1}\cap ... \cap A_1)}{P(A_{n-1}\cap ... \cap A_1)} \\ = ... = P(A_1\cap A_2\cap ... \cap A_n) = P(B) \end{aligned}$$

□

**Exemple 2.1.3.** Soit une urne avec $2n$ boule avec 2 boules identiques (même couleur) 2 à 2. On choisit 2 boules à la fois et on répete l'expérience jusqu'à tant qu'il n'y a plus de boules.

Soit $B = \{$chaque sorte de boules donne une paire de même couleur$\}$. On introduite des événements :

- $A_1 = \{$les 2 premières sont de même couleurs$\}$
- $A_2 = \{$les 2 boules suivant $A_1$ sont de même couleurs$\}$
- $\cdots$
- $A_n = \{$les 2 boules suivant $A_{n-1}$ sont de même couleurs$\}$

On a alors : $B = A_1 \cap ... \cap A_n$.

$$\begin{array}{l} P_(A_1) = \frac{n}{C_{2n}^2} \\ P_{A_1}(A_2) = \frac{n-1}{C_{2n-2}^2} \\ \vdots \end{array} \quad \text{et} \quad \begin{array}{rcl} P(B) & = & \dfrac{n}{C_{2n}^2}\dfrac{n-1}{C_{2n-1}^2}\cdots\dfrac{1}{C_2^2} \\ & = & \frac{n!\times 2^n}{2n(2n-1)(2n-2)...1} = \frac{n!2^n}{(2n)!} = \frac{1}{\prod 2k+1} \end{array}$$

**Proposition 2.1.2** (Probabilité totale)**.** *Soit* $(A_k)$ *qui forme une partition de* $\Omega$*.*

1. *disjoints*
2. $\bigcup_k A_k = \Omega$

*Soit* $B$ *un événement. Alors :*

$$P(B) = \sum_k P(A_k)P_{A_k}(B)$$

*Démonstration.* On a : $B = B \cap \Omega = B \cap \left( \bigcup_k A_k \right) = \bigcup_k (B \cap A_k)$. Alors :

$$P(B) = P\left( \bigcup_k (B \cap A_k) \right) = \sum_k P(B \cap A_k)$$

$\square$

**Exemple 2.1.4** (Retour sur l'**Exemple 2.1.2**)**.** On a alors :

$$P(B) = P(A_1)P_{A_1}(B) + P(A_2)P_{A_2}(B)$$

et :

$$P(A_1) = \frac{m}{m+n} \text{ et } P(A_2) = \frac{n}{m+n}$$

Donc :

$$P(B) = \frac{m}{m+n} \times \frac{m+1}{m+n+1} + \frac{n}{m+n} \times \frac{m}{m+n+1}$$

**Exemple 2.1.5.** On considère un echequier de 64 cases. On choisit deux cases de l'échequier et on y place deux fois. On dit que deux fois sont en position d'attaque si ils sont sur la même diagonale. Soit $B = \{$les deux fous sont en position d'attaque$\}$. On introduit des événements.

- $A_1 = \{$le premier fou est à l'extrémité de l'échequier$\}$
- $A_2 = \{$le premier fou est à une case de l'extrémité de l'échequier$\}$
- $A_3 = \{$le premier fou est à deux cases de l'extrémité de l'échequier$\}$
- $A_4 = \{$le premier fou est à trois cases de l'extrémité de l'échequier$\}$

En $A_1$, le premier fou a 7 possibilités d'être en position d'attaque avec le deuxième fou. En $A_2$, il en a 9. En $A_3$, il en a 11 et en $A_4$, 13. On a alors :

$$P(A_1) = \frac{28}{64},\ P(A_2) = \frac{20}{64},\ P(A_3) = \frac{12}{64},\ P(A_4) = \frac{4}{64}$$

$$P_{A_1}(B) = \frac{7}{63},\ P_{A_2}(B) = \frac{9}{63},\ P_{A_3}(B) = \frac{11}{63},\ P_{A_4}(B) = \frac{13}{63}$$

Alors :

$$P(B) = \frac{28}{64} \times \frac{7}{63} + \frac{20}{64} \times \frac{9}{63} = \frac{12}{64 \times 63} + \frac{4 \times 13}{64 \times 63}$$

**Proposition 2.1.3** (Formule de Bayes)**.** *Soit $A$ un événement de proabilité non nulle. Si les événements $H_i$ $(1 \leq i \leq n)$ forment une partition de $\Omega$ et aucun $P(H_i)$ n'est nul, on a pour tout $j = 1, ..., n$ :*

$$P_A(H_j) = \frac{P_{H_j}(A)P(H_j)}{\sum_{i=1}^{n} P_{H_i}(A)P(H_i)}$$

*Démonstration.* Par définition des probabilités conditionnelles on a :

$$P_A(H_j) = \frac{P(A \cap H_j)}{P(A)} = \frac{P_{H_j}(A)P(H_j)}{P(A)}$$

Et il ne reste plus qu'à dévelpper $P(A)$ en conditionnant par la partition $(H_i, 1 \leq i \leq n)$. $\square$

## 2.2 Indépendance

### 2.2.1 Indépendance de deux événements

Soit $A, B \in \Omega$. On suppose que $A$ est réalisé alors $P(B) = P_A(B)$ et $P(A) = P_B(A)$. On aura donc :

$$P(A \cap B) = P(A) \times P(B)$$

si $P(A) > 0$ et $P(B) > 0$.

**Définition 2.2.1.** Deux événements $A$ et $B$ sont indépendants si :

$$P(A \cap B) = P(A) \times P(B)$$

**Exemple 2.2.1.** Soit un jeu de 36 cartes. On considère :

$$A = \{\clubsuit\}$$

$$B = \{D\}$$

Alors :

$$P(A) = \frac{9}{36} \qquad P(B) = \frac{4}{36}$$

et :

$$A \cap B = \{D\clubsuit\}$$

$$P(A \cap B) = \frac{9 \times 4}{36 \times 36} = \frac{36}{36^2} = \frac{1}{36}$$

**Proposition 2.2.1.** *Si $A$ et $B$ sont indépendants alors $A^C$ et $B$ sont indépendants.*

*Démonstration.*

$$\begin{aligned} P(A^C)P(B) &= (1 - P(A))P(B) = P(A) - P(A)P(B) \\ &= P(A) - P(A \cap B) = P(A \cap B^C) \end{aligned}$$

□

*Conséquence.*

1. Si $A$ et $B$ sont indépendants alors $A$ et $B^C$ sont indépendants et $A^C$ et $B^C$ sont indépendants.
2. On a aussi $\Omega$ et $A$ sont indépendants et $\emptyset$ et $A$ sont indépendants.

### 2.2.2 Indépendance mutuelle

**Définition 2.2.2** (Fausse définition pour l'indépendance de plusieurs éléments)**.** Soit $A_1, ..., A_n$ un ensemble d'événements. On dit que $A_1, ..., A_n$ sont tous indépendants si $A_i, A_j$ sont indépendants pour $i \neq j$.

Mais on va voir quette définition n'est pas vraie.

**Exemple 2.2.2.** Soit $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $p\{\omega_i\} = \frac{1}{4}$

$$A_1 = \{\omega_1, \omega_2\}$$

$$A_2 = \{\omega_1, \omega_3\}$$

$$A_3 = \{\omega_1, \omega_3\}$$

On a :

$$P(A_i) = \frac{1}{2}$$

Mais pour $i \neq j$, $A_i, A_j$ sont indépendants car :

$$P(A_i \cap A_j) = P\{\omega_1^{=} = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$$

Mais $B = A_2 \cap A_3$ et $A_1$. On vérifie si oui ou non on a $P(B \cap A_1) = P(B) \cap P(A_1)$.

$$P(B) = \frac{1}{4} \qquad P(A_1) = \frac{1}{2}$$

$$P(B \cap A_1) = P\{\omega_1\} = \frac{1}{4} \neq \frac{1}{4} \times \frac{1}{2}$$

**Définition 2.2.3.** Soient $A_1, ..., A_n$ des événements. Ils sont indépendants si :

$$P(A_i \cap A_j) = P(A_i)P(A_j) \qquad i \neq j$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k) \qquad i \neq j \neq k$$

$$\cdots\cdots\cdots$$

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

**Proposition 2.2.2.** *Soient $A_1, ..., A_n$ des événements indépendants et $T_1, ..., T_k$ tel que :*

$$T_i \cap T_j = \emptyset$$

$$T_1 \cup ... \cup T_k = \{1, ..., n\}$$

*Supposons que $B_1$ soit une combinaison de $A_i$ avec $i \in T_1$, $B_2$ soit une combinaison de $A_i$ avec $i \in T_2$,...,$B_k$ soit une combinaison de $A_i$, $i \in T_k$ alors $B_1, B_2, ..., B_k$ sont indépendants.*

**Corollaire.** *On suppose $B_j = A_j$ ou $B_j = A_j^C$. Si $A_1, ..., A_n$ sont indépendants alors $B_1, ..., B_n$ sont indépendants.*

# Chapitre 3

# Variables aléatoires

## 3.1 Généralités

**Définition 3.1.1.** Soit $(\Omega, \mathcal{F}, P)$ est une espace probabilisée. On appelle variable aléatoire discrète sur $(\Omega, \mathcal{F}, P)$ toute application $X$ :

$$\begin{array}{rcccl} X & : & \Omega & \rightarrow & \mathbb{R} \\ & & \omega & \mapsto & X(\omega) \end{array}$$

vérifiant :

1. $\{\omega \,|\, X(\omega) \in \Omega\}$ est une partie au plus dénombrable de $\mathbb{R}$.
2. Pour tout $x \in X(\Omega)$, $A = \{\omega \in \Omega, X(\omega) = x\}$ fait partie de la famille $\mathcal{F}$ d'événements auxquels on peut attribuer une probabilité sur $P$.

Un cas simple est que $\Omega$ soit fini et dénombrable alors $\mathcal{F} = \mathcal{P}(\Omega)$.

**Exemple 3.1.1.** On suppose qu'on lance 2 dés en même temps :

$$\Omega = \{(i, j), i, j = 1...6\}$$

$$\mathcal{F} = \mathcal{P}(\Omega)$$

$X(i, j) = i + j$. Les valeurs possibles sont $\{2, ..., 12\}$. Quelle est la valeur de $P\{X = k\}$ ?

$$\begin{array}{ll} P\{X = 2\} = \frac{1}{36} & P\{X = 10\} = \frac{3}{36} \\ P\{X = 3\} = \frac{2}{36} & P\{X = 11\} = \frac{2}{36} \\ P\{X = 4\} = \frac{3}{36} & P\{X = 12\} = \frac{1}{36} \end{array}$$

**Définition 3.1.2** (Répartition ou loi)**.** Soit $F$ une fonction :

$$F(x) = P\{X \leq x\},\ x \in \mathbb{R}$$

On l'appelle fonction de répartition de $X$ avec $X = ]a, b]$

$$P\{X \in ]a, b]\} = F(b) - F(a)$$

*Démonstration.* Soit $A = \{X \leq a\}$, $B = \{X \leq b\}$ et $C = \{X \in ]a, b[\}$ alors :

$$A \cap C = \emptyset$$

$$A \cup C = B$$

$$P\{X \leq b\} = P(A \cup C) = P(A) + P(C) = P\{X \leq a\} + P\{X \in ]a, b]\}$$

$\square$

**Propriétés de la fonction de répartition**

**Propriété 3.1.1.** *On a :* $0 \leq F(x) \leq 1$.

**Propriété 3.1.2.** *Si* $x < y$ *alors* $F(x) \leq F(y)$.

**Propriété 3.1.3.** $F$ *est continue à droite.*

**Propriété 3.1.4.** $\lim_{x \to -\infty} F(x) = 0$ *et* $\lim_{x \to +\infty} F(x) = 1$

**Propriété 3.1.5.** $P\{X \in ]a, b]\} = F(b) - F(a)$

*Démonstration.* On vérifie d'abord la croissance de $F$ sur $\mathbb{R}$. Soient $s$ et $t$ deux réels quelconques tel que $s \leq t$. Tout $\omega$ vérifiant $X(\omega) \leq s$ vérifie a fortiori $X(\omega) \leq t$. Cette implication se traduit par l'inclusion d'événements $\{X \leq s\} \subset \{X \leq t\}$ d'où $P(X \leq s) \leq P(X \leq t)$, autrement dit $F(s) \leq F(t)$. Ainsi $F$ est croisate. Il en résulte qu'elle possède une limite à gauche et une limite à droite en tout point de $x$ de $\mathbb{R}$.

Le reste de la preuve repose sur la propriété de continuité monotone séquentielle de la probabilité qu'on rappelle :

*Rappel.* Soit $P$ une probabilité sur l'espace $(\Omega, \mathcal{F})$.

(i) Si $(A_n)_{n \geq \mathbb{N}^*}$ est une suite croissante d'événements (c'est-à-dire $\forall n \in \mathbb{N}^*, A_n \subset A_{n+1}$) alors :

$$\lim_{n \to +\infty} P(A_n) = P(A) \text{ où } A = \bigcup_{n \in \mathbb{N}^*} A_n$$

.

(ii) Si $(B_n)_{n \in \mathbb{N}^*}$ est une suite décroissante d'événements (c'est-à-dire $\forall n \in \mathbb{N}^* B_{n+1} \subset B_n$) alors :

$$\lim_{n \to +\infty} P(B_n) = P(B) \text{ où } B = \bigcap_{n \in \mathbb{N}^*} B_n$$

Soit $x \in \mathbb{R}$ fixé. Comme on est assuré de l'existence de la limite à droite de $F$ en ce point, pour montrer que cette limite vaut $F(x)$ et établir la continuité à droite $F$ en $x$, il suffit de vérifier que : $F(x) = \lim_{n \to +\infty} F\left(x + \frac{1}{n}\right)$. Comme $F(x + \frac{1}{n}) = P(X \leq x + \frac{1}{n})$, ceci résulte de la propriété $(ii)$ ci dessus appliquée aux événements $B_n = \{X \leq x + \frac{1}{n}\}$, en remarquant que :

$$\bigcap_{n \in \mathbb{N}^*} \left\{X \leq x + \frac{1}{n}\right\} = \{X \leq x \qquad (3.1)$$

En effet, pour tout $\omega \in \{X \leq x\}$, on a $X(\omega) \leq x \leq x + \frac{1}{n}$ pour tout $n \in \mathbb{N}^*$ et donc $\omega \in \bigcap_{n \in \mathbb{N}^*} B_n$ d'où $\{X \leq x\}$ est inclus dans l'intersection des $B_n$ $(n \in \mathbb{N}^*)$. Réciproquement, tout $\omega$ de cette intersection vérifie : $\forall n \in \mathbb{N}^*$, $X(\omega) \leq x + \frac{1}{n}$. Le passage à la limite quand $n$ tend vers l'infini conservant l'intersection des $B_n$ $(n \in \mathbb{N}^*)$ est incluse dans $\{X \in x\}$, ce qui achève la vérification de (3.1).

Comme $F$ est croissante, elle admet des limites en $-\infty$ et $+\infty$. Si $X(\Omega)$ est borné inférieurement [1], il est clair que $F(x) = 0$ pour tout $x$ assz petit et donc $\lim_{n \to -\infty} F(x) = 0$. Dans le cas général, on identifie la limite en $-\infty$ (dont on connait l'existence) grâce à une suite particulière :

$$\lim_{x \to -\infty} F(n) = \lim_{n \to +\infty} F(-n) = \lim_{n \to +\infty} P(X \leq -n)$$

[1]Attention, dans le cas général $x_0$ n'est pas nécessairement le plus petit élément de $X(\Omega)$, on peut très bien avoir par exemple $X(\Omega) = \mathbb{Z}$.

On utilise à nouveau la propriété $(ii)$ avec $B_n = \{X \leq -n\}$ en montrant que :

$$\bigcap_{n \in \mathbb{N}^*} \{X \leq -n\} = \emptyset \qquad (3.2)$$

En effet, soit $\omega$ un élément de cette intersection. Alors $X(\omega) \leq -n$ pour tout $n \in \mathbb{N}^*$ donc en passant à la limite, $X(\omega) = -\infty$ ce qui est impossible puisque $X$ ne prend que des valeurs finies[2]. Donc cette intersection est vide et par $(ii)$, $\lim\limits_{n \to +\infty} P(X \leq -n) = P(\emptyset) = 0$. La preuve de $\lim\limits_{x \to +\infty} F(x) = 1$ est analogue. □

## 3.2 Variables aléatoires discrètes

**Définition 3.2.1.** $X$ est une variable aléatoire discrète si $\exists V = \{x_1, x_2, ..., x_n\}$, $V \subset \mathbb{R}$ tel que $P\{X \in V\} = 1$. On a :

1) $\{X = x_k\} \cap \{X = x_n\} = \emptyset$, $k \neq n$

2) $\bigcup\limits_k \{X = x_k\} = \{X \in V\}$

**Proposition 3.2.1.** *Soit* $p_k = P\{X = x_k\}$, $\forall A \subset \mathbb{R}$

$$P\{X \in A\} = \sum_{k | x_k \in A} p_k$$

*Démonstration.* On a que :

$$P\{X \in A\} = \{\{X \in A\} \cap \{X \in V\}\} = P\left\{\{X \in A\} \cap \left\{\bigcup_k \{X = x_k\}\right\}\right\}$$

$$= P\left\{\bigcup_k \{X \in A\} \cap \{X = x_k\}\right\} = P\left\{\bigcup_{k | x_k \in A} \{X = x_k\}\right\} = \sum_{k | x_k \in A} p_k$$

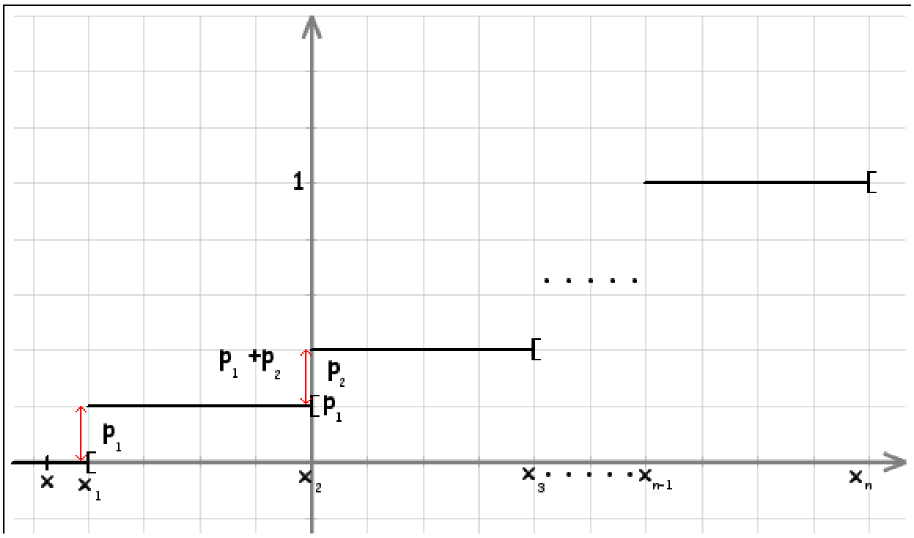
□

**Exemple 3.2.1.** $A = \mathbb{R}$, $\{X \in A\} = \Omega$. On a alors :

$$1 = \sum_k p_k$$

**Exemple 3.2.2.** Soit $\{x_k\}_{k=1,...,n}$, $x_1 < x_2 < ... < x_n$, $p_k = \{X = x_k\}$, $F(x) = P\{X \in A\} = \sum\limits_{k | x_k \in ]-\infty, x]} p_k$

[2]Ce qui signifie que $X(\Omega)$ soit borné...




## 3.3 Lois discrètes classiques

### 3.3.1 Loi de Bernouilli

**Définition 3.3.1.** La variable aléatoire $X$ suit la loit de Bernoulli de paramètre $p$ ($p \in [0,1]$) si elle ne prend que deux valeurs 0 et 1 avec :

$$P(X=1)=p \qquad P(X=0)=1-p=q$$

On notera $X \sim \mathcal{B}(p)$.

Si $A$ est un événements de probabilité $p$, son indicatrice définie par :

$$\mathbf{1}_A(\omega)=\begin{cases}1 \text{ si } \omega\in A\\ 0 \text{ si } \omega\notin A\end{cases} = \begin{cases}1 \text{ si } A \text{ est réalisé}\\ 0 \text{ si } A \text{ n'est pas réalisé}\end{cases}$$

est une variable aléatoire suivant la loi de Bernouilli de paramètre $p$. Réciproquement, si $X$ est un variable aléatoire de Bernouilli, on peut toujours écrire $X = \mathbf{1}_A$ en définissant $A = \{\omega \in \Omega, X(\omega)=1\}$.

### 3.3.2 Loi uniforme sur un ensemble fini de réels

**Définition 3.3.2.** La variable aléatoire $X$ suit la uniforme sur l'ensemble des réels $\{x_1,...,x_n\}$ si $P_X$ est l'équiprobabilité sur cet ensemble.

Autrement dit, l'ensemble des valeurs possibles de $X$ est $X(\Omega)=\{x_1,...,x_n\}$ et :

$$\forall k\in\{1,...,n\} \qquad P(X=x_k)=\frac{1}{n}$$

Par exemple, le nombre de points indiqué par un dé suit la loi uniforme sur $\{1,2,3,4,5,6\}$.

### 3.3.3 Lois binomiales

**Définition 3.3.3.** La variable aléatoire $X$ suit la loi binomiale de paramètres $n$ et $p$ ($n\in\mathbb{N}^*$ et $p\in[0,1]$) si l'ensemble des valeurs possibles est $X(\Omega)=\{0,1,...,n\}$ et :

$$\forall k\in\{0,1,...,n\} \qquad P(X=k)=C_n^k p^k(1-p)^{n-k}$$

*Notation.* $X \sim \mathcal{B}(n,p)$

La forme ci-dessus définit bien une loi de probabilité puisuqe les $C_n^k p^k (1-p)^{n-k}$ sont positifs et :

$$\sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$$

en appliquant la formule du binôme de Newton (d'où le nom de la loi). La loi binomiale $\mathcal{B}(n,p)$ est la loi du nombre de succès obtenus en une suite de $n$ épreuves répétées indépendantes avec pour chaque épreuve une probabilité de succès $p$.

De même, soit $A_1, ..., A_n$ une famille de $n$ événements mutuellement indépendants ayant tous même probabilité $p$ et notons $X_i$ la variable de Bernouilli indicatrice de $A_i$ :

$$X_i(\omega) = \begin{cases} 1 \text{ si } \omega \in A_i \\ 0 \text{ si } \omega \in A_i^C \end{cases}$$

Alors la variable aléatoire $S_n = \sum_{i=1}^{n} X_i$ suit la loi binomiale $\mathcal{B}(n,p)$.

### 3.3.4 Lois hypergéométriques

Alors que la loi binomiale intervient dans les tirages avec remise, la loi hypergéométrique correspond aux tirages sans remise.

**Exemple 3.3.1.** Dans une production totale de $N$ objets dont $M$ sont défectueux, on prélève au hasard un échantillon de $n$ objets (tirage sans remise). Soit $X$ le nombre aléatoire d'objets défectueux dans l'échantillon. Quelle est sa loi ?

On peut prendre comme espace $\Omega$ l'ensemble de tous les échantillons possibles (toutes les parties à $n$ éléments d'un ensemble de cardinal $N$) muni de l'équiprobabilité. Chaque échantillon a ainsi une probabilité $1/C_N^n$ d'être choisi. Les échantillons (événements élémentaires) réalisent l'événement $\{X = k\}$ sont ceux qui contiennent $k$ objets défectueux et $n-k$ objets défectueux. Ceci n'est réalisable que si $0 \leq k \leq M$ et $0 \leq n-k \leq N-M$. On dénombre ces échantillons. On les forme en choisissant $k$ objets défectueux dans une sous-population de $M$ et en complétant par $n-k$ objets non défectueux chosis dans une population de $N-M$. Il y en a donc $C_M^k \times C_{N-M}^{n-k}$. Finalement :

$$P(X = k) = \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n} \text{ si } \begin{cases} 0 \leq k \leq M \\ 0 \leq n-k \leq N-M \end{cases}$$

**Définition 3.3.4.** La loi définie par l'équation suivante :

$$P(X = k) = \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n} \text{ si } \begin{cases} 0 \leq k \leq M \\ 0 \leq n-k \leq N-M \end{cases}$$

s'appelle la loi hypergéométrique de paramètres $N$, $M$ et $n$. Notation : $X \sim \mathcal{H}(N, M, n)$. Le paramètre $N$ est l'effectif de la population totale, $M$ celui de la sous-population à laquelle on s'interesse et $n$ la taille de l'échantillon observé.

Pour une taille d'échantillon $n$ fixée, plus $N$ et $M$ sont grands, moins les tirages sans remise diffèrent des tirages avec remise. Plus précisement, la loi hypergéométrique converge vers la loi binomiale au sens suivant :

**Theorème 3.3.1.** *On suppose que quand $N$ tend vers $+\infty$, $M = M(N)$ tend vers $+\infty$ en vérifiant la condition :*

$$\lim_{N\to+\infty} \frac{M}{N} = p \text{ avec } 0 < p < 1$$

*Alors, $n$ restant fixé, la loi hypergéométrique $\mathcal{H}(N, M, n)$ converge la loi binomiale $\mathcal{B}(n, p)$, ce qui signifie que si $(X_N)_{N\geq 1}$ est une suite de variable aléatoire avec $X_N \sim \mathcal{H}(N, M, n)$ et $Y$ est une variable de loi binomiale $\mathcal{B}(n, p)$ alors :*

$$\forall k \in \{0, 1, ..., n\} \qquad \lim_{N\to+\infty} P(X_N = k) = P(Y = k)$$

*autrement dit :*

$$\forall k \in \{0, 1, ..., n\} \qquad \lim_{N\to+\infty} \frac{C_M^k \times C_{N-M}^{n-k}}{C_N^n} = C_n^k p^k (1-p)^{n-k}$$

### 3.3.5 Lois géométriques

**Exemple 3.3.2.** Une variable aléatoire $X$ suit la loi géométrique de paramètre $p \in ]0, 1[$, si $X(\Omega) = \mathbb{N}^*$ et :

$$\forall k \in \mathbb{N}^*, \; P(X = k) = (1-p)^{k-1} p$$

*Notation.* $X \sim \mathcal{G}(p)$

$\mathcal{G}(p)$ est la loi du nombre de tentatives en $n$ épreuves indépendantes.

### 3.3.6 Lois de Poisson

**Définition 3.3.5.** On dit que la variable aléatoire discrète $X$ suit la loi de Poisson de paramètre $\lambda > 0$, si l'ensemble des valeurs possibles est $X(\Omega) = \mathbb{N}$ et :

$$\forall k \in \mathbb{N} \qquad P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

*Notation.* $X \sim \mathcal{P}(\lambda)$

## 3.4 Variables aléatoires continues

### 3.4.1 Généralités

**Définition 3.4.1.** Soit $X$ une variable aléatoire à valeurs dans $\mathbb{R}$ et $f_X$ une densité de probabilité sur $\mathbb{R}$. On dit que $X$ est une variable aléatoire continue de densité $f_X$ si pour tout intervalle $A$ de $\mathbb{R}$ on a :

$$P(X \in A) = \int_A f_X(x)dx$$

La loi de la variable aléatoire $X$ est la loi continue sur $\mathbb{R}$ de densité $f_X$.

Pour déterminer la loi d'une variable aléatoire continue, il faut donc calculer sa densité. De manière équivalente, on déterminer la loi d'une variable continue en donnant la probabilité qu'elle appartienne à un intervalle $I$ quelconque.

Une variable aléatoire continue $X$, de densité $f_X$, tombe entre $a$ et $b$ avec une probabilité égale à :

$$P(a < X < b) = \int_a^b f_X(x)dx$$

Plus la densité $f_X$ est élevée au dessus d'un segment, plus les chances que $X$ a d'atteindre ce segment sont élevées, ce qui justifie le terme "densité".

On a :

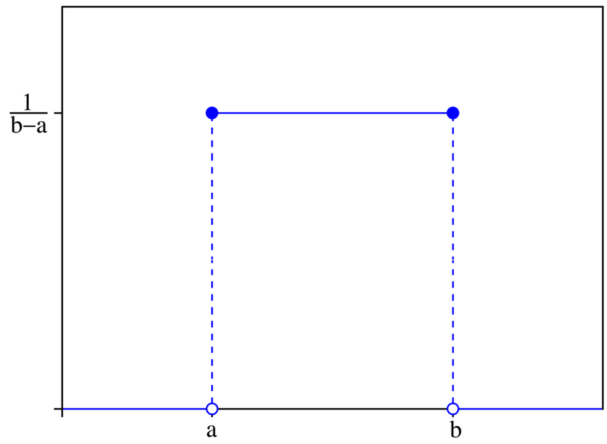
$$P(X \in [a,b]) = P(X \in [a,b[) = P(X \in ]a,b[) = P(X \in ]a,b])$$

### 3.4.2 Loi pour les variables aléatoires continues

#### Loi uniforme

$X \sim \mathcal{U}[a,b]$ alors :

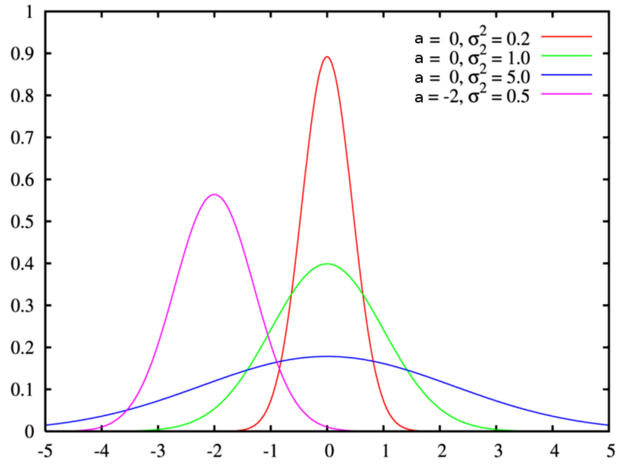
$$f_X(x) = \frac{1}{b-a}\mathbf{1}_{[a,b]}$$




#### Loi gausienne (loi normale)

$X \sim \mathcal{N}(a, \sigma^2)$, $a \in \mathbb{R}^2$, $\sigma \geq 0$, si :

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$$




### 3.4.3 Analogie entre variable aléatoire continue et discrète

$$F(x) = P\{X \leq x\} = P\{X \in ]-\infty, x]\} = \int_{]-\infty,x]} f_X(t)dt = \int_{-\infty}^{t} f_X(t)dt$$

## 3.5 Indépendance des variables aléatoires

### 3.5.1 Introduction

Soient $X$ et $Y$ deux variables aléatoires. On a :

$$\{X \in B\} \qquad \{Y \in C\}$$

On va dire que $X$ et $Y$ sont indépendants si les événements $\{X \in B\}$ et $\{Y \in C\}$ sont indépendants.

### 3.5.2 Généralités

**Définition 3.5.1.** Soients $X_1, ..., X_n$ variables aléatoires. On dit qu'ils sont indépendants si les événements $\{X_1 \in B_1\}$,...,$\{X_n \in B_n\}$ sont indépendants, $\forall B_1, ..., B_n \subset \mathbb{R}$.

**Theorème 3.5.1.** *$X_1, ..., X_n$ variables aléatoires discrètes et $\{a_j\}$ alors :*

$$X_1, ..., X_n \textit{ indépendants} \Leftrightarrow P\{X_1 = a_{j_1}, ..., X_n = a_{j_n}\} = \prod_{k=1}^{n} P\{X_k = a_{j_k}\}$$

*pour chaque $j_1, ..., j_n$.*

*Démonstration.* $(\Rightarrow)$ $B_k = \{a_{j_k}\}$ alors $\{X_1 = a_{j_1}\}$,...,$\{X_n = a_{j_n}\}$ sont indépendants. On a alors :

$$P\{X_1 = a_{j_1}, ..., X_n = a_{j_n}\} = \prod_{k=1}^{n} P\{X_k = a_{j_k}\}$$

$(\Leftarrow)$

$$P\{X_1 \in B_1, ..., X_n \in B_n\} = \prod P\{X_k \in B_k\}$$

On remarquera que :

$$\{X_1 \in B_1, ..., X_n \in B_n\} = \bigcap_{(*)} \{X_1 \in a_{j_1}, ..., X_n \in a_{j_n}\}$$

avec $(*) = (j_1, ..., j_n) \,|\, a_{j_k} \in B_k$ pour $k = 1, ..., n$.

$$\begin{aligned} P\{X_1 \in B_1, ..., X_n \in B_n\} &= \sum_{(*)} P\{X = a_{j_1}, ..., X_n = a_{j_n}\} \\ &= \sum_{(*)} \prod_{k=1}^{n} P\{X_k = a_{j_k}\} \\ &= \prod_{n} \sum P\{X_k = a_{j_k}\} \\ &= \prod_{k=1}^{n} P\{X_k \in B_k\} \rightarrow \text{ indépendants} \end{aligned}$$

car :

$$P\{A_i \cap A_j \cap A_k\} = P\{X_i \in B_i, X_j \in B_j, X_k \in B_k\} = P\{X_1 \in \mathbb{R}, ..., X_{i-1} \in \mathbb{R}, X_i \in B_i, X_{i+1} \in \mathbb{R}, ..$$

$\square$

## 3.6 Epreuves de Bernouilli

**Définition 3.6.1.** Soit $k$ un événement et $\varepsilon_k$ sa variable aléatoire dont les valeurs de $\varepsilon_k$ sont 0 et 1 (de probabilités respectives $q = 1 - p$ et $p$). On considère la suite $(\varepsilon_1, ..., \varepsilon_n)$ avec $\varepsilon_1, ..., \varepsilon_n$ indépendants. On appelle la suite de $n$ épreuves de Bernouilli l'ensemble des $\varepsilon_k$, $\forall k = 1, ..., n$. Les événements $\varepsilon_k$ suivent la loi de Bernouilli de paramètre $p$.

Soit $\nu_n$ le nombre de succès de ces $n$ événements alors :

$$\nu_n = \sum_{k=1}^{n} \varepsilon_k$$

**Proposition 3.6.1.** *$\nu_n \sim \mathcal{B}(n, p)$ c'est-à-dire :*

$$p_n(k) = P\{\nu_n = k\} = C_n^k p^k (1-p)^{n-k}, \qquad \forall k \in \{1, ..., n\}$$

*Démonstration.*

$$P\{\nu_k = k\} = \sum_{(*)} P\{\varepsilon_1 = a_1, \varepsilon_2 = a_2, ..., \varepsilon_n = a_n\} \qquad (\blacktriangle)$$

avec $(*) = (a_1, ..., a_n) \,|\, a_i = \{0, 1\}$ et $\displaystyle\sum_{i=1}^{n} a_i = k$.

$$\{\nu_k = k\} = \bigcap_{(*)} \{\varepsilon_1 = a_1, \varepsilon_2 = a_2, ..., \varepsilon_n = a_n\}$$

$$P\{\varepsilon_1 = a_1, ..., \varepsilon_n = a_n\} = \prod P\{\varepsilon_k = a_k\} = p^k (1-p)^{n-k}$$

$$(\blacktriangle) = p^k (1-p)^{n-k} \, \mathrm{card}(A)$$

avec :

$$A = \left\{ (a_1, ..., a_n) \middle| a_i = \begin{cases} 0 \\ 1 \end{cases} \text{et } \sum_{i=1}^{n} a_i = k \right\}$$

et $\mathrm{card}(A) = C_k^n$. □

**Theorème 3.6.2** (Poisson). *Soit $p_n(k)$ et $p = p_n$ tel que $np_n \to \lambda$. Alors $\forall k$, $p_n(k) \to \frac{\lambda^k}{k!} e^{-\lambda}$.*

*Démonstration.*

$$\begin{aligned} p_n(k) &= \tfrac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \tfrac{1}{k!} n(n-1)...(n-k+1) p^k (1-p)^{n-k} \\ &= \tfrac{1}{k!} \tfrac{(n-1)...(n-k+1)}{n^{k-1}} (np)^k (1-p)^{n-k} \\ &= \tfrac{1}{k} \underbrace{\left( \frac{n-1}{n} ... \frac{n-k+1}{n} \right)}_{A_n} \underbrace{(np)^k}_{B_n} \underbrace{(1-p)^{n-k}}_{C_n} \end{aligned}$$

On alors :

$$A_n = \frac{n-1}{n} \frac{n-2}{n} ... \frac{n-k+1}{n} = \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) ... \left(1 - \frac{k-1}{n}\right) \to 1$$

$$B_n = (np)^k = (np_n)^k \to \lambda^k$$

$$C_n = (1-p)^{n-k} = \left(1 + \frac{x_n}{n}\right)^n \frac{1}{(1 + \frac{x_n}{n})^n}$$

avec $x_n = np_n \to -\lambda \to e^{-\lambda}$. D'où la formule. □

**Theorème 3.6.3** (Théorème local de Moivre-Laplace). *Soit $S_n$ une variable aléatoire qui suit une loi binomiale de paramètre $p$ alors pour $n$ suffisamment grand la variable :*

$$Z_n = \frac{S_n - np}{\sqrt{npq}}$$

*converge en loi vers une loi normale centrée $\mathcal{N}(0,1)$.*

$$p_n(x) = \frac{1}{npq}\frac{1}{\sqrt{2\pi}}e^{-x_{nk}^2/2}$$

*c'est-à-dire :*

$$\max_{k\in I_n}\left|\frac{p_n(k)}{f_n(k)} - 1\right| \to 0$$

*avec $I_n = \{k \,|\, |x_{nk} \leq C\}$ et $x_{nk} = \frac{k-np}{\sqrt{npq}}$.*

# Chapitre 4

# Espérance mathématique

## 4.1 Généralités

### 4.1.1 Aspect discret

**Définition 4.1.1.** Soit $X$ une variable aléatoire vérifiant :

$$\sum_{x_k \in X(\Omega)} |x_k| P(X = x_k) < \infty$$

On appelle espérance mathématique de $X$ le réel $\mathbf{E}X$ défini par :

$$\mathbf{E}X = \sum_{x_k \in X(\Omega)} x_k P(X = x_k)$$

L'espérance de $X$ apparaît ainsi comme le barycentre des valeurs possibles de $X$ pondérées par leurs probabilités de réalisation.

**Propriété 4.1.1.** *1) Si $X = c$ alors $\mathbf{E}X = c$*

*2) Si $X > 0$ alors $\mathbf{E}X \geq 0$*

*3) $\mathbf{E}(aX) = a\mathbf{E}X$ (Linéarité de l'espérance)*

*Démonstration du 3).* $X = a_j \Leftrightarrow aX = aa_j$

$$\mathbf{E}(aX) = \sum_j aa_j p_j = a \sum_j a_j p_j = a\mathbf{E}X$$

□

*4) $\mathbf{E}(X + Y) = \mathbf{E}X + \mathbf{E}Y$*

*Démonstration du 4).* Soit $X = a_i$, $p_i = P\{X = a_i\}$ et $Y = b_j$, $q_j = P\{Y = b_j\}$. Soit :

$$A = \bigcup_j (A \cap \{y = b_j\})$$

On a :

$$p_i = P\{X = a_i\} = \sum_j P\{X = a_i, Y = b_j\}$$

$$q_j = p\{Y = b_j\} = \sum_i P\{X = a_i, Y = b_j\}$$

$$\begin{aligned}
\mathbf{E}X+\mathbf{E}Y &= \sum_i a_i p_i + \sum_j b_j a_j \\
&= \sum_i a_i \left( \sum_j P\{X=a_i, Y=b_j\} \right) + \sum_j b_j \left( \sum_i P\{X=a_i, Y=b_j\} \right) \\
&= \sum_i \sum_j a_i P\{X=a_i, Y=b_j\} + \sum_j \sum_i b_j P\{X=a_i, Y=b_j\} \\
&= \sum_i \sum_j (a_i+b_j) P\{X=a_i, Y=b_j\} = \mathbf{E}(X+Y)
\end{aligned}$$

□

*Remarque.* Si $\sum_i |a_i| p_i = \infty$ alors $X$ n'a pas d'espérance mathématiques.

*5)* $X \geq Y \Rightarrow \mathbf{E}X \geq \mathbf{E}Y$

*Démonstration du 5).*

$$X-Y \geq 0 \Rightarrow \mathbf{E}(X-Y) \geq 0 \Rightarrow \mathbf{E}X - \mathbf{E}Y = 0$$

□

*6)* $|\mathbf{E}X| = \mathbf{E}|X|$

*Démonstration du 6).*

$$-|X| \leq X \leq |X|$$

$$-\mathbf{E}|X| \leq \mathbf{E}X \leq \mathbf{E}|X| \Leftrightarrow |\mathbf{E}X| \leq \mathbf{E}|X|$$

□

**Exemple 4.1.1.** 1) $X = \begin{cases} 0, 1-p \\ 1, p \end{cases}$ , Bernouilli :

$$\mathbf{E}X = 0(1-p) + p = p$$

2) $X \sim \mathcal{B}(n,p)$, $k = 0, 1, ..., n$ :

$$P_n(k) = C_n^k p^k (1-p)^{n-k}$$

$$\mathbf{E}X = \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k}$$

*Remarque.* Si $X$ suit la même loi que $Y$ alors $\mathbf{E}X = \mathbf{E}Y$.

Soit $\nu_n \sim \mathcal{B}(n,p)$, $\mathbf{E}(X) = \mathbf{E}(\nu_n)$. On a que $\nu_n = \varepsilon_1 + ... + \varepsilon_n$ :

$$\begin{aligned}
\mathbf{E}X &= \mathbf{E}\nu_n = \mathbf{E}(\varepsilon_1 + ... + \varepsilon_n) \\
&= \mathbf{E}\varepsilon_1 + ... + \mathbf{E}\varepsilon_n = np = \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k}
\end{aligned}$$

3) Soit $X$ suivant une loi de Poisson, $k = 0, 1, 2, ...$

$$P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$$

On a :

$$\begin{aligned} \mathbf{E}X &= \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} \\ &= \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

### 4.1.2 Aspect continu

**Définition 4.1.2.** Soit $X$ une variable aléatoire continue de densité $p(x) \geq 0$. On a :

$$P\{X \in I\} = \int_I p(x) = 0$$

On note $X^h = hk$ si $X \in [kh, (k+1)h[$. On a alors que :

$$\begin{aligned} \mathbf{E}(X^h) &= \sum_{k\in\mathbb{Z}} kh\{X^h = kh\} = \sum_{k\in\mathbb{Z}} kh \int_{kh}^{kh+h} p(x)dx = \sum_{k\in\mathbb{Z}} \int_{kh}^{kh+h} khp(x)dx \\ &= \sum_{k\in\mathbb{Z}} \int_{-\infty}^{+\infty} \mathbf{1}_{\Delta_k}(x)p(x)dx = \int_{-\infty}^{\infty} \sum_{k\in\mathbb{Z}} \mathbf{1}_{\Delta_k}(x)p(x)dx \\ &= \int_{-\infty}^{\infty} g_h(x)p(x)dx = \int_{-\infty}^{\infty} xp(x)dx \end{aligned}$$

Elle existe si :

$$\int_{-\infty}^{\infty} |x|p(x)dx < \infty$$

### 4.1.3 Retour à l'aspect discret

**Proposition 4.1.2.** *$X$, $Y$ deux variables discrètes et indépendantes. Alors :*

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

*Démonstration.* Soit $X$ de valeurs $a_i$ et de probabilité $p_i$ et $Y$ de valeurs $b_j$ et de probabilité $q_j$ alors :

$$\begin{aligned} \mathbf{E}X\mathbf{E}Y &= \sum_i a_i p_i \sum_j b_j q_j = \sum_i \sum_j a_i b_j p_i q_j \\ &= \sum_i \sum_j a_i b_j P\{X = a_i\}P\{Y = b_j\} = \sum_i \sum_j P\{X = a_i, Y = b_j\} = \mathbf{E}(XY) \end{aligned}$$

□

**Proposition 4.1.3.** *Soit $X$ de valeurs $a_i$ et de probabilité $p_i$. Soit $f : \mathbb{R} \to \mathbb{R}$ tel que pour $Y = f(X)$, $\mathbf{E}Y$ existe. Alors :*

$$\mathbf{E}Y = \mathbf{E}f(X) = \sum_i f(a_i)p_i$$

*Démonstration.* Soit $Y$ de valeurs $b_j$ et de probabilités $q_j$ :

$$\mathbf{E}Y = \sum_j b_j P\{Y = b_j\} \qquad (*)$$

On a :

$$P\{Y = b_j\} = P\{f(X) = b_j\} \qquad (*')$$

On pose :

$$T_j = \{k | f(a_k) = b_j\}$$

Alors :

$$(*') = \sum_{k \in T_j} p_k$$

et :

$$(*) = \sum_j b_j \left( \sum_{k \in T_j} p_k \right) = \sum_j \left( \sum_{k \in T_j} f(a_k) p_k \right) = \sum_i f(a_i) p_i$$

□

**Exemple 4.1.2.** 1) Si $X$ suit la loi uniforme sur l'ensemble fini $\{x_1, ..., x_n\}$, $\mathbf{E}X$ est égale à la moyenne arithmétique des $x_i$.

2) Si $X$ suit la loi géométrique de paramétre $p > 0$ alors :

$$\mathbf{E}X = \frac{1}{p}$$

3) Si $X$ suit la loi hypergéométrique $\mathcal{H}(N, M, n)$ alors :

$$\mathbf{E}X = n\frac{M}{N}$$

Des exemples sont données dans le polycopié de cours.

**Exemple 4.1.3.** 1) $X \sim \mathcal{U}([a, b])$

$$p(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x)$$

$$\mathbf{E}X = \int_{-\infty}^{\infty} x p(x) dx = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{a+b}{2}$$

2) $X \sim \mathcal{N}(a, \sigma^2)$ alors $\mathbf{E}X = a$.

## 4.2 Moments

**Définition 4.2.1.** Les expressions $\mathbf{E}X^n$, $\mathbf{E}|X|^p$, $\mathbf{E}|X - \mathbf{E}X|^p$ sont des moments.

- $\mathbf{E}X^n$ est le moment de $X$ d'ordre $n$.
- $\mathbf{E}|X|^p$ est le moment absolu de $X$ d'ordre $p$.
- $\mathbf{E}|X - \mathbf{E}X|^p$ est le moment absolu centré d'ordre $p$.

**Définition 4.2.2.** On appelle variance de $X$ :

$$\mathrm{Var}\, X = \mathbf{E}|X - \mathbf{E}X|^2$$

**Propriété 4.2.1.** $\operatorname{Var} X = \mathbf{E}(X)^2 - (\mathbf{E}X)^2$

*Démonstration.*

$$\operatorname{Var} X = \mathbf{E}|X - \mathbf{E}X|^2 = \mathbf{E}(X^2 - 2X\mathbf{E}X + (\mathbf{E}X)^2) = \mathbf{E}(X^2) - 2\mathbf{E}X\mathbf{E}X + (\mathbf{E}X)^2$$

□

**Exemple 4.2.1.**

$$\varepsilon = \begin{cases} 0 & 1-p \\ 1 & p \end{cases}, \ \mathbf{E}\varepsilon = p$$

$$\operatorname{Var} \varepsilon = \mathbf{E}(\varepsilon^2) - p^2 = p - p^2 = p(1-p)$$

**Propriété 4.2.2.** $\operatorname{Var} X \geq 0$ *et* $\operatorname{Var} X = 0 \Leftrightarrow X = c$

**Propriété 4.2.3.** $\operatorname{Var}(cX) = c^2 \operatorname{Var} X$

*Démonstration.*

$$\mathbf{E}(cX - \mathbf{E}(cX)|^2 = \mathbf{E}|c - (X - \mathbf{E}X)|^2 = c^2\mathbf{E}|X - \mathbf{E}X|^2$$

□

**Propriété 4.2.4.** $\operatorname{Var}(X + a) = \operatorname{Var} X$

**Propriété 4.2.5.** *$X, Y$ indépendantes alors :* $\operatorname{Var}(X + Y) = \operatorname{Var} X + \operatorname{Var} Y$

*Démonstration.* Soit $a = \mathbf{E}X$ et $b = \mathbf{E}Y$

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X + Y - (a + b)) = \operatorname{Var}(X - a + Y - b) \qquad (*)$$

On note $X_1 = X - a$ et $Y_1 = Y - b$.

$$(*) = \mathbf{E}(X_1 + Y_1)^2 = \mathbf{E}X_1^2 + \mathbf{E}Y_1^2 + 2\mathbf{E}|X_1Y_1|$$

$$= \mathbf{E}X_1^2 + \mathbf{E}Y_1^2 = \operatorname{Var} X + \operatorname{Var} Y$$

□

**Exemple 4.2.2.** $X \sim \mathcal{B}(n, p)$, $\mathbf{E}X = np$

$$\operatorname{Var} X = \sum_{k=0}^{n} |k - np|^2 C_n^k p^k (1-p)^{n-k}$$

On a que $X \sim \nu_n$

$$\operatorname{Var} X = \operatorname{Var}(\nu_n) = \operatorname{Var}(\varepsilon_1 + ... + \varepsilon_n) = \operatorname{Var} \varepsilon_1 + ... + \operatorname{Var} \varepsilon_n = np(1-p)$$

## 4.3 Inégalité de Markov et Tchebychev

**Proposition 4.3.1** (Inégalité de Markov)**.** *Soit* $X \geq 0$*, alors* $\forall t \geq 0$

$$P\{X \leq t\} \leq \frac{\mathbf{E}X}{t}$$

*Démonstration dans le cas discret.* Soit $X$ de valeurs $a_i$ et de probabilité $p_i$. On a que $a_i \geq 0$

$$P\{X \leq t\} = \sum_{i|a_i \geq t} p_i \leq \sum_{i|a_i \geq t} \frac{a_i}{t} p_i \leq \frac{1}{t} \sum_i a_i p_i = \frac{\mathbf{E}X}{t}$$

□

**Proposition 4.3.2** (Inégalité de Tchebychev)**.** *Soit* $X$ *une variable alétaoire tel que* $\mathbf{E}X$ *et* $\operatorname{Var} X$ *existent alors* $\forall t \geq 0$ *:*

$$P\{|X - \mathbf{E}X| \geq t\} = \frac{\operatorname{Var} X}{t^2}$$

*Démonstration.*

$$P\{|X - \mathbf{E}X| \geq t\} = P\{\underbrace{|X - \mathbf{E}X|^2}_{Y} \geq \underbrace{t^2}_{s}\} = \frac{\mathbf{E}Y}{s} = \frac{\mathbf{E}|X - \mathbf{E}X|^2}{t^2} = \frac{\operatorname{Var} X}{t^2}$$

□

# Chapitre 5

# Loi des grands nombres

## 5.1 Loi des grands nombres

**Theorème 5.1.1** (Loi des grands nombres)**.** *Soit* $(X_n)$ *variables aléatoires de même loi,* $\mathbf{E}X_k = a$*,* $\operatorname{Var} X_k = \sigma^2 < \infty$*. Alors :*

$$\frac{S_n}{n} \xrightarrow{P} a$$

*Notation.* $\sum\limits_{k=1}^{n} X_k = S_n$

**Définition 5.1.1.** $Y_n \xrightarrow{P} Y$ si $\forall \varepsilon > 0$ :

$$P\{|Y_n - Y| > \varepsilon\} \xrightarrow[n\to+\infty]{} 0$$

*Démonstration.* Soit $\varepsilon > 0$

$$P\left\{\left|\frac{S_n}{n} - a\right| < \varepsilon\right\} \leq \mathbf{E}\left(\frac{S_n}{n} - a\right) \leq \frac{\operatorname{Var}\left(\frac{S_n}{n} - a\right)}{\varepsilon^2} = \frac{\frac{1}{n^2}\operatorname{Var} S_n}{\varepsilon^2} = \frac{1}{\varepsilon^2}\frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}(X_i)$$

$$= \frac{1}{\varepsilon^2}\frac{n\sigma^2}{n} = \frac{\sigma^2}{\varepsilon^2}\frac{1}{n} \xrightarrow[n\to+\infty]{} 0$$

□

**Exemple 5.1.1.** Soit $A$ de valeurs $\varepsilon_i$ de probabilité $p$ :

$$\varepsilon_i = \begin{cases} 1 \text{ si } A \text{ à la } i\text{-ème épreuve} \\ 0 \text{ si } A^C \text{ à la } i\text{-ème épreuve} \end{cases}$$

$$\frac{n_A}{n} = \frac{\varepsilon_1 + \ldots + \varepsilon_n}{n} \xrightarrow{P} \mathbf{E}\varepsilon_1 = p = P(A)$$