

# Probabilités discrètes. Variables aléatoires discrètes

## Plan du chapitre

<b>I - Probabilités discrètes</b> .....	<b>page 2</b>
1) Tribus .....	page 2
2) Espaces probabilisés .....	page 3
3) Propriétés des probabilités .....	page 3
4) Evénements négligeables, événements presque sûrs .....	page 6
5) Systèmes complets d'événements .....	page 6
6) Détermination d'une probabilité .....	page 7
<b>II - Probabilités conditionnelles. Indépendance</b> .....	<b>page 8</b>
1) Probabilités conditionnelles .....	page 8
1-a) Définition .....	page 8
1-b) Formule des probabilités composées .....	page 8
1-c) Formule des probabilités totales .....	page 9
1-d) Formule de BAYES .....	page 9
2) Evénements indépendants .....	page 9
2-a) Cas de deux événements .....	page 9
2-b) Familles d'événements mutuellement indépendants .....	page 10
<b>III - Variables aléatoires discrètes</b> .....	<b>page 11</b>
1) Loi de probabilité d'une variable aléatoire .....	page 11
2) Fonctions de variables aléatoires .....	page 11
3) Couples de variables aléatoires. n-uplets de variables aléatoires .....	page 12
4) Variables aléatoires indépendantes .....	page 12
<b>IV - Espérance, variance</b> .....	<b>page 12</b>
1) Espérance d'une variable aléatoire .....	page 12
2) Variance d'une variable aléatoire .....	page 12
3) Covariance d'un couple de variables aléatoires .....	page 12
4) Loi faible des grands nombres .....	page 12
<b>V - Variables usuelles</b> .....	<b>page 12</b>
1) Loi géométrique de paramètre $p$ .....	page 12
2) Loi de POISSON de paramètre $\lambda$ .....	page 12
2-a) Définition .....	page 12
2-b) Approximation de la loi binomiale par la loi de POISSON .....	page 12
<b>VI - Fonctions génératrices</b> .....	<b>page 12</b>
1) Fonction génératrice d'une variable aléatoire .....	page 12
2) Fonctions génératrices des variables usuelles .....	page 12
2-a) Fonction génératrice d'un variable aléatoire de BERNOULLI .....	page 12
2-b) Fonction génératrice d'un variable aléatoire binomiale .....	page 12
2-c) Fonction génératrice d'un variable aléatoire géométrique .....	page 12
2-d) Fonction génératrice d'un variable aléatoire de POISSON .....	page 12

# I - Probabilités discrètes

## 1) Tribus

Il s'agit de généraliser la notions d'événement exposée en maths sup. En maths sup, l'univers des possibles  $\Omega$  est fini et un événement (qui aura ensuite une probabilité une fois celle-ci définie) est une partie quelconque de  $\Omega$ . La probabilité d'un événement est alors la somme finie des probabilités des événements élémentaires constituant cet événement.

En maths spé, nous passons à la situation plus générale où  $\Omega$  est **au plus dénombrable** c'est-à-dire fini ou infini et dénombrable. Les probabilités à calculer seront alors des sommes (discrètes) de séries à termes positifs. Quand on continuera à généraliser, les sommes discrètes deviennent des sommes continues c'est-à-dire des intégrales et on sait qu'il existe des fonctions non intégrables même sur un segment. Il n'est donc plus possible à terme, de calculer la probabilité de n'importe quel partie de  $\Omega$ . La définition 1 ci-dessous, met en place la notion qui définit axiomatiquement les événements dont on calculera la probabilité : c'est la notion de **tribu**. En maths spé, cette notion n'est pas encore absolument indispensable dans la pratique mais son introduction est prévue par le programme officiel.

**DÉFINITION 1.** Soit  $\Omega$  un ensemble non vide au plus dénombrable, c'est-à-dire fini ou infini et dénombrable.

Une **tribu** ou  **$\sigma$ -algèbre** sur  $\Omega$  est un sous-ensemble  $\mathcal{A}$  de  $\mathcal{P}(\Omega)$  vérifiant :

- 1)  $\mathcal{A} \neq \emptyset$ ,
- 2)  $\forall A \in \mathcal{A}, {}^c A \in \mathcal{A}$  (stabilité par passage au complémentaire),
- 3)  $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}, \bigcup_{n=0}^{+\infty} A_n \in \mathcal{A}$  (stabilité par réunion dénombrable).

Un élément de  $\mathcal{A}$  s'appelle alors un **événement** et le couple  $(\Omega, \mathcal{A})$  est un **espace probabilisable**.

### Exemples.

- $\mathcal{A} = \{\emptyset, \Omega\}$  est une tribu sur  $\Omega$  appelée **tribu grossière**.
- $\mathcal{A} = \mathcal{P}(\Omega)$  est une tribu sur  $\Omega$ .
- Soit  $A \in \Omega$ .  $\mathcal{A} = \{\emptyset, A, {}^c A, \Omega\}$  est une tribu sur  $\Omega$ .
- On peut montrer qu'une intersection de tribus est une tribu. En conséquence, si  $\mathcal{B}$  est un sous-ensemble quelconque de  $\mathcal{P}(\Omega)$ , on peut montrer qu'il existe une plus petite tribu (au sens de l'inclusion) contenant  $\mathcal{B}$ . Elle est obtenue comme intersection de toutes les tribus contenant  $\mathcal{B}$  et s'appelle la **tribu engendrée par  $\mathcal{B}$** .  $\square$

Les axiomes de la définition 1 entraînent un certain nombre de propriétés obligatoirement vérifiées par une tribu :

**Théorème 1.** Soit  $\Omega$  un ensemble non vide au plus dénombrable, c'est-à-dire fini ou infini et dénombrable. Soit  $\mathcal{A}$  une tribu sur  $\Omega$ .

- 1)  $\Omega \in \mathcal{A}$  et  $\emptyset \in \mathcal{A}$ .
- 2)  $\mathcal{A}$  est stable par réunion finie :  $\forall n \in \mathbb{N}^*, \forall (A_k)_{1 \leq k \leq n} \in \mathcal{A}^n, \bigcup_{k=1}^n A_k \in \mathcal{A}$ .
- 3)  $\mathcal{A}$  est stable par intersection finie ou dénombrable.
- 4)  $\forall (A, B) \in \mathcal{A}^2, B \setminus A \in \mathcal{A}$ .

### Démonstration.

1) D'après l'axiome 1,  $\mathcal{A} \neq \emptyset$ . Donc, il existe  $A \in \mathcal{A}$ . D'après l'axiome 2,  ${}^c A \in \mathcal{A}$ . Donc, si on pose  $A_1 = {}^c A$  et pour  $n \geq 2$ ,  $A_n = A$ , alors d'après l'axiome 3,  $\Omega = \bigcup_{n=1}^{+\infty} A_n \in \mathcal{A}$ .

Mais alors, d'après l'axiome 2,  $\emptyset = {}^c \Omega \in \mathcal{A}$ .

2) Soient  $n \in \mathbb{N}^*$  puis  $(A_k)_{1 \leq k \leq n} \in \mathcal{A}^n$ . Pour  $k \in \mathbb{N}^*$ , on pose  $B_k = A_k$  si  $1 \leq k \leq n$  et  $B_k = \emptyset$  si  $k > n$ . Alors,  $\bigcup_{k=1}^n A_k = \bigcup_{k=1}^{+\infty} B_k \in \mathcal{A}$ .

3) Soit  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ . Pour  $n \in \mathbb{N}$ , on pose  $B_n = {}^c A_n$ . Alors, pour tout  $n \in \mathbb{N}$ ,  $B_n \in \mathcal{A}$  puis

$$\bigcap_{n=0}^{+\infty} A_n = \bigcap_{n=0}^{+\infty} {}^c B_n = {}^c \left( \bigcup_{n=0}^{+\infty} B_n \right) \in \mathcal{A}.$$

4) Soit  $(A, B) \in \mathcal{A}^2$ . Alors,  $B \setminus A = B \cap {}^c A \in \mathcal{A}$ .

Le vocabulaire sur les événements mis en place en première année se généralise à l'identique. Par exemple, deux événements  $A$  et  $B$  sont **disjoints** ou **incompatibles** si et seulement si  $A \cap B = \emptyset$ ,  $\Omega$  est l'**événement certain** et  $\emptyset$  est l'**événement impossible** ...

## 2) Espaces probabilisés

On définit maintenant la notion de probabilité sur un espace probabilisable :

**DÉFINITION 2.** Soit  $(\Omega, \mathcal{A})$  un espace probabilisable au plus dénombrable (c'est-à-dire soit  $\Omega$  un ensemble non vide au plus dénombrable et soit  $\mathcal{A}$  une tribu sur  $\Omega$ ).

Une **probabilité** sur  $(\Omega, \mathcal{A})$  est une application  $\mathbb{P}$  de  $\mathcal{A}$  dans  $[0, +\infty[$  vérifiant de plus :

- 1)  $\mathbb{P}(\Omega) = 1$ ,
- 2) Pour toute suite  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$  d'événements deux à deux disjoints,

$$\mathbb{P} \left( \bigcup_{n=0}^{+\infty} A_n \right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n) \quad (\sigma\text{-additivité de } \mathbb{P}).$$

Un **espace probabilisé** au plus dénombrable est un triplet  $(\Omega, \mathcal{A}, \mathbb{P})$  où  $(\Omega, \mathcal{A})$  est un espace probabilisable au plus dénombrable et  $\mathbb{P}$  est une probabilité sur  $(\Omega, \mathcal{A})$ .

**Commentaire.** La définition ci-dessus contient implicitement le fait que pour toute suite  $(A_n)_{n \in \mathbb{N}}$  d'événements deux à deux disjoints, la série de terme général  $\mathbb{P}(A_n)$ ,  $n \in \mathbb{N}$ , converge. En particulier, il est obligatoire que pour toute suite  $(A_n)_{n \in \mathbb{N}}$  d'événements deux à deux disjoints, on ait  $\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = 0$ .  $\square$

## 3) Propriétés des probabilités

On a déjà toutes les formules de maths sup qui se généralise à l'identique :

**Théorème 2.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable.

- 1)  $\mathbb{P}(\emptyset) = 0$ .
- 2) a)  $\forall (A, B) \in \mathcal{A}^2, (A \cap B = \emptyset \Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B))$ .  
b) Plus généralement, pour tout  $n \in \mathbb{N}^*$  puis tout  $n$ -uplet  $(A_k)_{1 \leq k \leq n} \in \mathcal{A}^n$  d'événements deux à deux disjoints,

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

- 3) a)  $\forall A \in \mathcal{A}, \mathbb{P}({}^c A) = 1 - \mathbb{P}(A)$ .  
b)  $\forall A \in \mathcal{A}, \mathbb{P}(A) \in [0, 1]$ .
- 4) a)  $\forall (A, B) \in \mathcal{A}^2, A \subset B \Rightarrow \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ .  
b)  $\forall (A, B) \in \mathcal{A}^2, A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$ .
- 5)  $\forall (A, B) \in \mathcal{A}^2, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ .
- 6) Pour tout  $n \geq 2$  puis pour tout  $(A_k)_{1 \leq k \leq n} \in \mathcal{A}^n$ ,  $\mathbb{P} \left( \bigcup_{k=1}^n A_k \right) \leq \sum_{k=1}^n \mathbb{P}(A_k)$ .

**Démonstration.**

1) Pour  $n \in \mathbb{N}$ , on pose  $A_n = \emptyset$ .

$$\mathbb{P}(\emptyset) = \mathbb{P} \left( \bigcup_{n=0}^{+\infty} A_n \right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(\emptyset)$$

et donc  $\sum_{n=1}^{+\infty} \mathbb{P}(\emptyset) = 0$  puis  $\mathbb{P}(\emptyset) = 0$ .

2) a) Soit  $(A, B) \in \mathcal{A}^2$  tel que  $A \cap B = \emptyset$ . On pose  $A_0 = A$ ,  $A_1 = B$  et pour  $n \geq 2$ ,  $A_n = \emptyset$ . Les  $A_n$ ,  $n \in \mathbb{N}$ , sont deux à deux disjoints et donc

$$\mathbb{P}(A \cup B) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \mathbb{P}(A) + \mathbb{P}(B).$$

b) Par récurrence sur  $n$ . Le cas  $n = 2$  est a).

3) a) Soit  $A \in \mathcal{A}$ . Puisque  $A \cap {}^c A = \emptyset$ ,

$$\mathbb{P}(A) + \mathbb{P}({}^c A) = \mathbb{P}(A \cup {}^c A) = \mathbb{P}(\Omega) = 1.$$

b) Soit  $A \in \mathcal{A}$ .  $\mathbb{P}(A) = 1 - \mathbb{P}({}^c A) \leq 1$ .

4) a) Soit  $(A, B) \in \mathcal{A}^2$  tel que  $A \subset B$ . Alors  $B = (B \setminus A) \cup A$  avec  $(B \setminus A) \cap A = \emptyset$ . Donc,

$$\mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(A \cup (B \setminus A)) = \mathbb{P}(B).$$

b) Soit  $(A, B) \in \mathcal{A}^2$  tel que  $A \subset B$ .  $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$ .

5) Soit  $(A, B) \in \mathcal{A}^2$ .  $(B \setminus (A \cap B)) \cap A = \emptyset$  avec  $A \cap B \subset B$  et donc

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (B \setminus (A \cap B))) = \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

6) Montrons par récurrence que pour tout  $n \geq 2$  puis pour tout  $(A_k)_{1 \leq k \leq n} \in \mathcal{A}^n$ ,  $\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \mathbb{P}(A_k)$ .

• Soient  $A_1$  et  $A_2$  deux événements.

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2).$$

• Soit  $n \geq 2$ . Supposons que pour tout  $(A_k)_{1 \leq k \leq n} \in \mathcal{A}^n$ ,  $\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \mathbb{P}(A_k)$ . Soit  $(A_k)_{1 \leq k \leq n+1} \in \mathcal{A}^{n+1}$ .

$$\begin{aligned} \mathbb{P}\left(\bigcup_{k=1}^{n+1} A_k\right) &= \mathbb{P}\left(\left(\bigcup_{k=1}^n A_k\right) \cup A_{n+1}\right) \\ &\leq \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) + \mathbb{P}(A_{n+1}) \quad (\text{d'après le cas } n = 2) \\ &\leq \sum_{k=1}^{n+1} \mathbb{P}(A_k) \quad (\text{par hypothèse de récurrence}). \end{aligned}$$

On a montré par récurrence que pour tout  $n \geq 2$  puis pour tout  $(A_k)_{1 \leq k \leq n} \in \mathcal{A}^n$ ,  $\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \mathbb{P}(A_k)$ .

En maths spé, on complète ces propriétés avec :

**Théorème 3.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements.

1) (continuité croissante) Si la suite  $(A_n)_{n \in \mathbb{N}}$  est croissante pour l'inclusion, alors la suite numérique  $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$  converge et de plus,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right).$$

2) (continuité décroissante) Si la suite  $(A_n)_{n \in \mathbb{N}}$  est décroissante pour l'inclusion, alors la suite numérique  $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$  converge et de plus,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{k=0}^{+\infty} A_k\right).$$

### Démonstration.

1) Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements, croissante pour l'inclusion. Donc,  $\forall k \in \mathbb{N}, A_k \subset A_{k+1}$ .

On se ramène au cas d'événements deux à deux disjoints : on pose  $B_0 = A_0$  et pour  $k \in \mathbb{N}$ , on pose  $B_k = A_k \setminus A_{k-1}$ . Les  $B_k$  sont deux à deux disjoints et, par récurrence,

$$\forall n \in \mathbb{N}^*, \bigcup_{k=0}^n A_k = A_n = A_0 \cup \left( \bigcup_{k=1}^n (A_k \setminus A_{k-1}) \right) = \bigcup_{k=0}^n B_k.$$

Ceci fournit déjà  $\bigcup_{k=0}^{+\infty} B_k = \bigcup_{k=0}^{+\infty} A_k$  et d'autre part, pour  $n \in \mathbb{N}^*$ ,

$$\mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{k=0}^n B_k\right) = \sum_{k=0}^n \mathbb{P}(B_k).$$

Puisque  $(B_k)_{k \in \mathbb{N}}$  est une suite d'événements deux à deux disjoints, la suite  $(\mathbb{P}(A_n))_{n \in \mathbb{N}} = \left(\sum_{k=0}^n \mathbb{P}(B_k)\right)_{n \in \mathbb{N}}$  converge et

$$\lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \lim_{n \rightarrow +\infty} \left(\sum_{k=0}^n \mathbb{P}(B_k)\right) = \sum_{k=0}^{+\infty} \mathbb{P}(B_k) = \mathbb{P}\left(\bigcup_{k=0}^{+\infty} B_k\right) = \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right).$$

2) Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements, décroissante pour l'inclusion. Si pour  $n \in \mathbb{N}$ , on pose  $A'_n = {}^c A_n$ , la suite  $(A'_n)_{n \in \mathbb{N}}$  est une suite d'événements, croissante pour l'inclusion. D'après 1), la suite  $(\mathbb{P}(A'_n))_{n \in \mathbb{N}}$  converge et donc la suite  $(\mathbb{P}(A_n))_{n \in \mathbb{N}} = (1 - \mathbb{P}(A'_n))_{n \in \mathbb{N}}$  converge. De plus,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbb{P}(A_n) &= 1 - \lim_{n \rightarrow +\infty} \mathbb{P}(A'_n) \\ &= 1 - \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A'_k\right) = 1 - \mathbb{P}\left(\bigcup_{k=0}^{+\infty} ({}^c A_k)\right) = 1 - \mathbb{P}\left({}^c \left(\bigcap_{k=0}^{+\infty} A_k\right)\right) \\ &= \mathbb{P}\left(\bigcap_{k=0}^{+\infty} A_k\right). \end{aligned}$$

**Commentaire.** Dans la démonstration précédente, dans le cas d'une suite d'événements  $(A_n)_{n \in \mathbb{N}}$  croissante pour l'inclusion, on aurait également pu constater que la suite  $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$  était une suite numérique croissante, majorée par 1 et donc convergente.  $\square$

**Théorème 4.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_n)_{n \in \mathbb{N}}$  une suite d'événements. Alors,

$$\mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

**Démonstration.** Pour  $n \in \mathbb{N}$ , posons  $A'_n = \bigcup_{k=0}^n A_k$  de sorte que pour  $n \in \mathbb{N}$ ,  $\bigcup_{k=0}^n A'_k = \bigcup_{k=0}^n A_k$ .

La suite  $(A'_n)_{n \in \mathbb{N}}$  est une suite croissante pour l'inclusion. Donc, la suite  $(\mathbb{P}(A'_n))_{n \in \mathbb{N}}$  converge vers  $\mathbb{P}\left(\bigcup_{k=0}^{+\infty} A'_k\right) = \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right)$ .

D'autre part, pour  $n \in \mathbb{N}$ , d'après le théorème 3,

$$\mathbb{P}(A'_n) = \mathbb{P}\left(\bigcup_{k=0}^n A_k\right) \leq \sum_{k=0}^n \mathbb{P}(A_k) \quad (*).$$

La série numérique de terme général  $\mathbb{P}(A_k)$ ,  $n \in \mathbb{N}$ , à termes positifs, converge dans  $\overline{\mathbb{R}}$  vers un réel positif ou vers  $+\infty$ . Quand  $n$  tend vers  $+\infty$  dans l'inégalité (\*), on obtient

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

(le second membre étant réel ou infini).

**Remarque.** Dans la planche d'exercices n° 37 de maths sup (« Dénombrables »), on détaille la *formule du crible* qui donne le calcul de  $\text{card} \left( \bigcup_{k=0}^n A_k \right)$ . Cette démonstration et la formule qui en résulte, peuvent être reproduite à l'identique pour les probabilités en remplaçant  $\text{card}$  par  $\mathbb{P}$ .

#### 4) Événements négligeables, événements presque sûrs

$\emptyset$  est l'événement impossible. Sa probabilité est nulle. Mais il dorénavant possible de trouver des événements non vides de probabilité nulle. Par exemple, si on choisit un entier au hasard, la probabilité que cet entier soit 0 est nulle. De même,  $\mathbb{P}(\Omega) = 1$  mais  $\mathbb{P}(A) = 1 \not\Rightarrow A = \Omega$ . Ceci nous conduit à poser la définition suivante :

**DÉFINITION 4.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $A \in \mathcal{A}$ .

$A$  est un **événement négligeable** si et seulement si  $\mathbb{P}(A) = 0$ .

$A$  est un **événement presque sûr** si et seulement si  $\mathbb{P}(A) = 1$ .

**Théorème 5.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ .

- Si  $B$  est négligeable, alors  $\mathbb{P}(A \cap B) = 0$  et  $\mathbb{P}(A \cup B) = \mathbb{P}(A)$ .
- Si  $B$  est presque sûr, alors  $\mathbb{P}(A \cap B) = \mathbb{P}(A)$  et  $\mathbb{P}(A \cup B) = 1$ .

**Démonstration.**

- Supposons  $\mathbb{P}(B) = 0$ . Puisque  $A \cap B \subset B$ ,  $\mathbb{P}(A \cap B) \leq \mathbb{P}(B) = 0$  puis  $\mathbb{P}(A \cap B) = 0$ . Mais alors,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(A).$$

- Supposons  $\mathbb{P}(B) = 1$ . Puisque  $B \subset A \cup B$ ,  $1 = \mathbb{P}(B) \leq \mathbb{P}(A \cup B)$  puis  $\mathbb{P}(A \cup B) = 1$ . Mais alors,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = \mathbb{P}(A).$$

**Théorème 6.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable.

Une réunion finie ou dénombrable d'événements négligeables est un événement négligeable.

Une réunion finie ou dénombrable d'événements presque sûrs est un événement presque sûr.

**Démonstration.**

- Soit  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ . Supposons chaque  $A_n$ ,  $n \in \mathbb{N}$ , négligeable. D'après le théorème 4 de la page précédente,

$$\mathbb{P} \left( \bigcup_{n=0}^{+\infty} A_n \right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \sum_{n=0}^{+\infty} 0 = 0.$$

et donc  $\mathbb{P} \left( \bigcup_{n=0}^{+\infty} A_n \right) = 0$ . Ceci démontre le résultat pour une réunion dénombrable d'événements négligeables. Le cas d'une famille finie d'événements négligeables s'obtient en prenant les  $A_n$  vides à partir d'un certain rang.

- Soit  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ . Supposons chaque  $A_n$ ,  $n \in \mathbb{N}$ , presque sûr. On en déduit que chaque  ${}^c A_n$  est négligeable et donc que  ${}^c \left( \bigcap_{n=0}^{+\infty} A_n \right) = \bigcup_{n=0}^{+\infty} {}^c A_n$  est négligeable puis que  $\bigcap_{n=0}^{+\infty} A_n$  est un événement presque sûr.

#### 5) Systèmes complets d'événements

**DÉFINITION 5.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ .

$(A_n)_{n \in \mathbb{N}}$  est un **système complet d'événements** si les événements  $A_n$ ,  $n \in \mathbb{N}$ , sont deux à deux disjoints et

$\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = 1$  (ou encore  $\mathbb{P} \left( \bigcup_{n=0}^{+\infty} A_n \right) = 1$  ou encore  $\bigcup_{n=0}^{+\infty} A_n$  est un événement presque sûr).

L'intérêt d'un système complet d'événements réside dans le théorème suivant :

**Théorème 7.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$  un système complet d'événements. Soit  $B \in \mathcal{A}$ . Alors,

$$\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B \cap A_n).$$

**Démonstration.** Puisque  $\bigcup_{n=0}^{+\infty} A_n$  est presque sûr, on a  $\mathbb{P}\left(B \cap \bigcup_{n=0}^{+\infty} A_n\right) = \mathbb{P}(B)$  d'après le théorème 5. Mais  $B \cap \bigcup_{n=0}^{+\infty} A_n = \bigcup_{n=0}^{+\infty} (B \cap A_n)$  et de plus, les événements  $B \cap A_n$ ,  $n \in \mathbb{N}$ , sont deux à deux disjoints. On en déduit que

$$\mathbb{P}(B) = \mathbb{P}\left(B \cap \bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(B \cap A_n).$$

## 6) Détermination d'une probabilité

Dans ce paragraphe, on se place dans le cas particulier où  $\mathcal{A} = \mathcal{P}(\Omega)$ .

**Théorème 8.** Soit  $\Omega$  un univers non vide, fini ou dénombrable.

1) Soit  $\mathbb{P}$  une probabilité sur l'espace probabilisable  $(\Omega, \mathcal{P}(\Omega))$ . Pour  $\omega \in \Omega$ , on pose  $p_\omega = \mathbb{P}(\{\omega\})$ . Alors,  $(p_\omega)_{\omega \in \Omega}$  est une famille de réels positifs, sommable, de somme 1. De plus, pour  $A \in \mathcal{P}(\Omega)$ ,

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega.$$

2) Inversement, soit  $(p_\omega)_{\omega \in \Omega}$  une famille de réels positifs indexée par  $\Omega$ , sommable, de somme 1. Il existe une probabilité  $\mathbb{P}$  et une seule sur l'espace probabilisable  $(\Omega, \mathcal{P}(\Omega))$  telle que

$$\forall \omega \in \Omega, \mathbb{P}(\{\omega\}) = p_\omega.$$

**Démonstration.**

1) Les  $p_\omega$ ,  $\omega \in \Omega$ , sont des réels positifs. Ensuite,  $\Omega = \bigcup_{\omega \in \Omega} \{\omega\}$  et donc, puisque les  $\{\omega\}$ ,  $\omega \in \Omega$ , sont deux à deux disjoints,

$$1 = \mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} p_\omega.$$

Ceci montre que la famille dénombrable  $(p_\omega)_{\omega \in \Omega}$  est sommable, de somme 1. Dit autrement,  $(\{\omega\})_{\omega \in \Omega}$  est un système complet d'événements. D'après le théorème 7, pour  $A \in \mathcal{P}(\Omega)$ ,

$$\mathbb{P}(A) = \sum_{\omega \in \Omega} \mathbb{P}(A \cap \{\omega\}) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \sum_{\omega \in A} p_\omega.$$

2) Soit  $(p_\omega)_{\omega \in \Omega}$  une famille de réels positifs indexée par  $\Omega$ , sommable, de somme 1. D'après 1), si  $\mathbb{P}$  est une probabilité sur  $(\Omega, \mathcal{P}(\Omega))$  telle que  $\forall \omega \in \Omega, \mathbb{P}(\{\omega\}) = p_\omega$ , on a nécessairement

$$\forall A \in \mathcal{P}(\Omega), \mathbb{P}(A) = \sum_{\omega \in A} p_\omega,$$

ce qui montre l'unicité de  $\mathbb{P}$ . Il s'agit alors de vérifier que  $\mathbb{P}$  ainsi définie, est bien une probabilité.

Pour chaque  $A \in \mathcal{P}(\Omega)$ , la famille  $(p_\omega)_{\omega \in A}$  est sommable en tant que sous-famille de la famille sommable  $(p_\omega)_{\omega \in \Omega}$  et de plus

$$0 \leq \mathbb{P}(A) = \sum_{\omega \in A} p_\omega \leq \sum_{\omega \in \Omega} p_\omega = 1.$$

Donc,  $\mathbb{P}$  est une application de  $\mathcal{P}(\Omega)$  dans  $[0, 1]$ . Ensuite,

$$\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p_{\omega} = 1.$$

Soit enfin  $(A_n)_{n \in \mathbb{N}}$  une suite d'éléments de  $\mathcal{P}(\Omega)$  deux à deux disjoints. L'événement  $A = \bigcup_{n=0}^{+\infty} A_n$  est une partie de

l'ensemble dénombrable  $\Omega$  et donc  $\bigcup_{n=0}^{+\infty} A_n$  est au plus dénombrable. On peut poser  $\bigcup_{n=0}^{+\infty} A_n = \{\omega_k, k \in I\}$  où l'ensemble des indices  $I$  est soit vide, soit un ensemble du type  $[[1, p]]$ ,  $p \in \mathbb{N}^*$ , soit  $\mathbb{N}$ . Pour chaque  $n$ , on note  $I_n$  l'ensemble des indices  $k \in I$  tels que  $\omega \in A_n$ . Les ensembles  $I_n$  sont deux à deux disjoints et leur réunion est  $I$ .

Puisque la famille  $(p_{\omega})_{\omega \in \Omega}$  est sommable, le théorème de sommation par paquets permet d'écrire

$$\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \sum_{n=0}^{+\infty} \left( \sum_{k \in I_n} p_{\omega_k} \right) = \sum_{k \in I} p_{\omega_k} = \mathbb{P}(A).$$

Puisque  $\mathbb{P}$  est une probabilité sur l'espace  $(\Omega, \mathcal{P}(\Omega))$  et que  $\forall \omega \in \Omega, \mathbb{P}(\omega) = p_{\omega}$ ,  $\mathbb{P}$  convient.

## II - Probabilités conditionnelles. Indépendance

### 1) Probabilités conditionnelles

#### a) Définition

**DÉFINITION 6.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ .

Si  $\mathbb{P}(A) \neq 0$ , la **probabilité de B sachant A**, notée  $\mathbb{P}_A(B)$ , est

$$\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

**Théorème 9.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $A \in \mathcal{A}$  tel que  $\mathbb{P}(A) \neq 0$ .

$\mathbb{P}_A$  est une probabilité sur l'espace probabilisable  $(\Omega, \mathcal{A})$ .

#### Démonstration.

• Puisque  $\mathbb{P}(A) \neq 0$ , pour tout  $B \in \mathcal{A}$ ,  $\mathbb{P}_A(B)$  existe dans  $\mathbb{R}$  et de plus  $\mathbb{P}_A(B) \geq 0$ .

•  $\mathbb{P}_A(\Omega) = \frac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$ .

• Soit  $(B_n)_{n \in \mathbb{N}}$  une suite d'événements deux à deux disjoints.

$$\mathbb{P}_A \left( \bigcup_{n=0}^{+\infty} B_n \right) = \frac{1}{\mathbb{P}(A)} \times \mathbb{P} \left( \bigcup_{n=0}^{+\infty} B_n \right) = \frac{1}{\mathbb{P}(A)} \sum_{n=0}^{+\infty} \mathbb{P}(B_n) = \sum_{n=0}^{+\infty} \mathbb{P}_A(B_n).$$

On a montré que  $\mathbb{P}_A$  est une probabilité sur l'espace probabilisable  $(\Omega, \mathcal{A})$ .

On note que si  $A$  est un événement tel que  $\mathbb{P}(A) \neq 0$  alors  $\mathbb{P}_A(A) = 1$ .

#### b) Formule des probabilités composées

Un résultat immédiat est :

**Théorème 10.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ .

Si  $\mathbb{P}(A) \neq 0$ ,  $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B)$ .

Si  $\mathbb{P}(B) \neq 0$ ,  $\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}_B(A)$ .

Plus généralement,



**Théorème 11.** (formule des probabilités composées)

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$  une suite d'événements tels que  $\forall n \in \mathbb{N}$ ,  $\mathbb{P}(A_0 \cap \dots \cap A_n) \neq 0$ .

Pour tout  $n \in \mathbb{N}^*$ ,

$$\mathbb{P}(A_0 \cap \dots \cap A_n) = \mathbb{P}_{A_0 \cap \dots \cap A_{n-1}}(A_n) \times \mathbb{P}_{A_0 \cap \dots \cap A_{n-2}}(A_{n-1}) \times \dots \times \mathbb{P}_{A_0}(A_1) \times \mathbb{P}(A_0).$$

**Démonstration.** Pour  $n \in \mathbb{N}$ , posons  $p_n = \mathbb{P}(A_0 \cap \dots \cap A_n)$ . Par hypothèse, pour tout  $n \in \mathbb{N}$ ,  $p_n \neq 0$ .

Soit  $n \in \mathbb{N}^*$ . Pour  $k \in \llbracket 0, n-1 \rrbracket$ ,

$$\mathbb{P}_{A_0 \cap \dots \cap A_k}(A_{k+1}) = \frac{\mathbb{P}(A_0 \cap \dots \cap A_{k+1})}{\mathbb{P}(A_0 \cap \dots \cap A_k)} = \frac{p_{k+1}}{p_k}.$$

Par suite,

$$\begin{aligned} \mathbb{P}(A_0 \dots \cap A_n) &= p_n \\ &= p_0 \prod_{k=0}^{n-1} \frac{p_{k+1}}{p_k} \text{ (produit télescopique)} \\ &= \mathbb{P}(A_0) \times \mathbb{P}_{A_0}(A_1) \times \dots \times \mathbb{P}_{A_0 \cap \dots \cap A_{n-2}}(A_{n-1}) \times \mathbb{P}_{A_0 \cap \dots \cap A_{n-1}}(A_n). \end{aligned}$$


---

**c) Formule des probabilités totales**

Le théorème 7 fournit la formule des probabilités totales :

**Théorème 12.** (formule des probabilités totales)

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_i)_{i \in I} \in \mathcal{A}^I$  un système complet d'événements, au plus dénombrable, tels que  $\forall i \in I$ ,  $\mathbb{P}(A_i) \neq 0$ .

Pour tout  $B \in \mathcal{A}$ ,

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \cap A_i) = \sum_{i \in I} \mathbb{P}_{A_i}(B) \times \mathbb{P}(A_i).$$

**d) Formule de BAYES****Théorème 13.** (formule de BAYES ou formule de probabilité des causes)

Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_j)_{j \in I} \in \mathcal{A}^I$  un système complet d'événements, au plus dénombrable, tels que  $\forall i \in I$ ,  $\mathbb{P}(A_i) \neq 0$ .

Soit  $B \in \mathcal{A}$  tel que  $\mathbb{P}(B) \neq 0$ . Alors, pour tout  $i \in I$ ,

$$\mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i) \times \mathbb{P}_{A_i}(B)}{\sum_{j \in I} \mathbb{P}(A_j) \times \mathbb{P}_{A_j}(B)}$$

**Démonstration.** D'après la formule des probabilités totales, pour  $i \in I$ ,

$$\mathbb{P}_B(A_i) = \frac{\mathbb{P}(B \cap A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i) \times \mathbb{P}_{A_i}(B)}{\sum_{j \in I} \mathbb{P}(A_j) \times \mathbb{P}_{A_j}(B)}$$


---

**2) Événements indépendants****a) Cas de deux événements**

DÉFINITION 7. Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ .

A et B sont **indépendants** si et seulement si  $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$ .

Un résultat immédiat est :

**Théorème 14.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ .

Si  $\mathbb{P}(A) \neq 0$ , A et B sont indépendants si et seulement si  $\mathbb{P}_A(B) = \mathbb{P}(B)$ .

**Théorème 15.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ . Si A et B sont indépendants, alors  $\bar{A}$  et B sont indépendants, A et  $\bar{B}$  sont indépendants et  $\bar{A}$  et  $\bar{B}$  sont indépendants.

**Démonstration.** Supposons A et B indépendants. Donc  $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$ . D'après la formule des probabilités totales,  $\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(\bar{A} \cap B)$  et donc

$$\begin{aligned}\mathbb{P}(\bar{A} \cap B) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A) \times \mathbb{P}(B) = \mathbb{P}(B) (1 - \mathbb{P}(A)) \\ &= \mathbb{P}(\bar{A}) \times \mathbb{P}(B),\end{aligned}$$

et donc  $\bar{A}$  et B sont indépendants. En appliquant à B et A, on a aussi A et  $\bar{B}$  sont indépendants. Enfin, en appliquant à A et  $\bar{B}$ , on a aussi  $\bar{A}$  et  $\bar{B}$  sont indépendants.

---

On note enfin de manière un peu anecdotique que :

**Théorème 16.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A, B) \in \mathcal{A}^2$ .

Si A est négligeable, alors A et B sont indépendants.

Si A est presque sûr, alors A et B sont indépendants.

**Démonstration.** Si A est négligeable, alors  $\mathbb{P}(A) = 0$  puis  $\mathbb{P}(A) \times \mathbb{P}(B) = 0$ . D'autre part,  $A \cap B \subset A$  et donc  $\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0$  puis  $\mathbb{P}(A \cap B) = 0 = \mathbb{P}(A) \times \mathbb{P}(B)$ . Ceci montre que A et B sont indépendants.

Si A est presque sûr, alors  $\bar{A}$  est négligeable puis  $\bar{A}$  et B sont indépendants d'après ce qui précède. On en déduit que A et B sont indépendants d'après le théorème précédent.

---

## b) Familles d'événements mutuellement indépendants

DÉFINITION 8. Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_i)_{i \in I} \in \mathcal{A}^I$  une famille non vide d'événements au plus dénombrable.

La famille  $(A_i)_{i \in I} \in \mathcal{A}^I$  est une famille d'**événements mutuellement indépendants** (ou plus simplement indépendants) si et seulement si pour toute partie finie non vide J de I,

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Un résultat immédiat est :

**Théorème 17.** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable. Soit  $(A_i)_{i \in I} \in \mathcal{A}^I$  une famille non vide d'événements au plus dénombrable.

Si les événements  $A_i$ ,  $i \in I$ , sont mutuellement indépendants, alors les événements  $A_i$  sont deux à deux indépendants.

⚠ Il faut prendre garde au fait que la réciproque est fautive ou encore deux à deux indépendants  $\nrightarrow$  mutuellement indépendants. Considérons par exemple un univers à quatre éléments  $\Omega = \{a, b, c, d\}$ , muni de la probabilité uniforme sur  $(\Omega, \mathcal{P}(\Omega))$ . Soient  $A = \{a, b\}$ ,  $B = \{b, c\}$  et  $C = \{c, a\}$ .

$\mathbb{P}(A \cap B) = \mathbb{P}(\{b\}) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = \mathbb{P}(A) \times \mathbb{P}(B)$ . Puis, par symétrie des rôles,

$\mathbb{P}(A \cap C) = \mathbb{P}(A) \times \mathbb{P}(C)$  et  $\mathbb{P}(B \cap C) = \mathbb{P}(B) \times \mathbb{P}(C)$ . Donc, les événements A, B et C sont deux à deux indépendants.

Maintenant,  $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) \neq \left(\frac{1}{2}\right)^3 = \mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C)$ . Donc, les événements A, B et C ne sont pas mutuellement indépendants.

### III - Variables aléatoires discrètes

#### 1) Loi de probabilité d'une variable aléatoire discrète

DÉFINITION 9. Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable et  $E$  un ensemble non vide.

Une **variable aléatoire discrète** sur  $\Omega$  à valeurs dans  $E$  est une application  $X$  de  $\Omega$  dans  $E$  telle que  $X(\Omega)$  soit une partie au plus dénombrable de  $E$  et telle que, pour tout  $x$  de  $E$ ,  $X^{-1}(\{x\}) \in \mathcal{A}$ .

Si  $E$  est une partie de  $\mathbb{R}$ , la variable aléatoire  $X$  est dite **réelle**.

#### Commentaire.

◇ Dans la pratique de maths sup et maths spé,  $E$  est le plus souvent  $\mathbb{N}$  ou  $\mathbb{Z}$ . C'est le cas des lois usuelles (uniforme, de BERNOULLI, géométrique, de POISSON) qui sont au programme.

◇ On rappelle la signification de  $X^{-1}(\{x\})$ . Si  $x$  est un élément de l'ensemble d'arrivée  $E$ ,  $X^{-1}(\{x\})$  est l'image réciproque du singleton  $\{x\}$  par l'application  $X$ , c'est-à-dire l'ensemble des antécédents de l'élément  $x$  de  $E$  par l'application  $X$  ou encore

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

$X^{-1}(\{x\})$  est donc un élément de  $\mathcal{P}(\Omega)$  et plus précisément un élément de  $\mathcal{A}$  d'après la définition. Dit autrement,  $X^{-1}(\{x\})$  est un événement que l'on écrit en pratique  $X = x$  pour simplifier la notation.

Plus généralement, si  $A$  est une partie de  $E$ , l'événement  $X^{-1}(A) = \{\omega \in \Omega / X(\omega) \in A\}$  se note plus simplement  $X \in A$ .

◇ Si  $X$  est réelle et  $x$  est un élément de  $E$ , on note plus simplement les événements  $X^{-1}(]x, +\infty[)$ ,  $X^{-1}([x, +\infty[)$ ,  $X^{-1}(]-\infty, x])$ ,  $X^{-1}(]-\infty, x])$  : ces événements sont respectivement notés,  $X > x$ ,  $X \geq x$ ,  $X < x$  et  $X \leq x$ . □

**Théorème 18.** Soient  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace probabilisé au plus dénombrable et  $E$  un ensemble non vide au plus dénombrable. Soit  $X$  une variable aléatoire sur cet espace à valeurs dans  $E$ .

L'application  $\mathbb{P}_X : \mathcal{P}(E) \rightarrow [0, 1]$  est une probabilité sur l'espace probabilisable  $(E, \mathcal{P}(E))$ .

$$A \mapsto \mathbb{P}(X \in A)$$

#### Démonstration.

• Vérifions tout d'abord que  $\mathbb{P}_X$  est bien une application de  $\mathcal{P}(E)$  dans  $[0, 1]$ . Il s'agit de vérifier que pour  $A \in \mathcal{P}(E)$ ,  $X \in A$  est effectivement un événement, c'est-à-dire un élément de la tribu  $\mathcal{A}$ .

Soit donc  $A \in \mathcal{P}(E)$ .  $A \cap X(\Omega)$  est une partie de  $X(\Omega)$  et est donc au plus dénombrable, c'est-à-dire vide, finie non vide ou dénombrable.

Si  $A \cap X(\Omega)$  est vide, alors  $X \in A = \emptyset \in \mathcal{A}$ .

Si  $A \cap X(\Omega)$  est finie non vide ou dénombrable, on peut poser  $A \cap X(\Omega) = \{x_i, i \in I\}$  où  $I$  est soit un ensemble de la forme  $[[1, p]]$ ,  $p \in \mathbb{N}^*$ , soit  $\mathbb{N}$ . Mais alors

$$X^{-1}(A) = \bigcup_{i \in I} X^{-1}(\{x_i\}) \in \mathcal{A}.$$

Dans tous les cas,  $X \in A$  est un événement.

•  $\mathbb{P}_X(E) = \mathbb{P}(X \in E) = \mathbb{P}(\Omega) = 1$ .

• Soit  $(A_n)_{n \in \mathbb{N}} \in (\mathcal{P}(E))^{\mathbb{N}}$ . Alors,  $(X \in A_n)_{n \in \mathbb{N}}$  est une suite d'événements (éléments de  $\mathcal{A}$ ) deux à deux disjoints (car  $X$  est une application). Donc, par  $\sigma$ -additivité de  $\mathbb{P}$ ,

$$\mathbb{P}_X \left( \bigcup_{n=0}^{+\infty} A_n \right) = \mathbb{P} \left( X \in \bigcup_{n=0}^{+\infty} A_n \right) = \mathbb{P} \left( \bigcup_{n=0}^{+\infty} X^{-1}(A_n) \right) = \sum_{n=0}^{+\infty} \mathbb{P}(X^{-1}(A_n)) = \sum_{n=0}^{+\infty} \mathbb{P}_X(A_n).$$

On a montré que  $\mathbb{P}_X$  est une probabilité sur l'espace probabilisable  $(E, \mathcal{P}(E))$ .

DÉFINITION 10. L'application  $\mathbb{P}_X$  du théorème 18 s'appelle la **loi de probabilité** de la variable aléatoire  $X$ .

**Commentaire.** La loi  $\mathbb{P}_X$  est entièrement déterminée par les *atomes de probabilité*  $\mathbb{P}_X(x) = \mathbb{P}(X \in \{x\}) = \mathbb{P}(X = x)$ ,  $x \in E$ . Dans la pratique, pour répondre à une question du type « déterminer la loi de probabilité de la variable  $X$  », ce sont les  $\mathbb{P}(X = x)$ ,  $x \in E$ , que l'on fournit. □

**Notation.** Quand deux variables aléatoires  $X$  et  $Y$  ont la même loi de probabilité, on écrit  $X \sim Y$ . Quand une variable aléatoire suit un loi de probabilité  $\mathcal{L}$  donnée, on écrit  $X \sim \mathcal{L}$ . □

- 2) *Fonctions de variables aléatoires*
- 3) *Couples de variables aléatoires.  $n$ -uplets de variables aléatoires*
- 4) *Variables aléatoires indépendantes*

## IV - Espérance, variance

- 1) *Espérance d'une variable aléatoire*
- 2) *Variance d'une variable aléatoire*
- 3) *Covariance d'un couple de variables aléatoires*
- 4) *Loi faible des grands nombres*

## V - Variables usuelles

- 1) *Loi géométrique de paramètre  $p$*
- 2) *Loi de POISSON de paramètre  $\lambda$* 
  - a) *Définition*
  - b) *Approximation de la loi binomiale par la loi de POISSON*

## VI - Fonctions génératrices

- 1) *Fonction génératrice d'une variable aléatoire*
- 2) *Fonctions génératrices des variables usuelles*
  - a) *Fonction génératrice d'un variable aléatoire de BERNOULLI*
  - b) *Fonction génératrice d'un variable aléatoire binomiale*
  - c) *Fonction génératrice d'un variable aléatoire géométrique*
  - d) *Fonction génératrice d'un variable aléatoire de POISSON*

DÉFINITION 1.

Théorème 1.

Démonstration.

---

Exercice 1.

Solution 1.

---

a) *Définition*