Chapitre XIII

Probabilités discrètes

Table des matières

Partie A : Dénombrabilité **3**
1. Généralités . . . 3
2. Opérations sur les ensembles dénombrables . . . 6
3. Exemples d'ensembles non dénombrables . . . 7

Partie B : Familles sommables **9**
1. Familles sommables de nombres réels positifs . . . 9
2. Familles sommables de nombres complexes . . . 12
3. Applications des familles sommables . . . 14
4. Exercices . . . 18

Partie C : Espaces probabilisés **22**
1. Espaces probabilisables . . . 22
2. Espaces probabilisés . . . 24
3. Propriétés élémentaires des probabilités . . . 27
4. Probabilité sur un univers au plus dénombrable . . . 30

Partie D : Probabilités conditionnelles **32**
1. Conditionnement . . . 32
2. Formules conditionnelles . . . 33
3. Événements indépendants . . . 37

Partie E : Variables aléatoires discrètes **39**
1. Variables aléatoires discrètes . . . 39
2. Lois usuelles sur un univers fini . . . 41
3. Lois usuelles sur un univers dénombrable . . . 42
4. Couples de variables aléatoires . . . 44
5. Variables aléatoires indépendantes . . . 46

Partie F : Espérance, variance **48**
1. Espérance . . . 48
2. Variance . . . 55
3. Covariance . . . 58
4. Loi faible des grands nombres . . . 59
5. Fonctions génératrices . . . 59

Dans ce chapitre, on utilisera la notation $\overline{A}$ pour désigner le complémentaire d'une sous-partie A d'un ensemble.

Partie A

Dénombrabilité

Dans cette partie, E et F désignent des ensembles quelconques.

1. Généralités

Définition 1. *Dénombrabilité*

— On dit que E est **au plus dénombrable** s'il est en bijection avec une partie de $\mathbb{N}$.
— On dit que E est **dénombrable** s'il est en bijection avec $\mathbb{N}$.

Exemple 1.

— Un ensemble fini est au plus dénombrable.

> En effet, si son cardinal est n, alors il est en bijection avec la partie finie de $\mathbb{N}$, $[\![0, n]\!]$.

— L'ensemble $\mathbb{N}$ et $\mathbb{N}^*$ sont dénombrables.

> Pour la dénombrabilité de $\mathbb{N}$ il suffit de considérer l'identité. Quant à $\mathbb{N}^*$, l'application $f : n \mapsto n+1$ est une bijection de $\mathbb{N}$ dans $\mathbb{N}^*$ (sa réciproque est $g : n \mapsto n-1$).

— L'ensemble $\mathbb{Z}$ est dénombrable.

> On définit l'application $f : \mathbb{N} \to \mathbb{Z}$ telle que, pour $n \in \mathbb{N}$
>
> $$f(n) = \begin{cases} \frac{n}{2} & \text{si } n \text{ pair} \\ -\frac{n+1}{2} & \text{si } n \text{ impair} \end{cases}$$
>
> i.e. $f(0) = 0$, $f(1) = -1$, $f(2) = 1$, $f(3) = -2$, $f(4) = 2$, etc...
> f est bijective de réciproque $f^{-1} : \mathbb{Z} \to \mathbb{N}$:
>
> $$f^{-1}(n) = \begin{cases} 2k & \text{si } k \geq 0 \\ -(2k+1) & \text{si } k < 0 \end{cases}$$

Exercice 1.

Montrer que $2\mathbb{N}$ est un ensemble dénombrable.

Correction.

On définit $f : \mathbb{N} \to 2\mathbb{N}$ par :

$$\forall\, n \in \mathbb{N}, \quad f(n) = 2n$$

Alors f est bijective : sa réciproque est $g : 2\mathbb{N} \to \mathbb{N}$ où, pour tout $k \in 2\mathbb{N}$, $g(k) = \frac{k}{2}$.

En fait, on a le résultat général suivant :

Proposition 1.

Soit A une partie infinie de $\mathbb{N}$. Alors A est dénombrable.

Démonstration.

Soit A une partie infinie de $\mathbb{N}$. Alors A est non vide - car l'ensemble vide est un ensemble fini (de cardinal 0), et A est non majoré dans $\mathbb{N}$ car, si on suppose par l'absurde qu'il est majoré par $n \in \mathbb{N}$, alors $A \subset [\![0, n]\!]$ qui est fini, contradiction.

On construit par récurrence une suite strictement croissante (u_n) à valeurs dans A qui parcourt, dans l'ordre, tous les éléments de A :

- **Initialisation.** Comme A est une partie non vide de $\mathbb{N}$, A possède un plus petit élément $\min(A)$. On note alors $u_0 = \min(A)$.
- **Hérédité.** Soit $n \in \mathbb{N}$. On suppose $u_0, ..., u_n$ construits. Comme A n'est pas majoré par u_n, l'ensemble $\{k \in A \mid k > u_n\}$ est une partie non vide de $\mathbb{N}$ et donc possède un plus petit élément que l'on note u_{n+1} i.e.

$$u_{n+1} = \min(\{k \in A \mid k > u_n\}).$$

Ainsi la fonction de $\mathbb{N}$ dans A associée à la suite (u_n), $\varphi : n \mapsto u_n$, est une bijection.

En effet, elle est :

- *injective.* Pour tout $n \in \mathbb{N}$, par construction, on a $u_{n+1} > u_n$. Par suite la suite (u_n) est strictement croissante et donc la fonction φ l'est aussi. Ainsi, φ est injective.
- *surjective.* Soit $k \in A$. Comme la suite (u_n) est strictement croissante, il existe $m \in \mathbb{N}$ tel que $u_{m+1} > k$. L'ensemble de ces m est une partie non vide de $\mathbb{N}$, donc il en existe un plus petit que l'on note n i.e. $u_{n+1} > k$ et $k \geq u_n$.
  Or, par définition, u_{n+1} est le plus petit élément de A strictement plus grand que u_n donc l'inégalité $u_{n+1} > k > u_n$ est impossible. Ainsi, $\varphi(n) = u_n = k$. Par suite, φ est surjective.

□

Proposition 2.

L'ensemble E est au plus dénombrable si, et seulement si, il existe une injection de E dans $\mathbb{N}$.

Démonstration.

Si E est au plus dénombrable, alors il existe une bijection f de E dans $A \subset E$. Par suite, $f : E \to \mathbb{N}$ est injective.

Réciproquement, s'il existe une injection f de E dans $\mathbb{N}$, alors $f : E \to f(E)$ est une bijection de

E dans la partie $f(E)$ de $\mathbb{N}$. Donc E est au plus dénombrable. $\square$

Exemple 2.

L'ensemble $\mathbb{N}^2$ est dénombrable : en effet, l'application $(n, m) \to 2^n 3^m$ est injective par unicité de la décomposition d'un nombre entier en facteurs premiers et de plus, $\mathbb{N}^2$ est infini.

Proposition 3.

Soit $A \subset E$. Si E est dénombrable, alors A est au plus dénombrable. Et de plus, si E est dénombrable et A est infini, alors A est dénombrable.

Démonstration.

On suppose E dénombrable. Alors il existe une bijection $f : E \to \mathbb{N}$. De plus, l'application $\mathrm{id}_{|A} : A \to E$ telle que, pour $a \in A$, $\mathrm{id}_{|A}(a) = a$ est injective. Par suite, l'application $g = \mathrm{id}_{|A} \circ f : A \to \mathbb{N}$ est injective. Il en résulte que A est au plus dénombrable.

De plus, si A est infini, $g : A \to g(A)$ est une bijection et $g(A)$ est donc une partie infinie de $\mathbb{N}$. Par suite, $g(A)$ est dénombrable, donc il existe une bijection $h : g(A) \to \mathbb{N}$. Ainsi, $g \circ h$ est une bijection de A dans $\mathbb{N}$. Donc A est dénombrable. $\square$

Proposition 4.

On suppose E non vide. Alors E est au plus dénombrable si, et seulement si, il existe une surjection de $\mathbb{N}$ sur E.

Démonstration.

- $(\Rightarrow)$. On suppose E au plus dénombrable. Alors il existe $A \subset \mathbb{N}$ et $f : A \to E$ une bijection. Comme E est non vide, on choisit $x_0 \in E$ et on construit l'application $\tilde{f}$: pour tout $n \in \mathbb{N}$,

$$\tilde{f}(n) = \begin{cases} f(n) & \text{si } n \in A \\ x_0 & \text{si } n \notin A. \end{cases}$$

  Alors on a $\tilde{f}(\mathbb{N}) = f(A) \cup \{x_0\} = f(A) = E$. Donc $\tilde{f}$ est une surjection de $\mathbb{N}$ sur E.
- $(\Leftarrow)$. On suppose qu'il existe $f : \mathbb{N} \to E$ une surjection. Montrons qu'il existe une application injective g de E dans $\mathbb{N}$. Comme f est surjective, pour tout $x \in E$, $f^{-1}(\{x\})$ est une partie non vide de $\mathbb{N}$. Ainsi, la fonction de E dans $\mathbb{N}$:

$$g : x \mapsto \min(f^{-1}(\{x\})$$

  est injective. En effet, pour $x, y \in E$ tels que $g(x) = g(y)$, on a $x = f(g(x)) = f(g(y)) = y$. $\square$

Exercice 2.

Montrer de nouveau que $\mathbb{N}^2$ est dénombrable en exhibant (graphiquement) une surjection de $\mathbb{N}^2$ sur $\mathbb{N}$.

2. Opérations sur les ensembles dénombrables

Proposition 5. *Produit cartésien*

Si E, F sont dénombrables, alors $E \times F$ est dénombrable.

Démonstration.

On suppose E, F dénombrables. Alors il existe des bijections $f : E \to \mathbb{N}$ et $g : F \to \mathbb{N}$. Alors la fonction $(x, y) \mapsto (f(x), g(y))$ est une bijection de $E \times F$ dans $\mathbb{N}^2$ qui est lui-même dénombrable. Ainsi, on peut construire une bijection de $\mathbb{N}$ sur $E \times F$. Par suite, $E \times F$ est dénombrable. $\square$

Exemple 3.

L'ensemble $\mathbb{Q}$ des nombres rationnels est dénombrable.

> En effet, l'application $f : \mathbb{Z} \times \mathbb{N}^* \to \mathbb{Q}$ telle que, pour $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$, $f(p, q) = \frac{p}{q}$ est surjective et l'ensemble $\mathbb{Z} \times \mathbb{N}^*$ est dénombrable comme produit cartésien d'ensembles dénombrables. Par suite, il existe une surjection de $\mathbb{N}$ sur $\mathbb{Q}$, d'où $\mathbb{Q}$ est dénombrable.

Remarque 1.

Par récurrence, on obtient qu'un produit cartésien d'une famille finie (non vide) d'ensembles dénombrables est dénombrable.

Exercice 3.

Soit A, B des parties de E. Montrer que si A, B sont dénombrables, alors $A \cup B$ est dénombrable.

On a la généralisation suivante :

Proposition 6. *Réunion*

Soit $(A_i)_{i \in I}$ une famille de parties de E. Si I est au plus dénombrable et si, pour tout $i \in I$, A_i est au plus dénombrable, alors $\bigcup_{i \in I} A_i$ est au plus dénombrable. Et de plus, s'il existe $i \in I$ tel que A_i est dénombrable, alors $\bigcup_{i \in I} A_i$ est dénombrable.

Démonstration.

On suppose I non vide, au plus dénombrable et pour tout $i \in I$, A_i au plus dénombrable. Alors, pour chaque $i \in I$, il existe une surjection f_i de $\mathbb{N}$ dans A_i (quitte à le supprimer de la liste, on peut supposer A_i non vide). Alors la fonction $f : I \times \mathbb{N} \to \bigcup_{i \in I} A_i$ telle quel, pour $(i, n) \in I \times \mathbb{N}$:

$$f(i, n) = f_i(n).$$

est surjective. En effet, pour tout $a \in \bigcup_{i \in I} A_i$, il existe $i \in \mathbb{I}$ tel que $a \in A_i$ et il existe $n \in \mathbb{N}$ tel que $a = f_i(n)$; par suite, $a) f(i, n)$.
Ainsi, comme $I \times \mathbb{N}$ est dénombrable, il en résulte que $\bigcup_{i \in I} A_i$ est dénombrable. □

3. Exemples d'ensembles non dénombrables

Proposition 7.

Il n'existe aucune surjection de E sur l'ensemble $\mathcal{P}(E)$ des parties de E.

Démonstration.

Soit $f : E \to \mathcal{P}(E)$ une application. On note $A = \{x \in E \mid x \notin f(x)\}$. Alors $A \subset \mathcal{P}(E)$ n'a pas d'antécédent par f. En effet, pour $x \in E$ on a :

- *1er cas :* $x \in A$. Alors $f(x) \neq A$ car comme $x \in A$, par définition de A, $x \notin f(x)$.
- *2nd cas :* $x \notin A$. Alors $f(x) \neq A$ car comme $x \notin A$, par définition de A, $x \in f(x)$.

Dans tous les cas, $f(x) \neq A$.

Il en résulte que f n'est pas surjective. □

Remarque 2.

Il en résulte que $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable : en effet, il n'existe a fortiori aucune bijection de $\mathbb{N}$ dans $\mathcal{P}(\mathbb{N})$.

Exercice 4.

1. Montrer que $\mathcal{P}(E)$ est en bijection avec $\{0, 1\}^E$.
2. On considère $X = \{0, 1\}^{\mathbb{N}}$. Montrer que X n'est pas dénombrable grâce à l'argument diagonal de Cantor : supposer par l'absurde qu'il existe une énumération (indexée par $\mathbb{N}$) de X, et construire un élément de X qui ne peut pas être dans cette énumération (indice : argument "diagonal" !).
3. En déduire qu'un produit dénombrable d'ensembles au plus dénombrables n'est pas dénombrable en général et également que $\mathcal{P}(\mathbb{N})$ n'est toujours pas dénombrable !

Proposition 8.

L'ensemble $\mathbb{R}$ des nombres réels n'est pas dénombrable.

Partie B

Familles sommables

1. Familles sommables de nombres réels positifs

Dans ce paragraphe, I désigne un ensemble dénombrable.

a. Définitions

Définition 2. *Famille sommable*

Soit $(a_i)_{i\in I}$ de réels positifs est **sommable** si l'ensemble

$$\left\{\sum_{j\in J} a_j \mid J \subset I \text{ et } J \text{ fini}\right\}$$

est majoré ; autrement dit si :

$$\text{il existe } M \geq 0 \text{ tel que, pour toute partie } J \text{ fini de } I,\ \textstyle\sum_{j\in J} a_j \leq M.$$

— Si la famille $(a_i)_{i\in I}$ est sommable, alors on appelle **somme de la famille** $(a_i)_{i\in I}$ et on note $\displaystyle\sum_{i\in I} a_i$ la quantité :

$$\sum_{i\in I} a_i = \sup\left(\left\{\sum_{j\in J} a_j \mid J \subset I \text{ et } J \text{ fini}\right\}\right).$$

— Si la famille $(a_i)_{i\in I}$ n'est pas sommable, on convient que $\sum_{i\in I} a_i = +\infty$.

Exemple 4.

La famille $(\frac{1}{2^n})_{n\in\mathbb{N}}$ est sommable.

En effet, pour toute partie finie J de $\mathbb{N}$, on a $J \subset [\![0, m]\!]$ où $m = \max(J)$, donc :

$$\sum_{j\in J} \frac{1}{2^j} \leq \sum_{i=0}^{m} \frac{1}{2^i} = 2 - \frac{1}{2^m} \leq 2.$$

(Démontrer que la somme de la famille est en fait bien égale à 2 grâce à la définition précédente)

Proposition 9.

Soit $(u_n)_{n\in\mathbb{N}}$ une suite de réels positifs. Alors la famille $(u_n)_{n\in\mathbb{N}}$ est sommable si, et seulement si, $\sum u_n$ converge. Dans ce cas, on a :

$$\sum_{n\in\mathbb{N}} u_n = \sum_{n=0}^{+\infty} u_n.$$

b. Propriétés

Proposition 10. *Sous-familles sommable*

Soit $(a_i)_{i\in I}$ une famille de réels positifs. Si $(a_i)_{i\in I}$ est sommable, alors pour tout $J \subset I$ la sous-famille $(a_j)_{j\in J}$ est sommable et

$$\sum_{j\in J} a_j \leq \sum_{i\in I} a_i.$$

Proposition 11.

Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles de réels positifs telles que pour tout $i \in I$, $a_i \leq b_i$. Si $(b_i)_{i\in I}$ est sommable, alors $(a_i)_{i\in I}$ est sommable et

$$\sum_{i\in I} a_i \leq \sum_{i\in I} b_i.$$

Proposition 12.

Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles de réels positifs et $\lambda, \mu \in \mathbb{R}_+$. Si $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ sont sommables, alors la famille $(\lambda a_i + \mu b_i)_{i\in I}$ est sommable et

$$\sum_{i\in I}(\lambda a_i + \mu b_i) = \lambda \sum_{i\in I} a_i + \mu \sum_{i\in I} b_i.$$

Proposition 13. *Changement d'indice*

Soit σ une bijection d'un ensemble J sur I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable si et seulement si, $(a_{\sigma(j)})_{j\in J}$ est sommable. Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}.$$

c. Sommation par paquets

Définition 3. *Partition*

Soit $(I_\omega)_{\omega\in\Omega}$ une famille de partie de I. On dit que $(I_\omega)_{\omega\in\Omega}$ est une **partition** de I si $\bigcup_{\omega\in\Omega} I_\omega = I$ et pour tout $\omega, \omega' \in \Omega$ avec $\omega \neq \omega'$, $I_\omega \cap I_{\omega'} = \emptyset$.

Exemple 5.

Les ensembles $2\mathbb{N}$ et $2\mathbb{N}+1$ forment une partition de $\mathbb{N}$.

Proposition 14. *Sommation par deux paquets*

Soit J, K des parties de I tels que (J, K) est une partition de I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable si, et seulement si $(a_j)_{j\in J}$ et $(a_k)_{i\in K}$ sont sommables. Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{j\in J} a_j + \sum_{k\in K} a_k.$$

Exercice 5.

On admet l'égalité $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. En déduire que $\sum_{n\geq 1} \frac{1}{(2n+1)^2}$ converge et déterminer sa somme.

On peut généraliser la proposition précédente de la façon suivante :

Théorème 1. *Sommation par paquets*

Soit $(I_\omega)_{\omega\in\Omega}$ une partition au plus dénombrable de I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

- pour tout $\omega \in \Omega$, la famille $(a_i)_{i\in I_\omega}$ est sommable de somme $\alpha_\omega = \sum_{i\in I_\omega} a_i$,
- la famille $(\alpha_\omega)_{\omega\in\Omega}$ est sommable.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{\omega\in\Omega} \left(\sum_{i\in I_\omega} a_i \right).$$

Corollaire 1.

Soit $(I_n)_{n\in\mathbb{N}}$ une partition au plus dénombrable de I et $(a_i)_{i\in I}$ une famille de réels positifs. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

- pour tout $n \in \mathbb{N}$, la famille $(a_i)_{i\in I_n}$ est sommable,
- la série $\sum \left(\sum_{i\in I_n} a_i \right)$ est convergente.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} a_i\right).$$

2. Familles sommables de nombres complexes

Dans ce paragraphe, I désigne un ensemble dénombrable.

a. Définitions

Définition 4. *Famille sommable de nombres complexes*

Soit $(a_i)_{i\in I}$ une famille de nombres complexes. On dit que la famille $(a_i)_{i\in I}$ est **sommable** si la famille $(|a_i|)_{i\in I}$ de réels positifs est sommable.

Remarque 3.

Ainsi, d'après ce qu'on a vu précédemment, une suite $(a_n)_{n\in\mathbb{N}}$ est sommable si, set seulement si, la série de terme général $(a_n)_{n\in\mathbb{N}}$ est absolument convergente.

Définition 5. *Somme d'une famille sommable de réels*

Soit $(a_i)_{i\in I}$ une famille sommable de nombres réels. On pose, pour chaque $i \in I$, $a_i^+ = \max(0, a_i)$ et $a_i^- = \min(0, a_i)$. On appelle **somme** de la famille $(a_i)_{i\in I}$ et on note $\sum_{i\in I} a_i$ la quantité :

$$\sum_{i\in I} a_i = \sum_{i\in I} a_i^+ + \sum_{i\in I} a_i^-.$$

Définition 6. *Somme d'une famille sommable de complexes*

Soit $(a_i)_{i\in I}$ une famille sommable de nombres complexes. On appelle **somme** de la famille $(a_i)_{i\in I}$ et on note $\sum_{i\in I} a_i$ la quantité :

$$\sum_{i\in I} a_i = \sum_{i\in I} \mathrm{Re}(a_i) + i\sum_{i\in I} \mathrm{Im}(a_i).$$

Proposition 15.

Soit $(a_n)_{n\in\mathbb{N}}$ une suite de nombres complexes. La famille $(a_n)_{n\in\mathbb{N}}$ est sommable si, et seulement

si, la série $\sum a_n$ est absolument convergente. Dans ce cas, on a :

$$\sum_{n\in\mathbb{N}} a_n = \sum_{i=0}^{+\infty} a_n.$$

b. Propriétés

Proposition 16. *Changement d'indice*

Soit σ une bijection d'un ensemble J sur I et $(a_i)_{i\in I}$ une famille de nombres complexes. Alors $(a_i)_{i\in I}$ est sommable si et seulement si, $(a_{\sigma(j)})_{j\in J}$ est sommable. Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{j\in J} a_{\sigma(j)}.$$

Remarque 4.

En combinant les deux précédentes propositions, on obtient que la somme d'une série absolument convergente est invariante par changement bijectif d'indice. Dans le cas d'une série semi-convergente, par exemple $\sum_{n\geq 1} \frac{(-1)^n}{n}$, on peut, en appliquant une permutation d'indice bien choisie, faire converger la série obtenue vers n'importe quel nombre réel, et même la faire diverger !

Proposition 17. *Inégalité triangulaire*

$(a_i)_{i\in I}$ une famille de nombres complexes. Si $(a_i)_{i\in I}$ est sommable, alors

$$|\sum_{i\in I} a_i| \leq \sum_{i\in I} |a_i|.$$

Proposition 18.

Soit $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ des familles à valeurs dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $\lambda, \mu \in \mathbb{K}$. Si $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ sont sommables, alors la famille $(\lambda a_i + \mu b_i)_{i\in I}$ est sommable et

$$\sum_{i\in I}(\lambda a_i + \mu b_i) = \lambda \sum_{i\in I} a_i + \mu \sum_{i\in I} b_i.$$

Question.

De quelle structure peut-on munir l'ensemble des familles sommables sur le corps $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$?

c. Sommation par paquets

Théorème 2. *Sommation par paquets*

Soit $(I_\omega)_{\omega\in\Omega}$ une partition au plus dénombrable de I et $(a_i)_{i\in I}$ une famille de complexes. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

— pour tout $\omega \in \Omega$, la famille $(a_i)_{i\in I_\omega}$ est sommable de somme $\alpha_\omega = \sum_{i\in I_\omega} a_i$,

— la famille $(\alpha_\omega)_{\omega\in\Omega}$ est sommable.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{\omega\in\Omega}\left(\sum_{i\in I_\omega} a_i\right).$$

Corollaire 2.

Soit $(I_n)_{n\in\mathbb{N}}$ une partition de I et $(a_i)_{i\in I}$ une famille de complexes. Alors $(a_i)_{i\in I}$ est sommable, si, et seulement si,

— pour tout $n \in \mathbb{N}$, la famille $(a_i)_{i\in I_n}$ est sommable,

— la série $\sum\left(\sum_{i\in I_n} |a_i|\right)$ est convergente.

Dans ce cas, on a :

$$\sum_{i\in I} a_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in I_n} a_i\right).$$

3. Applications des familles sommables

a. Sommes doubles

Définition-Proposition 7. *Partitions classiques de $\mathbb{N}^2$*

On appelle :

— **Partition verticale** de $\mathbb{N}^2$, la partition $(V_m)_{m\in\mathbb{N}}$ de $\mathbb{N}^2$ où, pour $m \in \mathbb{N}$,

$$V_m = \{m\} \times \mathbb{N} = \{(m,n) \mid n \in \mathbb{N}\}$$

.

— **Partition horizontale** de $\mathbb{N}^2$, la partition $(H_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$ où, pour $n \in \mathbb{N}$,

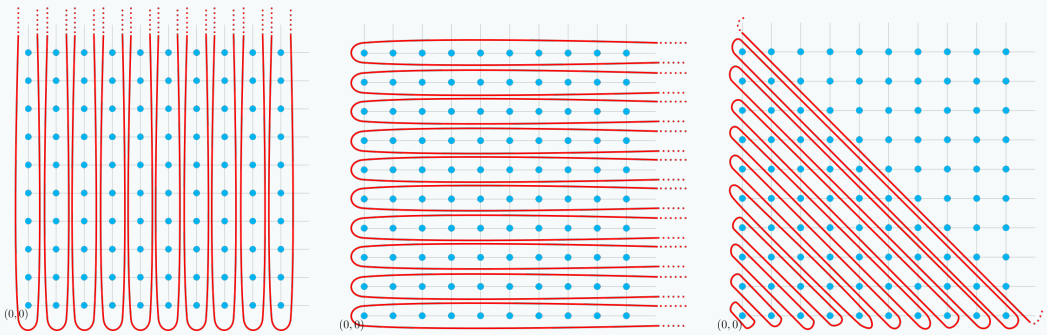
$$H_n = \mathbb{N} \times \{n\} = \{(m,n) \mid m \in \mathbb{N}\}$$

.

— **Partition diagonale** de $\mathbb{N}^2$, la partition $(\Delta_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$ où, pour $k \in \mathbb{N}$,

$$\Delta_k = \{(m,n) \in \mathbb{N}^2 \mid m+n = k\} = \{(p, k-p) \mid p \in [\![0,k]\!]\}$$

.



Partition verticale | Partition horizontale | Partition diagonale

Théorème 3. *Somme double "verticale"*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels positifs. Alors $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si :

— pour tout $m \in \mathbb{N}$, la série $\sum_{n\geq 0} a_{m,n}$ converge (somme sur V_m) et

— la série $\sum \left(\sum_{n=0}^{+\infty} a_{m,n} \right)$ converge.

Dans ce cas, on a :

$$\sum_{(n,m)\in\mathbb{N}^2} a_{m,n} = \sum_{m=0}^{+\infty} \left(\sum_{n=0}^{+\infty} a_{m,n} \right).$$

Théorème 4. *Somme double "horizontale"*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels positifs. Alors $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si :

— pour tout $n \in \mathbb{N}$, la série $\sum_{m\geq 0} a_{m,n}$ converge (somme sur H_n) et

— la série $\sum \left(\sum_{m=0}^{+\infty} a_{m,n} \right)$ converge.

Dans ce cas, on a :

$$\sum_{(n,m)\in\mathbb{N}^2} a_{m,n} = \sum_{n=0}^{+\infty} \left(\sum_{m=0}^{+\infty} a_{m,n} \right).$$

Théorème 5. *Somme double "diagonale"*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de réels positifs. Alors $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si :

— la série $\sum \left(\underbrace{\sum_{p=0}^{k} a_{p,k-p}}_{\text{somme sur } \Delta_k} \right)$ converge.

Dans ce cas, on a :

$$\sum_{(n,m)\in\mathbb{N}^2} a_{m,n} = \sum_{k=0}^{+\infty} \left(\sum_{p=0}^{k} a_{p,n-p} \right) = \sum_{k=0}^{+\infty} \left(\sum_{m+n=k} a_{m,n} \right).$$

Corollaire 3. *Somme double de complexes*

Soit $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ une famille de complexes. Si $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable alors :

$$\begin{aligned}
\sum_{(m,n)\in\mathbb{N}^2} a_{m,n} &= \sum_{n=0}^{+\infty} \left(\sum_{m=0}^{+\infty} a_{m,n} \right) \\
&= \sum_{m=0}^{+\infty} \left(\sum_{n=0}^{+\infty} a_{m,n} \right) \\
&= \sum_{k=0}^{+\infty} \left(\sum_{p=0}^{k} a_{p,n-p} \right) \\
&= \sum_{k=0}^{+\infty} \left(\sum_{m+n=k} a_{m,n} \right)
\end{aligned}$$

Exercice 6.

Soit $\alpha \in \mathbb{R}$. Étudier la sommabilité de la famille $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ où, pour $(m,n) \in \mathbb{N}^2$,

$$a_{m,n} = \frac{1}{(m+n+1)^\alpha}.$$

Correction.

Soit $\alpha \in \mathbb{R}$. On utilise le théorème de sommation par paquets sur la partition diagonale de $\mathbb{N}^2$ *(ce choix est guidé par la "vue" du $m+n$ qui sera constant sur les diagonale !)* i.e.

$$\mathbb{N}^2 = \bigsqcup_{k=0}^{+\infty} \Delta_k \text{ où } \Delta_k = \{(n, k-n) \mid n \in [\![0,k]\!]\}.$$

Testons les hypothèses de ce théorème :

La famille $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est une famille de réels positifs.

— Somme sur chaque paquet : soit $k \in \mathbb{N}$. La famille $(a_{m,n})_{(m,n)\in\Delta_k}$ est une famille finie, donc elle est sommable et on pose :

$$s_k = \sum_{(m,n)\in\Delta_k} \frac{1}{(m+n+1)^\alpha} = \sum_{n=0}^{k} \frac{1}{(k+1)^\alpha} = \frac{1}{(k+1)^{\alpha-1}}.$$

— Somme des sommes des paquets : la famille $(s_k)_{k\in\mathbb{N}}$ est une suite, elle est donc sommable si, et seulement si, $\sum s_k$ converge absolument.
On a $|s_k| = \dfrac{1}{(k+1)^{\alpha-1}} \underset{k\to+\infty}{\sim} \dfrac{1}{k^{\alpha-1}}$; or, d'après le critère de Riemann, $\sum_{k\geq 1} \frac{1}{k^{\alpha-1}}$ converge si, et seulement si, $\alpha - 1 > 1$ i.e. $\alpha > 2$. D'où, par comparaison, $\sum s_k$ converge absolument si, et seulement si, $\alpha > 2$.
Ainsi, la famille $(s_k)_{k\in\mathbb{N}}$ est sommable si, et seulement si, $\alpha > 2$.

Par suite, d'après le théorème de sommation par paquets, la famille $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si, et seulement si, $\alpha > 2$. Et dans ce cas, on a :

$$\sum_{(m,n)\in\mathbb{N}^2} a_{m,n} = \sum_{k=0}^{+\infty} s_k = \sum_{k=0}^{+\infty} \frac{1}{(k+1)^{\alpha-1}} = \sum_{k=1}^{+\infty} \frac{1}{k^{\alpha-1}} = \zeta(\alpha-1).$$

b. Produit de Cauchy

Proposition 19.

Soit $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ des suites de nombres complexes. Si $\sum a_n$ et $\sum b_n$ sont absolument convergentes, alors

$$\sum_{(m,n)\in\mathbb{N}^2} a_m b_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right).$$

Définition 8. *Produit de Cauchy*

Soit $\sum a_n$ et $\sum b_n$ des séries de nombres complexes. On appelle **produit de Cauchy** des séries $\sum a_n$ et $\sum b_n$ la série $\sum c_n$ de terme général défini, pour $n \in \mathbb{N}$, par :

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Théorème 6.

Soit $\sum a_n$ et $\sum b_n$ des séries de nombres complexes. Si $\sum a_n$ et $\sum b_n$ sont absolument convergentes, alors leur produit de Cauchy $\sum c_n$ est absolument convergent et on a :

$$\sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right) = \sum_{n=0}^{+\infty}\left(\sum_{p+q=n} a_p b_q\right).$$

Exercice 7.

Soit $z \in \mathbb{C}$ tel que $|z| < 1$. Montrer que $\displaystyle\sum_{n=0}^{+\infty}(n+1)z^n = \left(\sum_{n=0}^{+\infty} z^n\right)^2$. En déduire que

$$\sum_{n=0}^{+\infty}(n+1)z^n = \frac{1}{(z-1)^2}.$$

4. Exercices

Exercice 8.

Déterminer si la famille $(a_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable où pour $m, n \in \mathbb{N}$:

$$(1)\ a_{m,n} = \frac{1}{2^{m+n}}; \quad (2)\ a_{m,n} = \frac{1}{2^{m-n}}; \quad (3)\ a_{m,n} = \frac{1}{2^{(m+1)(n+1)}}; \quad (4)\ a_{m,n} = \frac{1}{2^{mn}};$$

$$(5)\ a_{m,n} = \frac{1}{2^{\max(m,n)}}; \quad (6)\ a_{m,n} = \frac{1}{2^{\min(m,n)}};$$

Si c'est le cas, écrire la somme de la famille comme une somme de série (puis déterminer sa valeur si la série est connue !)

Correction.

On pose, pour $m, n \in \mathbb{N}$,

$$V_m = \{(m,p) \in \mathbb{N}^2 \mid p \in \mathbb{N}\} \quad H_n = \{(p,n) \in \mathbb{N}^2 \mid p \in \mathbb{N}\} \quad \Delta_n = \{(p, n-p) \mid p \in [\![0,n]\!]\}.$$

1. On applique le théorème de sommation par paquets en utilisant la partition $(H_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$.
   - *Somme sur chaque paquet.* Soit $n \in \mathbb{N}$. On considère la famille $(\frac{1}{2^{m+n}})_{m\in\mathbb{N}}$. Il s'agit d'une suite de réels positifs donc elle est sommable si, et seulement si, la série $\sum_{m\geq 0}\frac{1}{2^{m+n}}$ converge.
     On note $a = \frac{1}{2^n}$. Pour tout $m \in \mathbb{N}$, on a $0 \leq \frac{1}{2^{m+n}} = \frac{a}{2^m}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison $\frac{1}{2}$ dans $]-1,1[$) donc $\sum_{m\geq 0}\frac{1}{2^{m+n}}$ converge. Ainsi la famille $(\frac{1}{2^{m+n}})_{m\in\mathbb{N}}$ est sommable et on a :
     $$s_n = \sum_{m\in\mathbb{N}}\frac{1}{2^{m+n}} = \sum_{m=0}^{+\infty}\frac{1}{2^{m+n}} = a\sum_{m=0}^{+\infty}\frac{1}{2^m} = a\frac{1}{1-\frac{1}{2}} = 2a = \frac{2}{2^n}.$$
   - *Somme des sommes sur les paquets.* La famille $(s_n)_{n\in\mathbb{N}}$ est une suite de réels positifs, elle est donc sommable si, et seulement si, la série $\sum s_n$ converge. On a, pour tout $n \in \mathbb{N}$, $s_n = 2\frac{1}{2^{n-1}}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison $\frac{1}{2}$ dans $]-1,1[$) donc $\sum s_n$ converge. Ainsi la famille $(s_n)_{n\in\mathbb{N}}$ est sommable et on a :
     $$\sum_{n\in\mathbb{N}} s_n = \sum_{n=0}^{+\infty} s_n = 2\frac{1}{1-\frac{1}{2}} = 4.$$

Il en résulte, d'après le théorème de sommation par paquets, que la famille $(\frac{1}{2^{m+n}})_{(m,n)\in\mathbb{N}^2}$ est sommable et on a :

$$\sum_{(m,n)\in\mathbb{N}^2} \frac{1}{2^{m+n}} = \sum_{n\in\mathbb{N}} s_n = 4.$$

Remarque : comme $(\frac{1}{2^{m+n}})_{(m,n)\in\mathbb{N}^2}$ est une famille sommable, on a toujours d'après le théorème de sommation par paquet mais sur la partition $(\Delta_k)_{k\in\mathbb{N}}$ de $\mathbb{N}^2$, la série $\displaystyle\sum_{k\geq 0}\sum_{(m,n)\in\Delta_k} \frac{1}{2^{m+n}} = \sum_{k\geq 0} \frac{k+1}{2^k}$ converge et on a :

$$\sum_{k\geq 0} \frac{k+1}{2^k} = \sum_{k\in\mathbb{N}}\sum_{(m,n)\in\Delta_k} \frac{1}{2^{m+n}} = \sum_{(m,n)\in\mathbb{N}^2} \frac{1}{2^{m+n}} = 4.$$

2. — *1ère façon :* en exhibant une une sous-famille non sommable.

   On considère la sous-famille $(\frac{1}{2^{n-n}})_{(n,n)\in\mathbb{N}^2} = (1)_{n\in\mathbb{N}}$. Il s'agit d'une suite qui est le terme général d'une série grossièrement divergente donc cette sous-famille n'est pas sommable. Par suite, la famille $(\frac{1}{2^{m-n}})_{(m,n)\in\mathbb{N}^2}$ n'est pas sommable.

   — *2ème façon :* en utilisant le théorème de sommation par paquets sur la partition $(V_m)_{m\in\mathbb{N}}$ de $\mathbb{N}^2$.

   - *Somme sur chaque paquet.* Soit $m \in \mathbb{N}$. On considère la famille $(\frac{1}{2^{m-n}})_{n\in\mathbb{N}}$. Il s'agit d'une suite de réels positifs donc elle est sommable si, et seulement si, la série $\sum_{n\geq 0} \frac{1}{2^{m-n}}$ converge.
     On note $a = \frac{1}{2^m} \neq 0$. Pour tout $n \in \mathbb{N}$, on a $0 \leq \frac{1}{2^{m-n}} = a2^m$ qui est le terme général d'une série grossièrement divergente. Ainsi, la famille $(\frac{1}{2^{m-n}})_{n\in\mathbb{N}}$ n'est pas sommable.

   Par suite, d'après le théorème de sommation par paquets, la famille $(\frac{1}{2^{m-n}})_{(m,n)\in\mathbb{N}^2}$ n'est pas sommable.

3. On applique le théorème de sommation par paquets en utilisant la partition $(H_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$.

   - *Somme sur chaque paquet.* Soit $n \in \mathbb{N}$. On considère la famille $(\frac{1}{2^{(m+1)(n+1)}})_{m\in\mathbb{N}}$. Il s'agit d'une suite de réels positifs donc elle est sommable si, et seulement si, la série $\sum_{m\geq 0} \frac{1}{2^{(m+1)(n+1)}}$ converge.
     Pour tout $m \in \mathbb{N}$, on a $0 \leq \frac{1}{2^{(m+1)(n+1)}} \leq \frac{1}{2^m}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison dans $]-1,1[$). Ainsi, par comparaison, $\sum_{m\geq 0} \frac{1}{2^{(m+1)(n+1)}}$ converge et donc la famille $(\frac{1}{2^{(m+1)(n+1)}})_{m\in\mathbb{N}}$ est sommable. De plus, on a :

   $$s_n = \sum_{(m,n)\in H_n} \frac{1}{2^{(m+1)(n+1)}} = \sum_{m=0}^{+\infty} \frac{1}{(2^{n+1})^{m+1}} = \frac{1}{2^{n+1}} \frac{1}{1-\frac{1}{2^{n+1}}} = \frac{1}{2^{n+1}-1}.$$

   - *Somme des sommes des paquets.* La famille $(s_n)_{n\in\mathbb{N}}$ est une suite de réels positifs, elle est donc sommable si, et seulement si, la série $\sum s_n$ converge.
     On a, pour tout $n \in \mathbb{N}$, $s_n \leq \frac{1}{2^{n+1}}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison $\frac{1}{2}$ dans $]-1,1[$) donc $\sum s_n$ converge. Ainsi la famille $(s_n)_{n\in\mathbb{N}}$ est sommable et on a :

   $$\sum_{n\in\mathbb{N}} s_n = \sum_{n=0}^{+\infty} s_n = \sum_{n=0}^{+\infty} \frac{1}{2^{n+1}-1} = \sum_{n=1}^{+\infty} \frac{1}{2^n-1}$$

Il en résulte, d'après le théorème de sommation par paquets, que la famille $(\frac{1}{2^{(m+1)(n+1)}})_{(m,n)\in\mathbb{N}^2}$ est sommable et on a :

$$\sum_{(m,n)\in\mathbb{N}^2} \frac{1}{2^{(m+1)(n+1)}} = \sum_{n\in\mathbb{N}} s_n = \sum_{n=1}^{+\infty} \frac{1}{2^n - 1}.$$

Remarque : voir exercice d'après !

4. On exhibe une une sous-famille non sommable.

   On considère la sous-famille $(\frac{1}{2^{0\times n}})_{(0,n)\in\mathbb{N}^2} = (1)_{n\in\mathbb{N}}$. Il s'agit d'une suite qui est le terme général d'une série grossièrement divergente donc cette sous-famille n'est pas sommable. Par suite, la famille $(\frac{1}{2^{mn}})_{(m,n)\in\mathbb{N}^2}$ n'est pas sommable.

5. On applique le théorème de sommation par paquets en utilisant la partition $(H_n)_{n\in\mathbb{N}}$ de $\mathbb{N}^2$.

   - *Somme sur chaque paquet.* Soit $n \in \mathbb{N}$. On considère la famille $(\frac{1}{2^{\max(m,n)}})_{m\in\mathbb{N}}$. Il s'agit d'une suite de réels positifs donc elle est sommable si, et seulement si, la série $\sum_{m\geq 0} \frac{1}{2^{\max(m,n)}}$ converge.
     Pour tout $m \geq n$, on a $0 \leq \frac{1}{2^{\max(m,n)}} = \frac{1}{2^m}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison $\frac{1}{2}$ dans $]-1,1[$). Ainsi, par comparaison, $\sum_{m\geq 0} \frac{1}{2^{\max(m,n)}}$ converge et donc la famille $(\frac{1}{2^{\max(m,n)}})_{m\in\mathbb{N}}$ est sommable. De plus, on a :
     $$s_n = \sum_{(m,n)\in H_n} \frac{1}{2^{\max(m,n)}} = \sum_{m=0}^{n-1} \frac{1}{2^n} + \sum_{m=n}^{+\infty} \frac{1}{2^m} = \frac{n}{2^n} + \frac{1}{2^n}\frac{1}{1-\frac{1}{2}} = \frac{n+2}{2^n}.$$

   - *Somme des sommes des paquets.* La famille $(s_n)_{n\in\mathbb{N}}$ est une suite de réels positifs, elle est donc sommable si, et seulement si, la série $\sum s_n$ converge.
     On a, pour tout $n \in \mathbb{N}$, $s_n = \frac{n+2}{2^n}$ qui est le terme général d'une série convergente (par exemple avec la régle de d'Alemebert ou avec le critère de Riemann en multipliant par n^2) donc $\sum s_n$ converge. Ainsi la famille $(s_n)_{n\in\mathbb{N}}$ est sommable et on a :
     $$\sum_{n\in\mathbb{N}} s_n = \sum_{n=0}^{+\infty} s_n = \sum_{n=0}^{+\infty} \frac{n+2}{2^n} = 6$$
     Pour la dernière égalité, utiliser la remarque finale de la question 1 ou encore utiliser la somme de la série entière $\sum nx^n$.

   Il en résulte, d'après le théorème de sommation par paquets, que la famille $(\frac{1}{2^{\max(m,n)}})_{(m,n)\in\mathbb{N}^2}$ est sommable et on a :
   $$\sum_{(m,n)\in\mathbb{N}^2} \frac{1}{2^{\max(m,n)}} = \sum_{n\in\mathbb{N}} s_n = 6.$$

6. On exhibe une une sous-famille non sommable.

   On considère la sous-famille $(\frac{1}{2^{\min(0,n)}})_{(0,n)\in\mathbb{N}^2} = (1)_{n\in\mathbb{N}}$. Il s'agit d'une suite qui est le terme général d'une série grossièrement divergente donc cette sous-famille n'est pas sommable. Par suite, la famille $(\frac{1}{2^{\min(m,n)}})_{(m,n)\in\mathbb{N}^2}$ n'est pas sommable.

Exercice 9.

Pour un entier $n \in \mathbb{N}^*$, on note $d(n) = \#\{d \in \mathbb{N} \mid d|n\}$ le nombre de diviseurs positifs de n. Montrer que les séries $\sum_{n\geq 1} \frac{\delta(n)}{2^n}$ et $\sum_{n\geq 1} \frac{\delta(n)}{n^2}$ convergent et que :

$$\sum_{n=1}^{+\infty} \frac{d(n)}{2^n} = \sum_{n=1}^{+\infty} \frac{1}{2^n - 1} \quad \text{et} \quad \sum_{n=1}^{+\infty} \frac{d(n)}{n^2} = \frac{\pi^4}{36}.$$

Correction.

Pour $n \in \mathbb{N}^*$, on note $P_n = \{(p,q) \in \mathbb{N}^2 \mid pq = n\}$ et on remarque que $d(n) = \#P_n$. De plus, $(P_n)_{n\in\mathbb{N}^*}$ forme une partition de $(\mathbb{N}^*)^2$. On applique le théorème de sommation par paquets aux familles de termes généraux $a_{p,q} = \frac{1}{2^{pq}}$ et $b_{p,q} = \frac{1}{p^2q^2}$: tout d'abord sur la partition verticale $(V_p)_{p\in\mathbb{N}^*}$ où $V_p = \{(p,q) \in (\mathbb{N}^*)^2\}$ avec $p \in \mathbb{N}^*$ puis sur la partition $(P_n)_{n\in\mathbb{N}^*}$ précédente.

- *Somme sur chaque paquet.* Soit $p \in \mathbb{N}^*$. On considère les familles $(\frac{1}{2^{pq}})_{(p,q)\in V_p} = (\frac{1}{2^{pq}})_{q\in\mathbb{N}^*}$ et $(\frac{1}{p^2q^2})_{(p,q)\in V_p} = (\frac{1}{p^2q^2})_{q\in\mathbb{N}^*}$. Il s'agit de suites de réels positifs donc elles sont sommables si, et seulement si, les séries associées convergent.
  Pour tout $q \in \mathbb{N}^*$, on a $0 \leq \frac{1}{2^{pq}} = \frac{1}{(2^p)^q}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison $\frac{1}{2^p}$ dans $]-1,1[$). Ainsi, $\sum_{q\geq 1} \frac{1}{2^{pq}}$ converge et donc la famille $(\frac{1}{2^{pq}})_{q\in\mathbb{N}^*}$ est sommable. De plus, on a :

  $$\alpha_p = \sum_{(p,q)\in V_p} \frac{1}{2^{pq}} = \sum_{q=1}^{+\infty} \frac{1}{(2^p)^q} = \frac{1}{2^p} \frac{1}{1 - \frac{1}{2^p}} = \frac{1}{2^p - 1}.$$

  Pour tout $q \in \mathbb{N}^*$, on a $0 \leq \frac{1}{p^2q^2} \leq \frac{1}{q^2}$ qui est le terme général d'une série convergente (critère de Riemann $2 > 1$). Ainsi, par comparaison, $\sum_{q\geq 1} \frac{1}{p^2q^2}$ converge et donc la famille $(\frac{1}{p^2q^2})_{q\in\mathbb{N}^*}$ est sommable. De plus, on a :

  $$\beta_p = \sum_{(p,q)\in V_p} \frac{1}{p^2q^2} = \frac{1}{p^2} \sum_{q=1}^{+\infty} \frac{1}{q^2} = \frac{\pi^2}{6} \frac{1}{p^2}.$$

- *Somme des sommes des paquets.*
  - La famille $(\alpha_p)_{n\in\mathbb{N}^*}$ est une suite de réels positifs, elle est donc sommable si, et seulement si, la série $\sum_{p\geq 1} = \alpha_p$ converge.
    On a, pour tout $p \in \mathbb{N}^*$, $\alpha_p = \frac{1}{2^p - 1} \leq \frac{1}{2^p}$ qui est le terme général d'une série convergente (terme d'une série géométrique de raison $\frac{1}{2}$ dans $]-1,1[$) donc par comparaison $\sum_{p\geq 1} \alpha_p$ converge. Ainsi la famille $(\alpha_p)_{p\in\mathbb{N}^*}$ est sommable et on a :

    $$\sum_{p\in\mathbb{N}^*} \alpha_p = \sum_{p=1}^{+\infty} \frac{1}{2^p - 1}.$$

  - La famille $(\beta_p)_{n\in\mathbb{N}^*}$ est une suite de réels positifs, elle est donc sommable si, et seulement si, la série $\sum_{p\geq 1} = \beta_p$ converge.
    On a, pour tout $p \in \mathbb{N}^*$, $\beta_p = \frac{\pi^2}{6} \frac{1}{p^2}$ qui est le terme général d'une série convergente (critère de Riemann $2 > 1$) donc par comparaison $\sum_{p\geq 1} \beta_p$ converge. Ainsi la famille

$(\beta_p)_{p\in\mathbb{N}^*}$ est sommable et on a :

$$\sum_{p\in\mathbb{N}^*}\beta_p=\frac{\pi^2}{6}\sum_{p=1}^{+\infty}\frac{1}{p^2}=\left(\frac{\pi^2}{6}\right)^2=\frac{\pi^4}{36}.$$

Il en résulte, d'après le théorème de sommation par paquets, que les familles les familles $(\frac{1}{2^{pq}})_{(p,q)\in(\mathbb{N}^*)^2}$ et $(\frac{1}{p^2q^2})_{(p,q)\in(\mathbb{N}^*)^2}$ sont sommables et on a :

$$\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{2^{pq}}=\sum_{p=1}^{+\infty}\frac{1}{2^p-1}\quad\text{et}\quad\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{p^2q^2}=\frac{\pi^4}{36}.$$

De plus, toujours d'après le théorème de sommation par paquets, comme ces familles sont sommables, leurs sommes sur $(\mathbb{N}^*)^2$ sont égales à leurs sommes des sommes sur la partition $(P_n)_{n\in\mathbb{N}^*}$. Or, pour $n\in\mathbb{N}^*$, on a :

$$\sum_{(p,q)\in P_n}\frac{1}{2^{pq}}=\sum_{p|n}\frac{1}{2^n}=\frac{d(n)}{2^n}\quad\text{et}\quad\sum_{(p,q)\in P_n}\frac{1}{p^2q^2}=\sum_{p|n}\frac{1}{n^2}=\frac{d(n)}{n^2}$$

Et donc :

$$\sum_{p=1}^{+\infty}\frac{1}{2^p-1}=\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{2^{pq}}=\sum_{n\in\mathbb{N}^*}\sum_{(p,q)\in P_n}\frac{1}{2^{pq}}=\sum_{n=1}^{+\infty}\frac{d(n)}{2^n}.$$

et

$$\frac{\pi^4}{36}=\sum_{(p,q)\in(\mathbb{N}^*)^2}\frac{1}{p^2q^2}=\sum_{n\in\mathbb{N}^*}\sum_{(p,q)\in P_n}\frac{1}{p^2q^2}=\sum_{n=1}^{+\infty}\frac{d(n)}{n^2}.$$

Partie C

Espaces probabilisés

1. Espaces probabilisables

a. Définitions

Définition 9. *Tribu*

Soit Ω un ensemble non vide et $\mathcal{T}\subset\mathcal{P}(\Omega)$. On dit que $\mathcal{T}$ est une **tribu** sur Ω si :

- $\Omega\in\mathcal{T}$;
- $\mathcal{T}$ est stable par passage au complémentaire i.e. pour tout $A\in\mathcal{T}$, $\overline{A}\in\mathcal{T}$;
- $\mathcal{T}$ est stable par réunion dénombrable i.e. pour toute suite $(A_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathcal{T}$,

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}.$$

Remarque 5.

On trouve parfois la terminologie σ-algèbre pour désigner une tribu et la notation $\mathcal{A}$ au lieu de $\mathcal{T}$.

Définition 10. *Espace probabilisable*

Soit Ω un ensemble non vide et $\mathcal{T}$ une tribu sur Ω. Le couple $(\Omega, \mathcal{T})$ est appelé **espace probabilisable**.

Exemple 6.

Soit Ω un ensemble non vide.

- $\mathcal{T} = \{\emptyset, \Omega\}$ est une tribu appelée **tribu grossière** ;
- $\mathcal{T} = \mathcal{P}(\Omega)$ est une tribu ;
- Si $A \subset \Omega$, $\mathcal{T} = \{\emptyset, A, \overline{A}, \Omega\}$ est une tribu.

Exercice 10.

Soit Ω un ensemble non vide et $\mathcal{F} \subset \mathcal{P}(\Omega)$.

1. Montrer que l'intersection de toutes les tribus contenant $\mathcal{F}$ (en existe-t-il ?) est une tribu. On note cette tribu $\sigma(\mathcal{F})$ et on l'appelle **tribu engendrée par** $\mathcal{F}$.
2. Montrer que $\sigma(\mathcal{F})$ est la plus petite tribu contenant $\mathcal{F}$.

Remarque 6. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On utilise la terminologie suivante :

- Ω est l'**univers** ;
- les éléments A de $\mathcal{T}$ sont les **événements** ;

b. Les événements

Proposition 20.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On a les propriétés suivantes :

- $\emptyset \in \mathcal{T}$;
- $\mathcal{T}$ est stable par réunion finie ;

— $\mathcal{T}$ est stable par intersection au plus dénombrable.
— pour tout $A, B \in \mathcal{T}$, $A \setminus B \in \mathcal{T}$.

Démonstration.

— On a $\Omega \in \mathcal{T}$ donc $\emptyset = \overline{\Omega} \in \mathcal{T}$.
— Soit $n \in \mathbb{N}^*$ et $A_1, ..., A_n \in \mathcal{T}$. On considère la suite $(A_k)_{k\in\mathbb{N}}$ telle que $A_k = A_n \in \mathcal{T}$ pour tout entier $k > n$. Par stabilité de $\mathcal{T}$ par union dénombrable, on a $\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{T}$. Par suite, on a :

$$A_1 \cup ... \cup A_{n-1} \cup A_n = A_1 \cup ...A_{n-1} \cup \left(\bigcup_{\substack{k\in\mathbb{N} \\ k\geq n}} A_n \right) = \bigcup_{k\in\mathbb{N}} A_k \in \mathcal{T}.$$

— Pour toute famille $(A_i)_{i\in I}$ au plus d'éléments de $\mathcal{T}$, $(\overline{A_i})_{i\in I}$ est une famille au plus dénombrable d'éléments de $\mathcal{T}$ donc, par stabilité de $\mathcal{T}$ par réunion au plus dénombrable et par passage au complémentaire :

$$\bigcap_{i\in I} A_i = \overline{\bigcup_{i\in I} \overline{A_i}} \in \mathcal{T}.$$

— Soit $A, B \in \mathcal{T}$. Par stabilité de $\mathcal{T}$ par intersection finie et par passage au complémentaire, on a :

$$A \setminus B = A \cap \overline{B} \in \mathcal{T}.$$

□

Remarque 7. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $A, B \in \mathcal{T}$ des événements. On utilise la terminologie suivante concernant les événements :

— Ω est l'événement **certain** et $\emptyset$ l'événement **impossible** ;
— $\overline{A}$ est l'événement **contraire** de A ;
— si $A \cap B = \emptyset$, on dit que A et B sont des événements **incompatibles**.

Définition 11. *Système complet d'événements*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $(A_i)_{i\in I}$ une famille au plus dénombrable d'événements. On dit que la famille $(A_i)_{i\in I}$ est un **système complet d'événements** si elle forme une partition de Ω.

2. Espaces probabilisés

a. Définition

Définition 12. *Probabilité*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et $P : \mathcal{T} \to [0,1]$ une application. On dit que P est une **probabilité** sur $(\Omega, \mathcal{T})$ si :

i) $P(\Omega) = 1$;

ii) pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux disjoints, la série $\sum P(A_n)$ converge et :

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

Remarque 8. *Vocabulaire*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{T})$. Pour $A \in \mathcal{T}$, on dit que $P(A)$ est **la probabilité de l'événement** A.

Définition 13. *Espace probabilisé*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable et P une probabilité sur $(\Omega, \mathcal{T})$. Le triplet $(\Omega, \mathcal{T}, P)$ est appelé **espace probabilisé**.

Définition 14. *Événement négligeable/presque sûr*

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $A \in \mathcal{T}$. On dit que :

— l'événement A est **négligeable** si $P(A) = 0$.

— l'événement A est **presque sûr** si $P(A) = 1$.

b. Propriétés

Proposition 21.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $A, B \in \mathcal{T}$. On a les propriétés suivantes :

i) $P(\emptyset) = 0$;

ii) si A et B sont incompatibles, $P(A \cup B) = P(A) + P(B)$;

iii) $P(\overline{A}) = 1 - P(A)$;

iv) si $A \subset B$, $P(B \setminus A) = P(B) - P(A)$;

v) si $A \subset B$, $P(A) \leq P(B)$;

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

Démonstration.

i) Notons $c = P(\emptyset)$. Pour tout $n \in \mathbb{N}$, on pose $A_n = \emptyset \in \mathcal{T}$. Alors, pour tout $n, m \in \mathbb{N}$ avec $n \neq m$, $A_n \cap A_m = \emptyset \cap \emptyset = \emptyset$ donc $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à

deux disjoints. Par suite, la série $\sum P(A_n) = \sum c$ converge. Or, une suite constante est terme général d'une série convergente si, et seulement s'il s'agit de la suite nulle, d'où $P(\emptyset) = c = 0$.

ii) On suppose A et B incompatibles. Alors, pour $A_0 = A$, $A_1 = B$ et $A_n = \emptyset$ pour tout entier $n \geq 2$, $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à deux disjoints et donc, $\sum P(A_n)$ converge et on a, en utilisant i) :

$$\begin{aligned}
P(A \cup B) &= P\left(A_0 \cup A_1 \cup \emptyset\right) \\
&= P\left(A_0 \cup A_1 \cup \bigcup_{\substack{n\in\mathbb{N} \\ n\geq 2}} A_n\right) \\
&= P\left(\bigcup_{n\in\mathbb{N}} A_n\right) \\
&= \sum_{n=0}^{+\infty} P(A_n) \\
&= P(A_1) + P(A_2) + \sum_{n=2}^{+\infty} \underbrace{P(A_n)}_{P(\emptyset)=0} \\
P(A \cup B) &= P(A) + P(B).
\end{aligned}$$

iii) Comme $A \in \mathcal{T}$, $\overline{A} \in \mathcal{T}$ et comme A et $\overline{A}$ sont incompatibles, d'après ii), on a :

$$1 = P(\Omega) = P(A \cup \overline{A}) = P(A) + P(\overline{A}).$$

D'où $P(\overline{A}) = 1 - P(A)$.

iv) On suppose $A \subset B$. On a alors, par les diverses stabilités d'une tribu : $\overline{B} \in \mathcal{T}$, $B \setminus A = B \cap \overline{A} \in \mathcal{T}$ et $A \cup \overline{B} = \overline{B \setminus A} \in \mathcal{T}$; donc, $\overline{B}$ et A étant incompatibles, on a, en utilisant ii) et iii) :

$$\begin{aligned}
P(B \setminus A) &= 1 - P\left(A \cup \overline{B}\right) \\
&= 1 - (P(A) + P(\overline{B})) \\
&= 1 - (P(A) + 1 - P(B)) \\
P(B \setminus A) &= P(B) - P(A).
\end{aligned}$$

v) On suppose $A \subset B$. Comme P est à valeurs dans $[0, 1]$ et $B \setminus A \in \mathcal{T}$, on a $P(B \setminus A) \geq 0$. Or, d'après iv), on a $P(B \setminus A) = P(B) - P(A)$, d'où $P(B) - P(A) \geq 0$ et donc $P(A) \leq P(B)$.

vi) Par les diverses stabiltés d'une tribu, $A \cup B$, $A \cap B$, $A \setminus (A \cap B)$ et $B \setminus (A \cap B)$ appartiennent à $\mathcal{T}$. De plus, $(A \setminus (A \cap B), B \setminus (A \cap B), A \cap B)$ forme une partition de $A \cup B$ donc, $A \cap B$ étant inclus dans A et B, d'après ii) et iv) :

$$\begin{aligned}
P(A \cup B) &= P\left((A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B)\right) \\
&= P(A \setminus (A \cap B)) + P(B \setminus (A \cap B)) + P(A \cap B) \\
&= P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B) \\
P(A \cup B) &= P(A) + P(B) - P(A \cap B).
\end{aligned}$$

□

Exercice 11.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$. Montrer que si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements, alors :

1. $\sum P(A_n)$ converge et $\displaystyle\sum_{n=0}^{+\infty} P(A_n) = 1$.
2. pour tout $B \in \mathcal{T}$, $\sum P(A_n \cap B)$ converge et $\displaystyle\sum_{n=0}^{+\infty} P(A_n \cap B) = P(B)$.

3. Propriétés élémentaires des probabilités

Théorème 7. *Continuité croissante/décroissante*

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$.

- Si $(A_n)_{n\in\mathbb{N}}$ est croissante i.e. pour tout $n \in \mathbb{N}$, $A_n \subset A_{n+1}$, alors la suite $(P(A_n))_{n\in\mathbb{N}}$ converge et :
$$P(\bigcup_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P(A_n).$$
- Si $(A_n)_{n\in\mathbb{N}}$ est décroissante i.e. pour tout $n \in \mathbb{N}$, $A_{n+1} \subset A_n$, alors la suite $(P(A_n))_{n\in\mathbb{N}}$ converge et :
$$P(\bigcap_{n\in\mathbb{N}} A_n) = \lim_{n\to+\infty} P(A_n).$$

Démonstration.

- Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements. Alors, pour tout $n \in \mathbb{N}$, on a $P(A_n) \leq P(A_{n+1})$ et $P(A_n) \leq 1$. Ainsi $(P(A_n))_{n\in\mathbb{N}}$ est une suite croissante et majorée de réels donc elle converge.

  On note $A_{-1} = \emptyset \in \mathcal{T}$.
  La série télescopique $\sum(P(A_n) - P(A_{n-1}))$ étant de même nature que la suite $(P(A_n))_{n\in\mathbb{N}}$, cette série converge, et on a :
$$\sum_{n=0}^{+\infty}(P(A_n) - P(A_{n-1})) = \lim_{n\to+\infty} P(A_n) - P(A_{-1}) = \lim_{n\to+\infty} P(A_n).$$
  On pose, pour $n \in \mathbb{N}$, $B_n = A_n \setminus A_{n-1} \in \mathcal{T}$. Pour tout $n \in \mathbb{N}$, on a $P(B_n) = P(A_n) - P(A_{n-1})$ car, pour $n = 0$, $A_{-1} = \emptyset \subset A_0$ et pour $n \geq 1$, la suite $(A_n)_{n\in\mathbb{N}}$ est croissante.
  De plus, pour $n, m \in \mathbb{N}$ avec $n > m$, on a, toujours par croissance, $B_m \subset A_m \subset A_{n-1}$ car $n - 1 \geq m$, donc :
$$B_m \cap B_n = B_m \cap (A_n \cap \overline{A_{n-1}}) = A_n \cap \underbrace{(B_m \cap \overline{A_{n-1}})}_{=\emptyset} = \emptyset.$$

Ainsi, les B_n sont des événements deux à deux disjoints. Par suite, $\sum P(B_n)$ converge et on a :

$$P\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n=0}^{+\infty} \underbrace{P(B_n)}_{=P(A_n)-P(A_{n-1})} = \lim_{n\to+\infty} P(A_n).$$

Or, on a, par un raisonnement par double inclusion :

$$\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n,$$

donc :

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = P\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \lim_{n\to+\infty} P(A_n).$$

- Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements. Alors, pour tout $n \in \mathbb{N}$, $\overline{A_n} \in \mathcal{T}$ et $\overline{A_{n+1}} \subset \overline{A_n}$ car $A_n \subset A_{n+1}$. Par suite, $(\overline{A_n})_{n\in\mathbb{N}}$ est une suite croissante d'événements et donc, d'après la point précédent, $(P(\overline{A_n}))_{n\in\mathbb{N}}$ converge et :

$$P\left(\bigcup_{n\in\mathbb{N}} \overline{A_n}\right) = \lim_{n\to+\infty} P(\overline{A_n}).$$

Or, pour tout $n \in \mathbb{N}$, $P(A_n) = 1 - P(\overline{A_n})$ donc, par combinaison linéaire de suites convergentes, $(P(A_n))_{n\in\mathbb{N}}$ converge et $\lim_{n\to+\infty} P(A_n) = 1 - \lim_{n\to+\infty} P(\overline{A_n})$.

De plus, on a $\bigcap_{n\in\mathbb{N}} \overline{A_n} = \overline{\bigcup_{n\in\mathbb{N}} A_n}$, donc :

$$P\left(\bigcap_{n\in\mathbb{N}} A_n\right) = 1 - P\left(\bigcup_{n\in\mathbb{N}} \overline{A_n}\right) = 1 - \lim_{n\to+\infty} P(\overline{A_n}) = \lim_{n\to+\infty} P(A_n).$$

□

Proposition 22.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.

- pour $n \in \mathbb{N}$, on a :

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

- On a :

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

où $\sum_{n=0}^{+\infty} P(A_n) \in \mathbb{R}_+ \cup \{+\infty\}$.

Démonstration.

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements. Alors, par stabilité d'une tribu par réunion au plus dénombrable, pour $n \in \mathbb{N}$, $B_n = \bigcup_{k=0}^{n} A_k$ est un événement et $\bigcup_{n\in\mathbb{N}} A_n$ est un événement.

- On démontre ce résultat par récurrence sur $\mathbb{N}$.
  — Initialisation. On a

$$P(B_0) = P(A_0) \leq P(A_0) = \sum_{k=0}^{0} P(A_k).$$

  donc la propriété est vraie au rang $n = 0$.
  — Hérédité. Soit $n \in \mathbb{N}$. On suppose $P(B_n) \leq \sum_{k=0}^{n} P(A_k)$.
  On a $B_{n+1} = B_n \cup A_{n+1}$, donc :

$$P(B_{n+1}) = P(B_n) + P(A_{n+1}) - \underbrace{P(B_n \cap A_{n+1})}_{\geq 0} \leq P(B_n) + P(A_{n+1}).$$

  Ainsi, par hypothèse de récurrence,

$$P(B_{n+1}) \geq \underbrace{P(B_n)}_{\leq \sum_{k=0}^{n} P(A_k)} + P(A_{n+1}) \leq \sum_{k=0}^{n+1} P(A_k).$$

  Ce qui achève le raisonnement par récurrence.
  Ainsi, pour tout $n \in \mathbb{N}$,

$$P\left(\bigcup_{k=0}^{n} A_k\right) = P(B_n) \leq \sum_{k=0}^{n} P(A_k).$$

- Si $\sum P(A_n)$ diverge, l'inégalité est vérifiée. On suppose que $\sum P(A_n)$ converge. La suite $(B_n)_{n\in\mathbb{N}}$ est une suite croissante d'évènements, donc par continuité croissante, $(P(B_n))_{n\in\mathbb{N}}$ converge et on a :

$$P\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \lim_{n\to+\infty} P(B_n).$$

  Ainsi, d'après le point précédent et le théorème des gendarmes, on a :

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

  Puis on conclut en remarquant que $\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} A_n$.

□

Corollaire 4.

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathcal{T}$. Si, pour tout $n \in \mathbb{N}$, A_n est négligeable, alors $\bigcup_{n\in\mathbb{N}} A_n$ est négligeable.

Démonstration.

On applique le deuxième point de la proposition précédente à la suite $(A_n)_{n\in\mathbb{N}}$; on a :

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

D'où $P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 0$ et donc $\bigcup_{n\in\mathbb{N}} A_n$ est négligeable. □

4. Probabilité sur un univers au plus dénombrable

a. Caractérisation

Théorème 8.

Soit Ω un ensemble non vide au plus dénombrable.

i) Soit P une probabilité sur $(\Omega, \mathcal{P}(\Omega))$ et on pose, pour $\omega \in \Omega$, $p_\omega = P(\{\omega\})$. Alors, pour tout $\omega \in \Omega$, $p_\omega \geq 0$ et :

$$\sum_{\omega\in\Omega} p_\omega = 1.$$

ii) Si $(p_\omega)_{\omega\in\Omega}$ est une famille sommable de réels positifs telle que $\sum_{\omega\in\Omega} p_\omega = 1$, alors il existe une unique probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ telle que, pour tout $\omega \in \Omega$, $P(\{\omega\}) = p_\omega$. Elle est définie, pour $A \in \mathcal{P}(\Omega)$, par :

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

Démonstration.

Comme, pour tout $\omega, \omega' \in \Omega$, $\{\omega\} \in \mathcal{P}(\Omega)$, $\{\omega\} \cap \{\omega'\} = \emptyset$ et $\bigcup_{\omega\in\Omega}\{\omega\} = \Omega$, la famille au plus dénombrable $(\{\omega\})_{\omega\in\Omega}$ est un sytème complet d'évènements.

i) La famille au plus dénombrable $(\{\omega\})_{\omega\in\Omega}$ étant un sytème complet d'évènements, la famille $(P(\{\omega\}))_{\omega\in\Omega}$ est sommable et on a :

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

ii) Soit $A \in \mathcal{P}(\Omega)$ un évènement. Alors la famille $(p_\omega)_{\omega\in A}$ est sommable comme sous-famille de la famille sommable $(p_\omega)_{\omega\in\Omega}$.

On définit alors l'application $P : \mathcal{P}(\Omega) \to \mathbb{R}$, pour $A \in \mathcal{P}(\Omega)$, par :

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

Montrons que P est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.

— On a, pour tout $A \in \mathcal{P}(\Omega)$, $P(A) = \sum_{\omega\in A} p_\omega \geq 0$ car $p_\omega \geq 0$ pour tout $\omega \in \Omega$ et $P(A) = \sum_{\omega\in A} p_\omega \leq \sum_{\omega\in\Omega} p_\omega = 1$ car $(p_\omega)_{\omega\in A}$ est une sous-famille de la famille sommable $(p_\omega)_{\omega\in\Omega}$. Ainsi, P est à valeurs dans $[0, 1]$.

— on a $P(\Omega) = \sum_{\omega\in\Omega} p_\omega = 1$.

— Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux disjoints. On pose $A = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{P}(\Omega)$. Alors $(p_\omega)_{\omega\in A}$ est une famille sommable, donc, d'après le théorème de sommation par paquets appliqué à la partition $(A_n)_{n\in\mathbb{N}}$ de A, la famille $(P(A_n))_{n\in\mathbb{N}} = (\sum_{\omega\in A_n} p_\omega)_{n\in\mathbb{N}}$ est sommable et donc $\sum P(A_n)$ converge, et de plus, on a :

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = P(A) = \sum_{\omega\in A} p_\omega = \sum_{n\in\mathbb{N}}\sum_{\omega\in A_n} p_\omega = \sum_{n=0}^{+\infty} P(A_n).$$

Il en résulte que P est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.
Finalement, on a, pour tout $\omega \in \Omega$, $\{\omega\} \in \mathcal{P}(\Omega)$ et $P(\{\omega\}) = \sum_{\omega'\in\{\omega\}} p_{\omega'} = p_\omega$.
Par construction, P est unique.

□

b. Exemples de probabilités sur $\mathbb{N}$

Exemple 7.

1. **Loi géométrique.** Soit $p \in]0,1[$. On définit la probabilité P suivante sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par $p_n = P(\{n\}) = (1-p)^n p$ pour $n \in \mathbb{N}$.
2. **Loi de Poisson.** Soit $\lambda \in \mathbb{R}_+^*$. On définit la probabilité P suivante sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ par $p_n = P(\{n\}) = e^{-\lambda}\dfrac{\lambda^n}{n!}$ pour $n \in \mathbb{N}$.

Pour justifier la bonne définition de ces probabilités, d'après le théorème précédent, il suffit de vérifier que la famille $(p_n)_{\in\mathbb{N}}$ est sommable et de somme 1.

1. La suite $(p_n)_{n\in\mathbb{N}} = ((1-p)^n p)_{n\in\mathbb{N}}$ est une suite géométrique de raison $1-p \in]0,1[$ donc $\sum p_n$ converge et ainsi, $(p_n)_{\in\mathbb{N}}$ est sommable car la suite est à terme positifs. De plus, on a :
$$\sum_{n\in\mathbb{N}} p_n = p\sum_{n=0}^{+\infty}(1-p)^n = \frac{p}{1-(1-p)} = 1.$$
Donc il s'agit bien d'une probabilité.
2. La suite $(p_n)_{n\in\mathbb{N}} = (e^{-\lambda}\frac{\lambda^n}{n!})_{n\in\mathbb{N}}$ est le terme général d'une série convergente (par exemple en utilisant la règle de D'Alembert) et ainsi, $(p_n)_{\in\mathbb{N}}$ est sommable car la suite est à terme positifs. De plus, on a :
$$\sum_{n\in\mathbb{N}} p_n = e^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!} = e^{-\lambda}.e^{\lambda} = 1.$$
Donc il s'agit bien d'une probabilité.

Partie D

Probabilités conditionnelles

Dans cette partie, $(\Omega, \mathcal{T}, P)$ désigne un espace probabilisé.

1. Conditionnement

Définition 15. *Probabilité conditionnelle*

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. Pour $A \in \mathcal{T}$, on appelle **probabilité (conditionnelle) de** A **sachant** B et on note $P_B(A)$ ou encore $P(A|B)$ la quantité :

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

L'application $P_B : \mathcal{T} \to [0,1]$ définie par :

$$P_B : A \mapsto P_B(A)$$

est appelée **probabilité conditionnelle sachant** B.

Théorème 9.

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. La probabilité conditionnelle $P_B : \mathcal{T} \to \mathbb{R}$ sachant B est une probabilité sur l'espace probabilisable $(\Omega, \mathcal{T})$.

Démonstration.

Montrons que P_B une probabilité sur $(\Omega, \mathcal{T})$.

- Pour tout $A \in \mathcal{T}$, $P_B(A) = \frac{P(A \cap B)}{P(B)} \geq 0$ car P est à valeurs positives et comme $A \cap B \subset B$, on a $P(A \cap B) \leq P(B)$ d'où $P_B(A) = \frac{P(A \cap B)}{P(B)} \leq 1$. Par suite, P_B est à valeurs dans $[0,1]$.
- On a $P_B(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.
- Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements deux à deux disjoints. On a :

$$\left(\bigcup_{n \in \mathbb{N}} A_n \right) \cap B = \bigcup_{n \in \mathbb{N}} (A_n \cap B)$$

De plus, pour tout $n, m \in \mathbb{N}$ avec $n \neq m$, $(A_n \cap B) \cap (A_m \cap B) = \underbrace{(A_n \cap A_m)}_{=\emptyset} \cap B = \emptyset$;

donc $(A_n \cap B)_{n \in \mathbb{N}}$ est une suite d'événements deux à deux disjoints.

Par suite, $\sum P(A_n \cap B)$ converge, donc $\sum P_B(A_n) = \sum \frac{P(A_n \cap B)}{P(B)}$ également et on a :

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

Il en résulte que P_B est une une probabilité sur $(\Omega, \mathcal{T})$. □

2. Formules conditionnelles

a. Formule des probabilités composées

Théorème 10. *Formule des probabilités composées*

i) Soit $A, B \in \mathcal{T}$ tel que $P(B) > 0$. On a :

$$P(A \cap B) = P(A|B)P(B).$$

ii) Soit $n \in \mathbb{N}$ avec $n \geq 2$ et $A_1, \ldots, A_n \in \mathcal{T}$ tels que $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$. On a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times \ldots \times P\left(A_n \middle| \bigcap_{i=1}^{n-1} A_i\right).$$

Démonstration.

i) Soit $A, B \in \mathcal{T}$ tel que $P(B) > 0$. On a $P(A|B) = \frac{P(A\cap B)}{P(B)}$ donc :

$$P(A \cap B) = P(A|B)P(B).$$

ii) Montrons, par récurrence sur $\mathbb{N} \setminus \{0, 1\}$, que, pour tout entier $n \geq 2$, la propriété $\mathcal{P}_n =$"pour tout $A_1, \ldots, A_n \in \mathcal{T}$ tels que $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0$, on a :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \prod_{k=1}^{n-1} P\left(A_{k+1} \middle| \bigcap_{i=1}^{k-1} A_i\right).$$

— Initialisation. On a, d'après i) :

$$P\left(\bigcap_{i=1}^{2} A_i\right) = P(A_1 \cap A_2) = P(A_2|A_1)P(A_1) = P(A_1) \times \prod_{k=1}^{2-1} P\left(A_{k+1} \middle| \bigcap_{i=1}^{k-1} A_i\right).$$

Donc $\mathcal{P}_2$ est vraie.

— Hérédité. Soit un entier $n \geq 2$. On suppose $\mathcal{P}_n$ vraie. Soit $A_1, \ldots, A_{n+1} \in \mathcal{T}$ tels que

$P\left(\bigcap_{i=1}^{n} A_i\right) > 0$. Alors, comme $\bigcap_{i=1}^{n} A_i \subset \bigcap_{i=1}^{n-1} A_i$, on a :

$$0 < P\left(\bigcap_{i=1}^{n} A_i\right) \leq P\left(\bigcap_{i=1}^{n-1} A_i\right).$$

Ainsi, par hypothèse de récurrence,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \prod_{k=1}^{n-1} P\left(A_{k+1} \middle| \bigcap_{i=1}^{k-1} A_i\right).$$

De plus, on a, d'après i) :

$$P\left(\bigcap_{i=1}^{n+1} A_i\right) = P\left(A_{n+1} \cap \left(\bigcap_{i=1}^{n} A_i\right)\right) = P\left(A_{n+1} \middle| \bigcap_{i=1}^{n} A_i\right) P(\left(\bigcap_{i=1}^{n} A_i\right).$$

Ainsi, il en résulte que :

$$\begin{aligned} P\left(\bigcap_{i=1}^{n+1} A_i\right) &= P(A_1) \prod_{k=1}^{n-1} P\left(A_{k+1} \middle| \bigcap_{i=1}^{k-1} A_i\right) \times P\left(A_{n+1} \middle| \bigcap_{i=1}^{n} A_i\right) \\ P\left(\bigcap_{i=1}^{n+1} A_i\right) &= P(A_1) \prod_{k=1}^{n} P\left(A_{k+1} \middle| \bigcap_{i=1}^{k-1} A_i\right) \end{aligned}$$

Ce qui achève le raisonnement par récurrence.

Il en résulte que la propriété $\mathcal{P}_n$ est vraie pour tout $n \in \mathbb{N} \setminus \{0, 1\}$.

□

b. Formule des probabilités totales

Théorème 11. *Formule des probabilités totales*

Soit $(B_i)_{i \in I}$ un système complet d'événements. Pour tout événement $A \in \mathcal{T}$, on a :

$$P(A) = \sum_{i \in I} P(A|B_i)P(B_i).$$

avec la convention $P(A|B_i)P(B_i) = 0$ *si* $P(B_i) = 0$.

Démonstration.

Comme $(B_i)_{i \in I}$ est un système complet d'événements, on a :

$$\bigcup_{i \in I} B_i = \Omega \quad \text{et} \quad B_i \cap B_j = \emptyset \text{ si } i \neq j.$$

Par suite, on a :

$$A = \bigcup_{i \in I} (A \cap B_i) \text{ avec } (A \cap B_i) \cap (A \cap B_j) = \emptyset \text{ si } i \neq j.$$

En effet,

$$\bigcup_{i \in I} (A \cap B_i) = \bigcup_{i \in I} A \cap \bigcup_{i \in I} B_i = A \cap \Omega = A;$$

et

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset.$$

Par suite, par les propriétés de la fonction de probabilité P, on a :

$$P(A) = P\left(\bigcup_{i \in I} (A \cap B_i)\right) = \sum_{i \in I} P(A \cap B_i) = \sum_{i \in I} P(A|B_i)P(B_i).$$

□

Exercice 12.

Soit $B \in \mathcal{T}$ tel que $P(B) > 0$ et $P(\overline{B}) > 0$. Montrer que pour tout $A \in \mathcal{T}$, $P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$.

c. Formule de Bayes

Théorème 12. *Formule de Bayes*

Soit $A \in \mathcal{T}$ tel que $P(A) > 0$.

i) Soit $B \in \mathcal{T}$ tel que $P(B) > 0$. On a :

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

ii) Soit $(B_i)_{i \in I}$ un système complet d'événements. On a, pour tout $j \in I$:

$$P(B_j) = \frac{P(A|B_j)P(B_j)}{\sum_{i \in I} P(A|B_i)P(B_i)}.$$

d. Exercice d'application des formules

Exercice 13.

Dans une population, on donne la probabilité définie, pour $n \in \mathbb{N}$ par p_n qu'une famille ait n enfants où, α étant un réel strictement positif :

$$p_n = \alpha \frac{2^n}{n!}.$$

On suppose qu'il est équiprobable qu'un enfant soit une fille ou un garçon.

1. Déterminer le nombre α.
2. Déterminer la probabilité qu'une famille ait au moins un garçon.
3. Déterminer la probabilité qu'une famille ait exactement 2 garçons et 3 filles.
4. On cherche à déterminer la probabilité qu'une famille ait exactement deux enfants sachant qu'elle a au moins deux filles.
   a) Déterminer la probabilité qu'une famille ait exactement deux filles.
   b) Déterminer la probabilité qu'une famille ait au moins deux filles.
   c) Conclure.

Correction.

1. Soit Ω l'ensemble des compositions (d'enfants) des familles de la population. La suite d'événements $(A_n)_{n\in\mathbb{N}}$ définie par :
$$A_n\text{="La famille possède exactement } n \text{ enfants".}$$
étant un système complet d'événements, on a :
$$1 = P(\Omega) = P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} A_n = \sum_{n=0}^{+\infty} p_n.$$
Par suite,
$$1 = \sum_{n=0}^{+\infty} p_n = \alpha \sum_{n=0}^{+\infty} \frac{2^n}{n!} = \alpha e^2.$$
D'où $\alpha = e^{-2}$.
2. On note B l'événement :
$$B\text{="La famille possède uniquement des filles (ou aucun enfant)".}$$
Alors $\overline{B}$="La famille possède au moins un garçon" est l'événement dont la probabilité est recherchée.
On a, d'après la formule des probabilités totales :
$$P(B) = \sum_{n=0}^{+\infty} P(B|A_n)P(A_n) = \sum_{n=0}^{+\infty} \frac{1}{2^n} p_n = e^{-2}.e = \frac{1}{e}.$$
3. On note C l'événement recherché, D l'événement "la famille possède 2 garçons et 3 filles (au moins)" et on reprend la suite (A_n) de la question précédente. Alors on a $C = D \cap A_2$ donc :
$$P(C) = P(D \cap A_5) = P(D|A_5)P(A_5) = \left(\binom{5}{2}\frac{1}{2^2}\frac{1}{2^3}\right).e^{-2}\frac{2^5}{5!} = \frac{1}{12e^2}.$$
4. On note F l'événement "une famille a au moins deux filles" et on reprend la suite (A_n) de la question 1.

a) Avec les notations précédentes, ce qu'on doit calculer est :

$$P(F \cap A_2) = P(F|A_2)P(A_2) = \frac{1}{4}.e^{-2}\frac{4}{2} = \frac{1}{2e^2}.$$

b) D'après la formule des probabilités totales, on a :

$$P(F) = 1 - P(\overline{F}) = 1 - \sum_{n=0}^{+\infty} P(\overline{F}|A_n)P(A_n).$$

et on a :

$$P(\overline{F}|A_n) = \binom{n}{0}\frac{1}{2^0}\frac{1}{2^n} + \binom{n}{1}\frac{1}{2^1}\frac{1}{2^{n-1}} = (n+1)\frac{1}{2^n}.$$

Ainsi, on a :

$$\begin{aligned} P(\overline{F}) &= \sum_{n=0}^{+\infty} P(\overline{F}|A_n)P(A_n) \\ &= \sum_{n=0}^{+\infty}(n+1)\frac{1}{2^n}p_n \\ &= e^{-2}\sum_{n=0}^{+\infty}\frac{n+1}{n!} \\ &= e^{-2}\left(\sum_{n=1}^{+\infty}\frac{n}{n!} + \sum_{n=0}^{+\infty}\frac{1}{n!}\right) \\ &= 2e^{-2}\sum_{n=0}^{+\infty}\frac{1}{n!} \\ &= \frac{2}{e}. \end{aligned}$$

D'où $P(F) = 1 - \frac{2}{e}$.

c) On a :

$$P(A_2|F) = \frac{P(A_2 \cap F)}{P(F)} = \frac{1}{2e^2}.\frac{1}{1-\frac{2}{e}} = \frac{1}{2e(e-2)}.$$

3. Événements indépendants

a. Définition et premières propriétés

Définition 16. *Événements indépendants*

Soit $A, B \in \mathcal{T}$. On dit que A et B sont **indépendants** si

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

On justifie le choix de la terminologie d'"indépendance" par la proposition suivante :

Proposition 23.

Soit $A, B \in \mathcal{T}$ avec $P(B) > 0$. Les événements A et B sont indépendants si, et seulement si, $P(A|B) = P(A)$.

Proposition 24.

Soit $A, B \in \mathcal{T}$. Si A et B sont indépendants, alors
- $\overline{A}$ et $\overline{B}$ sont indépendants ;
- $\overline{A}$ et B sont indépendants (et A et $\overline{B}$ également).

b. Indépendance d'une famille d'événement

Définition 17. *Indépendance deux à deux/Indépendance*

Soit $(A_i)_{i\in I}$ une famille à valeurs dans $\mathcal{T}$.

- On dit que les événements A_i, pour $i \in I$ sont **deux à deux indépendants** si, pour tous $i, j \in \mathbb{I}$ avec $i \neq j$, on a :

$$P(A_i \cap A_j) = P(A_i)P(A_j).$$

- On dit que les événements A_i, pour $i \in I$ sont **indépendants** (ou *mutuellement indépendants*) si, pour toute partie finie J de I, on a :

$$P\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} P(A_i).$$

Remarque 9.

Il est clair que l'indépendance d'une famille d'événements implique l'indépendance deux à deux de ces événements mais **la réciproque est fausse !**

En effet, si on considère deux lancers de dé et les trois événements suivants sont deux à deux indépendants, mais pas indépendants !
- A_1="La somme des deux lancers est pair" ;
- A_2="Le premier lancer est pair" ;
- A_3="Le second lancer est pair".

Partie E

Variables aléatoires discrètes

1. Variables aléatoires discrètes

a. Définition

Définition 18. *Variable aléatoire discrète*

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$. On dit que la fonction X est une **variable aléatoire discrète** sur $(\Omega, \mathcal{T})$ si :

i) l'ensemble $X(\Omega)$ est au plus dénombrable ;

ii) pour tout $x \in X(\Omega)$, $X^{-1}(\{x\})$ est un événement de $\mathcal{T}$, i.e.

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

De plus, si $E \subset \mathbb{R}$, on dit que X est une variable aléatoire discrète **réelle**.

Proposition 25.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$ une variable aléatoire discrète. Pour tout $A \subset E$, $X^{-1}(A) \in \mathcal{T}$.

Pour des questions de lisibilité et de simplicité, on emploie les notations suivantes :

Notation 1.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable, E un ensemble et $X : \Omega \to E$ une variable aléatoire discrète.

— Pour $A \subset E$, l'événement $X^{-1}(A)$ est noté $(X \in A)$.

— Pour $x \in E$, l'événement $X^{-1}(\{x\})$ est noté :

$$(X = x).$$

— Si $E \subset \mathbb{R}$, pour $x \in E$, les événements $X^{-1}(]-\infty, x[)$, $X^{-1}(]-\infty, x])$, $X^{-1}(]x, +\infty[)$ et $X^{-1}([x, +\infty[)$ sont notés respectivement :

$$(X < x), \qquad (X \leq x), \qquad (X > x), \qquad (X \geq x).$$

b. Loi de probabilité d'une variable aléatoire

Définition 19. *Loi de probabilité*

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. On appelle **loi de probabilité de** X l'application $P_X : \mathcal{P}(X(\Omega)) \to [0,1]$ définie par :

$$P_X : A \mapsto P(X \in A).$$

Proposition 26.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Le triplet $(X(\Omega), \mathcal{P}(X(\Omega)), P_X)$ est un espace probabilisé.

Proposition 27.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. Alors la famille $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements. En particulier, on a :

$$\sum_{x \in X(\Omega)} P(X = x) = 1.$$

Remarque 10.

— D'après ce qui précède, si X une variable aléatoire *discrète* sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, la loi de probabilité P_X est entièrement déterminée par les événements $(X = x)$ pour $x \in X(\Omega)$.

Ainsi, pour déterminer la loi de probabilité P_X de la variable aléatoire X, (il faut et) il suffit de déterminer les valeurs de $P(X = x)$ pour chaque $x \in X(\Omega)$.

— Si X est une variable aléatoire discrète à valeurs dans un ensemble au plus dénombrable E, $X(\Omega)$ n'est pas forcément égal à E. Dans la suite, on étendra la loi P_X de X à $\mathcal{P}(E)$ tout entier, en posant, pour $x \in E \setminus X(\Omega)$, $P_X(\{x\}) := 0$ ou encore $P(X = x) = 0$. Ainsi, P_X définit une loi de probabilité sur $(E, \mathcal{P}(E))$.

Définition 20. *Loi de probabilité discrète*

Soit E un ensemble au plus dénombrable et $(p_x)_{x \in E}$ une famille de réels positifs telle que $\sum_{x \in E} p_x = 1$. On appelle **loi (de probabilité) discrète sur** E la probabilité sur $(E, \mathcal{P}(E))$ définie par la famille $(p_x)_{x \in E}$. On notera souvent $\mathcal{L}$ une telle loi.

Notation 2.

Soit E un ensemble au plus dénombrable, $\mathcal{L}$ une loi discrète sur E définie par une famille $(p_x)_{x \in E}$ et X, Y des variables aléatoires discrètes à valeurs dans E.

— On dit que X **suit la loi discrète** $\mathcal{L}$ et on note $X \sim \mathcal{L}$ si $P_X = \mathcal{L}$ i.e. si pour tout

$x \in E$,

$$P(X = x) = p_x.$$

— On dit que X et Y **suivent la même loi** ou encore que X et Y sont **équidistribuées** et on note $X \sim Y$ si $P_X = P_Y$ i.e. si pour tout $x \in E$,

$$P(X = x) = P(Y = y).$$

c. Fonction d'une variable aléatoire discrète

Proposition-Notation 28.

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, F un ensemble et $f : X(\Omega) \to F$ une fonction. On note $f(X)$ la variable aléatoire discrète $f \circ X$.

Proposition 29.

Soit E, F des ensembles, X, Y des variables aléatoires discrètes à valeurs dans un ensemble E et $f : E \to F$ une fonction. Si $X \sim Y$, alors $f(X) \sim f(Y)$.

2. Lois usuelles sur un univers fini

a. Loi uniforme

Définition 21. *Loi uniforme*

Soit E un ensemble fini. On dit qu'une variable aléatoire X **suit une loi uniforme sur** E et on note $X \sim \mathcal{U}(E)$ si $X(\Omega) = E$ et, pour tout $x \in E$,

$$P(X = x) = \frac{1}{\#E}.$$

Si $E = [\![1, n]\!]$ pour un certain $n \in \mathbb{N}^*$, on notera simplement $\mathcal{U}(n)$ la loi uniforme sur $[\![1, n]\!]$.

b. Loi de Bernoulli

Définition 22. *Loi de Bernoulli*

Soit $p \in]0, 1[$. On dit qu'une variable aléatoire X **suit une loi de Bernoulli de paramètre** p et on note $X \sim \mathcal{B}(p)$ si $X(\Omega) = \{0, 1\}$ et

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

c. Loi binomiale

Définition 23. *Loi binomiale*

Soit $n \in \mathbb{N}^*$ et $p \in]0,1[$. On dit qu'une variable aléatoire X **suit une loi binomiale de paramètres** n **et** p et on note $X \sim \mathcal{B}(n,p)$ si $X(\Omega) = [\![0,n]\!]$ et, pour tout $k \in [\![0,n]\!]$,

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

3. Lois usuelles sur un univers dénombrable

a. Loi géométrique

Définition 24. *Loi géométrique*

Soit $p \in]0,1[$. On dit qu'une variable aléatoire X **suit une loi géométrique de paramètre** p et on note $X \sim \mathcal{G}(p)$ si $X(\Omega) = \mathbb{N}^*$ et, pour tout $k \in \mathbb{N}^*$,

$$P(X=k) = p(1-p)^{k-1}.$$

Remarque 11.

Une variable aléatoire suivant une loi géométrique de paramètre p modélise le temps d'attente du premier succès d'une suite d'épreuves de Bernoulli indépendantes de paramètre p.

Exercice 14.

Soit $p \in]0,1[$ et X une variable aléatoire qui suit une loi géométrique de paramètre p. Déterminer la valeur de $P(X > k)$ pour $k \in \mathbb{N}^*$.

Correction.

On a $X \sim \mathcal{G}(p)$ donc $X(\Omega) = \mathbb{N}^*$ et pour tout $k \in \mathbb{N}^*$,

$$P(X=k) = p(1-p)^{k-1}.$$

On remarque que :

$$(X > k) = (X = k+1) \cup (X = k+2) \cup \ldots = \bigcup_{n=k+1}^{+\infty} (X = n)$$

Les évènements $(X = n)$ étant incompatibles, on a alors :

$$P(X > k) = \sum_{n=k+1}^{+\infty} P(X=n) = \sum_{n=k+1}^{+\infty} p(1-p)^{n-1} = p(1-p)^k \underbrace{\sum_{n=0}^{+\infty} (1-p)^n}_{=\frac{1}{p}} = (1-p)^k$$

Proposition 30.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ et que $P(X = n) \neq 0$ pour tout $n \in \mathbb{N}^*$. Alors X suit une loi géométrique si, et seulement si, pour tous $k, l \in \mathbb{N}^*$,

$$P(X > k + l | X > k) = P(X > l).$$

Démonstration.

On remarque tout d'abord que, pour tous $k, l \in \mathbb{N}^*$, $(X > k+l) \subset (X > k)$ d'où $(X > k+l) \cap (X > k) = (X > k + l)$, et ainsi :

$$P(X > k + l | X > k) = \frac{P((X > k + l) \cap (X > k))}{P(X > k)} = \frac{P(X > k + l)}{P(X > k)}.$$

($\Rightarrow$) On suppose qu'il existe $p \in]0, 1[$ tel que $X \sim \mathcal{G}(p)$. Alors, d'après l'exercice précedent, pour tous $k, l \in \mathbb{N}^*$:

$$P(X > k + l | X > k) = \frac{P(X > k + l)}{P(X > k)} = \frac{(1 - p)^{k+l}}{(1 - p)^k} = (1 - p)^l = P(X > l).$$

($\Leftarrow$) On suppose que pour tous $k, l \in \mathbb{N}^*$,

$$\left(\frac{P(X > k + l)}{P(X > k)} =\right) P(X > k + l | X > k) = P(X > l).$$

On pose $p = 1 - P(X > 1) \in]0, 1[$ (car les $P(X = n)$ sont non nuls) et pour $n \in \mathbb{N}^*$, $u_n = P(X > n) > 0$. Alors on a :

$$\frac{u_{n+1}}{u_n} = \frac{P(X > n + 1)}{P(X > n)} \underbrace{=}_{k=n \text{ et } l=1} P(X > 1) = 1 - p$$

Ainsi, $(u_n)_{n \in \mathbb{N}^*}$ est une suite géométrique de raison $1 - p$ et donc, pour tout $n \in \mathbb{N}^*$:

$$P(X > n) = u_n = (1 - p)^{n-1} u_1 = (1 - p)^{n-1} P(X > 1) = (1 - p)^n$$

On remarque que la formule précédente est toujours valable pour $n = 0$ car X étant à valeurs dans $\mathbb{N}^*$, $P(X > 0) = 1$.

Par suite, comme, pour $k \in \mathbb{N}^*$,

$$(X = k) = (X \leq k) \cap (X > k - 1) = \overline{(X > k)} \cap (X > k - 1) = (X > k - 1) \setminus (X > k)$$

et $(X > k) \subset (X > k - 1)$, on a :

$$\begin{aligned} P(X = k) &= P(X > k - 1) - P(X > k) \\ &= (1 - p)^{k-1} - (1 - p)^k \\ &= (1 - p)^{k-1}(1 - (1 - p)) \\ P(X = k) &= p(1 - p)^{k-1} \end{aligned}$$

Il en résulte que $X \sim \mathcal{G}(p)$.

□

b. Loi de Poisson

Définition 25. *Loi de Poisson*

Soit $\lambda \in \mathbb{R}_+^*$. On dit qu'une variable aléatoire X **suit une loi de Poisson de paramètre** λ et on note $X \sim \mathcal{P}(\lambda)$ si $X(\Omega) = \mathbb{N}$ et, pour tout $k \in \mathbb{N}$,

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

Exercice 15.

Soit $\lambda \in \mathbb{R}_+^*$, $(p_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $]0,1[$ et $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour chaque $n \in \mathbb{N}$, $X_n \sim \mathcal{B}(n, p_n)$.
Montrer que si $np_n \xrightarrow[n\to+\infty]{} \lambda$, alors, pour tout $k \in \mathbb{N}$,

$$P(X_n = k) \xrightarrow[n\to+\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Remarque 12.

Pour $n \in \mathbb{N}$ et $p \in]0,1[$ tels que $\lambda = np$ est proche de 0, une variable aléatoire qui suit une loi binomiale de paramètres n et p, suit approximativement une loi de Poisson de paramètre $\lambda = np$.

4. Couples de variables aléatoires

Dans ce paragraphe, $(\Omega, \mathcal{T}, P)$ désigne un espace probabilisé. Les variables aléatoires considérées sont définies sur cet espace.

a. Généralités

Définition 26. *Couple de variable aléatoires discrètes*

Soit X, Y des variables aléatoires discrètes sur un espace probabilisable. On appelle **couple de variables aléatoires discrètes sur** $(\Omega, \mathcal{T}, P)$ et on note (X, Y) l'application $\omega \mapsto (X(\omega), Y(\omega))$.
De plus, pour $(A, B) \subset X(\Omega) \times Y(\Omega)$, on note $(X \in A, Y \in B)$ l'évènement $((X, Y) \in A \times B) = (X \in A) \cap (Y \in B)$.

Remarque 13.

En particulier, pour $(x, y) \in X(\Omega) \times Y(\Omega)$, la notation est :

$$(X = x, Y = y) = ((X, Y) = (x, y)) = (X = x) \cap (Y = y).$$

Proposition 31.

Soit X, Y des variables aléatoires discrètes à valeurs dans E, F respectivement. Alors le couple de variables aléatoires (X, Y) est une variable aléatoire discrète à valeurs dans $E \times F$.

Proposition 32.

Soit (X, Y) un couple de variables aléatoires discrètes. Alors la famille d'événements :

$$((X = x) \cap (Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$$

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

b. Loi conjointe

Définition 27. *Loi conjointe*

Soit (X, Y) un couple de variables aléatoires discrètes. On appelle **loi conjointe de** (X, Y) et on note $P_{(X,Y)}$ la loi de probabilité du couple (X, Y).

Proposition 33.

Soit (X, Y) un couple de variables aléatoires discrètes. La loi $P_{(X,Y)}$ est déterminée par la donnée des $P(X = x, Y = y)$ pour chaque couple $(x, y) \in X(\Omega) \times Y(\Omega)$.

c. Lois marginales

Définition 28. *Loi marginales*

Soit (X, Y) un couple de variables aléatoires discrètes. On appelle **première loi marginale du couple** la loi de probabilité de X et **deuxième loi marginale du couple** la loi de probabilité de Y.

Théorème 13.

Soit (X, Y) un couple de variables aléatoires discrètes. On a :

- pour tout $x \in X(\Omega)$, $P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y)$.
- pour tout $y \in Y(\Omega)$, $P(Y = y) = \sum_{x \in X(\Omega)} P(X = x, Y = y)$.

d. Lois conditionnelles

Définition 29. *loi conditionnelle*

Soit (X, Y) un couple de variables aléatoires discrètes et $x \in X(\Omega)$ tel que $P(X = x) \neq 0$. On appelle **loi conditionnelle de** Y **sachant** $X = x$ la probabilité notée $P_{(X=x)}$ sur $(\Omega, \mathcal{T})$ déterminée, pour $y \in Y(\Omega)$, par :

$$P_{(X=x)}(Y = y) = P(Y = y|X = x).$$

Exercice 16.

Soit (X, Y) un couple de variables aléatoires discrètes tel que, pour tout $x \in X(\Omega)$, $P(X = x) \neq 0$.

1. Soit $(x, y) \in X(\Omega) \times Y(\Omega)$. Montrer que

$$P((X = x) \cap (Y = y)) = P(X = x)P(Y = y|X = x).$$

2. Soit $y \in Y(\Omega)$. Montrer que

$$P(Y = y) = \sum_{x \in X(\Omega)} P(X = x)P(Y = y|X = x).$$

Remarque 14.

On généralise toutes les notions précédentes aux cas d'un n-uplet de variables aléatoires.

5. Variables aléatoires indépendantes

a. Couple de variables aléatoires indépendantes

Définition 30. *Couple de variables aléatoires indépendantes*

Soit (X, Y) un couple de variables aléatoires discrètes. On dit que X et Y sont **indépendantes** et on note $X \perp\!\!\!\perp Y$ si, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$,

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

Proposition 34.

Soit (X, Y) un couple de variables aléatoires discrètes. Les variables X et Y sont indépendantes si, et seulement si, pour tout $(A, B) \subset X(\Omega) \times Y(\Omega)$,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Proposition 35.

Soit (X, Y) un couple de variables aléatoires discrètes et f, g des fonctions définies sur $X(\Omega)$ et $Y(\Omega)$ respectivement. Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.

Proposition 36. *Loi d'une somme de variables aléatoires discrètes*

Soit (X, Y) un couple de variables aléatoires discrètes à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors, pour tout $n \in \mathbb{N}$,

$$P(X + Y = n) = \sum_{k=0}^{n} P(X = k)P(Y = n - k).$$

Exercice 17.

Soit $\lambda, \mu \in \mathbb{R}_+^*$ et (X, Y) un couple de variables aléatoires discrètes tel que $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$. Montrer que si X et Y sont indépendantes, alors $X + Y \sim \mathcal{P}(\lambda + \mu)$.

b. Famille finie de variables aléatoires indépendantes

Définition 31. *Variables aléatoires indépendantes*

Soit $n \in \mathbb{N}^*$ et $(X_k)_{k \in [\![1,n]\!]}$ une famille de variables aléatoires discrètes. On dit que les variables $X_1, \ldots, X_n$ sont **indépendantes** (ou mutuellement indépendantes) si, pour tout $(x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega)$:

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{k=1}^{n} P(X_k = x_k).$$

Proposition 37.

Soit $n \in \mathbb{N}^*$ et $(X_k)_{k \in [\![1,n]\!]}$ une famille de variables aléatoires discrètes. Les variables $X_1, \ldots, X_n$ sont indépendantes si, et seulement si, pour tout $(A_1, \ldots A_n) \subset X_1(\Omega) \times \ldots \times X_n(\Omega)$,

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = \prod_{k=1}^{n} P(X_k \in A_k).$$

Proposition 38. *Lemme des coalitions*

Soit $n, m \in \mathbb{N}^*$ avec $m < n$, $(X_k)_{k \in [\![1,n]\!]}$ une famille de variables aléatoires discrètes et f, g des fonctions définies sur $X_1(\Omega) \times \ldots \times X_m(\Omega)$ et $X_{m+1}(\Omega) \times \ldots \times X_n(\Omega)$ respectivement. Si $X_1, \ldots, X_n$ sont indépendantes, alors $f(X_1, \ldots, X_m)$ et $g(X_{m+1}, \ldots, X_n)$ sont indépendantes.

Exercice 18.

Soit $m \in \mathbb{N}^*$, $n_1, ..., n_m \in \mathbb{N}$, $p \in [0,1]$ et $X_1, ..., X_m$ des variables aléatoires telles que, pour $k \in [\![1, m]\!]$, $X_k \sim \mathcal{B}(n_k, p)$.
On suppose $X_1, ..., X_n$ indépendantes. Déterminer la loi de $Y = X_1 + \ldots + X_n$.

c. Famille infinie de variables aléatoires indépendantes

Définition 32.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires discrètes. On dit que les variables X_n sont **indépendantes** si, pour toute partie finie $I = \{i_1, ..., i_k\}$ de $\mathbb{N}$, les variables $X_{i_1}, \ldots, X_{i_k}$ sont indépendantes.

Le théorème suivant (dont la démonstration est admise) va nous permettre de pouvoir "parler sans crainte" de suites de variables aléatoires indépendantes suivant chacune une loi prescrite :

Théorème 14. *Théorème d'existence de Kolmogorov*

Si $(\mathcal{L}_n)_{n\in\mathbb{N}}$ est une suite de lois discrètes, alors il existe un espace probabilisé et une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires discrètes indépendantes sur cet espace telle que, pour tout $n \in \mathbb{N}$, $X_n \sim \mathcal{L}_n$.

Exemple 8.

Soit $p \in [0,1]$ et $(X_n)_{n\in\mathbb{N}^*}$ une suites de variables aléatoires indépendantes suivant chacune une loi de Bernouilli de paramètre p - *le théorème précédent (et il n'est là que pour cela) garantit que cette suite est bien définie !*
La variable $T = \min\{n \in \mathbb{N}^* \cup \{\infty\} \mid X_n = 1\}$ suit une loi géométrique de paramètre p.

Partie F

Espérance, variance

Sauf mention contraire, les variables aléatoires discrètes considérées dans cette partie sont définies sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ et sont à valeurs réelles.

1. Espérance

a. Définitions et exemples

Définition 33. *Espérance*

Soit X une variable aléatoire discrète .

- Si X est à valeurs dans $\mathbb{R}_+$, on appelle **espérance de** X et on note $E(X)$ la quantité (potentiellement infinie) :
$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$
- Si X est à valeurs dans $\mathbb{R}$, on dit que X **admet une espérance finie** et on note $X \in L^1$, si la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas, on appelle **espérance de** X et on note $E(X)$ la quantité :
$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

Exemple 9.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $X \in L^1$ et $E(X) = \dfrac{1}{p}$,

Si $X \sim \mathcal{G}(p)$, alors pour tout $k \in \mathbb{N}^*$, $P(X = k) = p(1-p)^{k-1}$. D'après le critère de D'Alembert, $u_k = kp(1-p)^{k-1}$ est le terme général d'une série convergente car $0 < 1 - p < 1$. Par suite, $X \in L^1$ et

$$E(X) = \sum_{k=1}^{+\infty} kp(1-p)^{k-1}.$$

Calculons la valeur de cette somme. La série entière $\sum_{k\geq 0} x^k$ a pour rayon de convergence 1, donc sa somme $f : x \mapsto \frac{1}{1-x}$ est C^1 sur $]-1, 1[$ et on a, pour tout $x \in]-1, 1[$:

$$\begin{aligned}
\frac{1}{(1-x)^2} = f'(x) &= \frac{d}{dx} \sum_{k=0}^{+\infty} x^k \\
&= \sum_{k=0}^{+\infty} \frac{d}{dx} x^k \\
&= \sum_{k=0}^{+\infty} kx^{k-1}
\end{aligned}$$

Par suite, on a, comme $0 < 1 - p < 1$:

$$\begin{aligned}
E(X) &= \sum_{k=1}^{+\infty} kp(1-p)^{k-1} \\
&= p \sum_{k=0}^{+\infty} k(1-p)^{k-1} \\
&= pf'(1-p) = \frac{p}{(1-(1-p))^2} \\
E(X) &= \frac{1}{p}
\end{aligned}$$

— Si $X \sim \mathcal{P}(\lambda)$ alors $X \in L^1$ et $E(X) = \lambda$.

Si $X \sim \mathcal{P}(\lambda)$, alors pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$. D'après le critère de D'Alembert, $u_k = k\dfrac{\lambda^k}{k!}$ est le terme général d'une série convergente. Par suite, $X \in L^1$ et

$$E(X) = \sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!}.$$

Calculons la valeur de cette somme. On a :

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

Exercice 19.

1. Soit $c \in \mathbb{R}$ et X une variable aléatoire discrète presque sûrement égale à c. Calculer l'espérance de X.
2. Déterminer les espérances des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Exercice 20.

Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ telle que, pour $n \in \mathbb{N}^*$, $P(X = n) = \dfrac{\alpha}{n^2}$ pour un certain $\alpha > 0$.

1. Déterminer α.
2. Quelle est l'espérance de X ?

b. Propriétés de l'espérance

Théorème 15. *Théorème de transfert*

Soit X une variable aléatoire discrète et $f : X(\Omega) \to \mathbb{R}$ une fonction. La variable aléatoire $f(X)$ est d'espérance finie si, et seulement si, la famille $(f(x)P(X = x))_{x\in X(\Omega)}$ est sommable. Dans

ce cas, on a :

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$

Démonstration.

On considère la variable Y la variable aléatoire discrète $f(X)$ et on pose, pour $y \in Y(\Omega)$:

$$I_y = f^{-1}(\{y\}) = \{x \in X(\Omega) \mid f(x) = y\}.$$

On remarque qu'on a, pour tout $y \in Y(\Omega)$, $(X \in I_y) = (Y = y)$ par définition de I_y.
La famille $(I_y)_{y \in Y(\Omega)}$ forme une partition de $X(\Omega)$ et d'après le théorème de sommation par paquets appliqué à la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ et à la partition précédente, on a :
la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable si, et seulement si,

— Somme sur chaque paquet. Pour tout $y \in Y(\Omega)$, $(f(x)P(X = x))_{x \in I_y} = (yP(X = x))_{x \in I_y}$ est sommable. Or c'est le cas ici car $(yP(X = x))_{x \in I_y}$ est une sous-famille de la famille $(yP(X = x))_{x \in X(\Omega)}$ qui est sommable car $(P(X = x))_{x \in X(\Omega)}$ l'est. De plus, on a, en notant s_y la somme de la famille $(f(x)P(X = x))_{x \in I_y}$:

$$s_y = \sum_{x \in I_y} f(x)P(X = x) = y \sum_{x \in I_y} P(X = x) = yP(X \in I_y) = yP(Y = y).$$

— Sommes des sommes des paquets. La famille $(s_y)_{y \in Y(\Omega)} = (yP(Y = y))_{y \in Y(\Omega)}$ est sommable. Or, c'est le cas si, et seulement si, Y est d'espérance finie.

Il en résulte que la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable si, et seulement si, $Y = f(X)$ est d'espérance finie. De plus, dans ce cas, on a, toujours le théorème de sommation par paquets :

$$E(f(X)) = E(Y) = \sum_{y \in Y(\Omega)} \underbrace{P(Y = y)}_{= s_y} = \sum_{y \in Y(\Omega)} \sum_{x \in I_y} f(x)P(X = x) = \sum_{x \in X(\Omega)} f(x)P(X = x).$$

□

Théorème 16. *Linéarité de l'espérance*

Soit X, Y des variables aléatoires discrètes et $\lambda, \mu \in \mathbb{R}$. Si $X, Y \in L^1$, alors la variable aléatoire discrète $\lambda X + \mu Y \in L^1$ et on a :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Démonstration.

On considère la fonction $f : (x, y) \mapsto \lambda x + \mu y$.
Montrons que la famille $(f(x, y)P((X, Y) = (x, y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable.
On remarque que, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a :

$$f(x, y)P(X = x, Y = y) = \lambda x P(X = x, Y = y) + \mu y P(X = x, Y = y)$$

On étudie alors séparemment les familles $(xP(X=x,Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ et $(yP(X=x,Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$.

- On applique le théorème de sommation par paquets à la famille $(xP(X=x,Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ sur la partition verticale $(V_x)_{x\in X(\Omega)}$ où $V_x=\{(x,y)\mid y\in Y(\Omega)\}$ pour $x\in X(\Omega)$.
  - — Somme sur chaque paquet. Soit $x\in X(\Omega)$. Pour tout $y\in Y(\Omega)$, $P(X=x,Y=y)\leq P(Y=y)$ car $(X=x,Y=y)\subset(Y=y)$, donc :

    $$P(X=x,Y=y)\leq P(Y=y).$$

    Or, la famille $(P(Y=y))_{y\in Y(\Omega)}$ est sommable donc, par comparaison, $(P(X=x,Y=y))_{y\in Y(\Omega)}$ est sommable.
    Par suite, $(xP(X=x,Y=y))_{y\in Y(\Omega)}$ est sommable et on a, en notant s_x la somme de cette famille, d'après la formule des probabilités totales :

    $$\begin{aligned} s_x &= \sum_{y\in Y(\Omega)} xP(X=x,Y=y) \\ &= x\sum_{y\in Y(\Omega)} P(X=x,Y=y) \\ s_x &= xP(X=x) \end{aligned}$$

  - — Sommes des sommes des paquets. La famille $(s_x)_{x\in X(\Omega)}=(xP(X=x)_{x\in X(\Omega)}$ est sommable car $X\in L^1$.
    Par suite, $(xP(X=x,Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est une famille sommable et on a :

$$\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xP(X=x,Y=y)=\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} xP(X=x,Y=y)=\sum_{x\in X(\Omega)} xP(X=x)=E(X).$$

- On applique le théorème de sommation par paquets à la famille $(yP(X=x,Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ sur la partition horizontale $(H_y)_{y\in Y(\Omega)}$ où $H_y=\{(x,y)\mid x\in X(\Omega)\}$ pour $y\in Y(\Omega)$.
  - — Somme sur chaque paquet. Soit $y\in Y(\Omega)$. Pour tout $x\in X(\Omega)$, $P(X=x,Y=y)\leq P(X=x)$ car $(X=x,Y=y)\subset(X=x)$, donc :

    $$P(X=x,Y=y)\leq P(X=x).$$

    Or, la famille $(P(X=x))_{x\in X(\Omega)}$ est sommable donc, par comparaison, $(P(X=x,Y=y))_{x\in X(\Omega)}$ est sommable.
    Par suite, $(yP(X=x,Y=y))_{x\in X(\Omega)}$ est sommable et on a, en notant s_y la somme de cette famille, d'après la formule des probabilités totales :

    $$\begin{aligned} s_y &= \sum_{x\in X(\Omega)} yP(X=x,Y=y) \\ &= y\sum_{x\in X(\Omega)} P(X=x,Y=y) \\ s_y &= yP(Y=y) \end{aligned}$$

  - — Sommes des sommes des paquets. La famille $(s_y)_{y\in Y(\Omega)}=(yP(Y=y)_{y\in Y(\Omega)}$ est sommable car $Y\in L^1$.

Par suite, $(yP(X=x,Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est une famille sommable et on a :

$$\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} yP(X=x,Y=y) = \sum_{y\in Y(\Omega)}\sum_{x\in X(\Omega)} yP(X=x,Y=y) = \sum_{y\in Y(\Omega)} yP(Y=y) = E(Y).$$

Il en résulte que $(f(x,y)P((X,Y)=(x,y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est une famille sommable comme combinaison linéaire de famille sommable. Ainsi, d'après le théorème de transfert, $\lambda X+\mu Y = f(X,Y)$ est d'espérance finie et, par linéarité de la somme, on obtient :

$$\begin{aligned}
E(\lambda X+\mu Y) &= E(f(x,y)) \\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} f(x,y)P(X=x,Y=y) \\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} (\lambda x+\mu y)P(X=x,Y=y) \\
&= \lambda \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xP(X=x,Y=y) + \mu \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} yP(X=x,Y=y) \\
E(\lambda X+\mu Y) &= \lambda E(X)+\mu E(Y).
\end{aligned}$$

D'où la linéarité de l'espérance. □

Proposition 39. *Positivité/Croissance*

Soit X, Y des variables aléatoires discrètes.

- Si X est à valeurs positives, alors $E(X)\geq 0$ et $E(X)=0$ implique $X=0$ presque sûrement.
- Si $X,Y\in L^1$ et $X\leq Y$, alors $E(X)\leq E(Y)$.

Démonstration.

- On suppose X à valeurs positives. Alors on a :

$$E(X) = \sum_{x\in X(\Omega)} \underbrace{x}_{\geq 0}\underbrace{P(X=x)}_{x\geq 0} \geq 0$$

($E(X)$ étant potentiellement $+\infty$ ici).
On suppose de plus $X\in L^1$ et $E(X)=0$. Alors, on a, pour tout $x\in X(\Omega)$, comme X est à valeurs positives :

$$0\leq xP(X=x)\leq \sum_{x'\in X(\Omega)} x'P(X=x') = E(X) = 0,$$

donc $xP(X=x)=0$. Ainsi, pour tout $x\in X(\Omega)\setminus\{0\}$, $P(X=x)=0$, d'où :

$$P(X=0) = 1-\sum_{x\in X(\Omega)\setminus\{0\}} P(X=x) = 1.$$

D'où $X=0$ presque sûrement.

— On suppose $X \leq Y$ et $X, Y \in L^1$. Alors $Z = Y - X \in L^1$ est positive, donc, d'après le point précédent et par linéarité de l'espérance :

$$E(Y) - E(X) = E(Z) \geq 0,$$

d'où $E(X) \leq E(Y)$.

□

Proposition 40.

Soit X une variable aléatoire discrète. Alors $X \in L^1$ si, et seulement si, $|X| \in L^1$. Dans ce cas on a $|E(X)| \leq E(|X|)$.

Théorème 17. *Espérance d'un produit de variables indépendantes*

Soit X, Y des variables aléatoires discrètes. Si X et Y sont **indépendantes** et d'espérance finie, alors la variable aléatoire XY est d'espérance finie et on a :

$$E(XY) = E(X)E(Y).$$

Théorème 18. *Inégalité de Markov*

Soit X une variable aléatoire discrète à valeurs positives. Si $X \in L^1$, alors, pour tout $a > 0$, on a :

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

Démonstration.

On suppose $X \in L^1$. Soit $a > 0$. On pose $Y = \mathbb{1}_{(X \geq a)}$ i.e. pour $\omega \in \Omega$,

$$Y(\omega) = \begin{cases} 1 & \text{si } X(\omega) \geq a \\ 0 & \text{si } X(\omega) < a \end{cases}$$

Alors Y est une variable aléatoire à valeurs dans $\{0, 1\}$ et donc d'espérance finie ; et on a :

$$E(X) = 0 \times P(Y = 0) + 1 \times P(Y = 1) = P(Y = 1) = P(X \geq a).$$

De plus, on remarque que $X \geq aY$; en effet, pour tout $\omega \in \Omega$:
— si $X(\omega) \geq a$, alors $Y(\omega) = 1$, et donc $X(\omega) \geq aY(\omega)$;
— si $X(\omega) < a$, alors $Y(\omega) = 0$ et comme X est à valeurs positives, $X(\omega) \geq 0 = aY(\omega)$.
Par croissance de l'espérance, on a donc $E(aY) \leq E(X)$ et donc

$$aP(X \geq a) = aE(Y) = E(aY) \leq E(X)$$

D'où le résultat en divisant par $a > 0$. □

2. Variance

a. Moments

Définition 34. *Moments*

Soit X une variable aléatoire discrète et $r \in \mathbb{N}$. On dit que X **admet un moment d'ordre** r et on note :

$$X \in L^r.$$

si la variable aléatoire discrète X^r admet une espérance finie. Dans ce cas, on appelle **moment d'ordre** r la quantité $E(X^r)$.

0n-1

Proposition 41.

Soit X une variable aléatoire discrète. Si $X \in L^2$, alors $X \in L^1$.

Théorème 19. *Inégalité de Cauchy-Schwarz*

Soit X, Y des variables aléatoires discrètes. Si $X, Y \in L^2$ alors $XY \in L^1$ et on a :

$$E(XY) \leq \sqrt{E(X^2)}.\sqrt{E(Y^2)}.$$

b. Variance, écart-type

Définition 35. *Variance*

Soit X une variable aléatoire discrète telle que $X \in L^2$. On appelle **variance de** X et on note $V(X)$ la quantité :

$$V(X) = E\left((X - E(X))^2\right).$$

Théorème 20. *Formule de Kœnig-Huygens*

Soit X une variable aléatoire discrète. Si $X \in L^2$, alors on a :

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

Exemple 10.

Soit X une variable aléatoire discrète.

- Si $X \sim \mathcal{G}(p)$ alors $X \in L^2$ et $V(X) = \dfrac{1-p}{p^2}$,
- Si $X \sim \mathcal{P}(\lambda)$ alors $X \in L^2$ et $V(X) = \lambda$.

On suppose que $X \sim \mathcal{P}(\lambda)$. Alors

- $X(\Omega) = \mathbb{N}$
- pour tout $k \in \mathbb{N}$, $P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$.

$X \in L^2$ si X^2 est d'espérance finie, ce qui est équivalent, d'après le théorème de transfert, à : la famille $(k^2 P(X = k))_{k\in\mathbb{N}}$ est sommable. Comme cette famille est une suite à termes positifs, il suffit alors de montrer que c'est le terme général d'une série convergente.
Or, pour tout $k \geq 1$, $u_k = k^2 P(X = k) = k^2 e^{-\lambda}\frac{\lambda^k}{k!} > 0$ et on a :

$$\frac{u_{k+1}}{u_k} = \frac{\lambda(k+1)}{k^2} \xrightarrow[k\to+\infty]{} 0 < 1$$

donc, d'après la régle de D'Alembert, $\sum u_k$ converge.
Par suite, $(k^2 P(X = k))_{k\in\mathbb{N}}$ est sommable et donc $X \in L^2$.
Ainsi, X admet une variance finie et également une espérance finie. Calculons tout d'abord $E(X(X-1))$ (la variable $X(X-1)$ est bien d'espérance finie car X^2 et X sont d'espérance finie et on a $X(X-1) = X^2 - X$) :

$$\begin{aligned}
E(X(X-1)) &= \sum_{k=0}^{+\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} \\
&= \lambda^2 e^{-\lambda}\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!} \\
&= \lambda^2 e^{-\lambda}\underbrace{\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!}}_{=e^{\lambda}} \\
E(X(X-1)) &= \lambda^2.
\end{aligned}$$

On a, par linéarité : $E(X(X-1)) = E(X^2) - E(X)$ et d'après le théorème de Kœnig-Huygens, $V(X) = E(X^2) - E(X)^2$; par suite :

$$V(X) = E(X(X-1)) + E(X) - E(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Exercice 21.

Déterminer les variances des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Proposition 42.

Soit $\lambda, \mu \in \mathbb{R}$ et X une variable aléatoire discrète. Si $X \in L^2$, alors on a :

$$V(\lambda X + \mu) = \lambda^2 V(X).$$

Définition 36. *Écart-type*

Soit X une variable aléatoire discrète telle que $X \in L^2$. On appelle **écart-type de** X et on note $\sigma(X)$ la quantité :

$$\sigma(X) = \sqrt{V(X)}.$$

Définition 37. *Variable centrée réduite*

Soit X une variable aléatoire discrète telle que $X \in L^2$. On dit que la variable aléatoire X est **centrée réduite** si $E(X) = 0$ et $\sigma(X) = 1$.

Proposition 43.

Soit X une variable aléatoire discrète telle que $X \in L^2$ et $V(X) \neq 0$. Alors la variable aléatoire $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est une variable aléatoire centrée réduite.

Théorème 21. *Inégalité de Bienaymé-Tchebychev*

Soit X une variable aléatoire discrète. Si $X \in L^2$, alors on a, pour tout $\varepsilon > 0$:

$$P\left(|X - E(X)| \geq \varepsilon\right) \leq \frac{V(X)}{\varepsilon^2}.$$

3. Covariance

a. Généralités

Définition 38. *Covariance*

Soit (X, Y) un couple de variables aléatoires discrètes tel que $X, Y \in L^2$. On appelle **covariance du couple** (X, Y) et on note $\mathrm{Cov}(X, Y)$ la quantité :

$$\mathrm{Cov}(X, Y) = E\left((X - E(X))(Y - E(Y))\right).$$

Théorème 22. *Formule de Kœnig-Huygens*

Soit (X, Y) un couple de variables aléatoires discrètes. Si $X, Y \in L^2$, alors on a :

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Proposition 44.

Soit (X, Y) un couple de variables aléatoires discrètes tel que $X, Y \in L^2$. Si X et Y sont indépendantes, alors $\mathrm{Cov}(X, Y) = 0$.

b. Variance d'une somme

Théorème 23.

Soit (X, Y) deux variables aléatoires discrètes. Si $X, Y \in L^2$, alors :

$$V(X+Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y).$$

Soit $n \in \mathbb{N}^*$. On a, plus généralement, si $(X_1, ..., X_n)$ est un n-uplet de variables aléatoires discrètes, on a :

$$V(X_1 + \ldots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leq i < j \leq n} \mathrm{Cov}(X_i, X_j).$$

Corollaire 5.

Soit (X, Y) deux variables aléatoires discrètes telles que $X, Y \in L^2$. Si X et Y sont indépendantes, on a :

$$V(X+Y) = V(X) + V(Y).$$

4. Loi faible des grands nombres

Théorème 24. *Loi faible des grands nombres*

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires discrètes, deux à deux indépendantes, de même loi et admettant un moment d'ordre 2. On pose $m = E(X_1)$ et, pour $n \in \mathbb{N}^*$, $S_n = X_1 + \ldots + X_n$. Alors on a, pour tout $\varepsilon > 0$:

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow[n \to +\infty]{} 0.$$

5. Fonctions génératrices

a. Définition et premières propriétés

Définition 39. *Fonction génératrice*

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. On appelle **fonction génératrice de**

X et on note G_X la fonction :

$$G_X : t \mapsto E(t^X).$$

Théorème 25.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. La fonction génératrice G_X de X est définie sur $[-1,1]$ et on a, pour tout $t \in [-1,1]$:

$$G_X(t) = \sum_{n=0}^{+\infty} t^n P(X=n).$$

Démonstration.

Pour $t \in [-1,1]$, on considère la famille $(t^n P(X=n))_{n\in\mathbb{N}}$. On remarque que pour tout $t \in [-1,1]$ et tout $n \in \mathbb{N}$

$$|t^n P(X=n)| = |t|^n P(X=n) \leq P(X=n)$$

Or $P(X=n)$ est le terme général d'une série convergente (dont la somme vaut 1), donc par comparaison, $\sum t^n P(X=n)$ converge absolument.
Par suite, $(t^n P(X=n))_{n\in\mathbb{N}}$ est une famille sommable, et ainsi, d'après le théorème de transfert, la variable aléatoire t^X est d'espérance finie. Alors, pour $t \in [-1,1]$ $G_X(t) = E(t^X)$ est bien défini et on a :

$$G_X(t) = E(t^X) = \sum_{n=0}^{+\infty} t^n P(X=n)$$

□

Proposition 45.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$. La fonction génératrice de X est de classe C^∞ sur $]-1,1[$ et on a, pour tout $k \in \mathbb{N}$,

$$G_X^{(k)}(0) = k!P(X=k).$$

Corollaire 6.

Soit X, Y des variables aléatoires discrètes à valeurs dans $\mathbb{N}$.

- La loi de probabilité P_X de X est entièrement déterminée par sa fonction génératrice G_X.
- les variables X et Y ont la même loi si, et seulement si, $G_X = G_Y$.

Proposition 46.

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice. On a :

— La variable X appartient à L^1 si, et seulement si, G est dérivable en 1. Dans ce cas, on a :

$$G'_X(1) = E(X).$$

— La variable X appartient à L^2 si, et seulement si, G est deux fois dérivable en 1. Dans ce cas, on a :

$$G''_X(1) = E(X(X-1)).$$

Exercice 22.

Soit $X \in L^2$ une variable aléatoire discrète à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice. Exprimer $E(X^2)$ et $V(X)$ en fonction des dérivées de G_X en 1.

Exemple 11.

Soit X une variable aléatoire discrète.

— Si $X \sim \mathcal{G}(p)$ alors $G_X(t) = \dfrac{pt}{1-(1-p)t}$,

On suppose $X \sim \mathcal{G}(p)$ avec $p \in]0,1[$. Alors :

$$\begin{cases} X(\Omega) = \mathbb{N}^* \\ P(X = k) = p(1-p)^{k-1} \text{ pour tout } k \in \mathbb{N}^* \end{cases}$$

D'après le théorème 25, G_x est définie sur $[-1,1]$ mais on remarque, en posant $a_n = P(X = n) > 0$ pour $n \geq 1$, que $\sum_{n\geq 1} a_n t^n$ est de rayon de convergence $\frac{1}{1-p}$. En effet, comme $|\frac{a_{n+1}}{a_n}| = (1-p) \xrightarrow[n\to+\infty]{} (1-p)$, d'après la règle de D'Alembert pour les séries entières, $R = \frac{1}{1-p}$.
Ainsi, si $X \sim \mathcal{G}(p)$, G_X est définie sur $]-\frac{1}{1-p}, \frac{1}{1-p}[$ (intervalle plus gros que $[-1,1]$ car $1-p<1$). De plus, pour tout $t \in]-\frac{1}{1-p}, \frac{1}{1-p}[$, on a :

$$\begin{aligned} G_X(t) &= \sum_{n=1}^{+\infty} t^n P(X=n) \\ &= \sum_{n=1}^{+\infty} t^n p(1-p)^{n-1} \\ &= pt \sum_{n=1}^{+\infty} (t(1-p))^{n-1} \\ &= pt \sum_{n=0}^{+\infty} ((1-p)t)^n \text{ et } t(1-p) \in]-1,1[\\ &= pt \frac{1}{1-(1-p)t} \\ G_X(t) &= \frac{pt}{1-(1-p)t} \end{aligned}$$

— Si $X \sim \mathcal{P}(\lambda)$ alors $G_X(t) = e^{\lambda(t-1)}$.

On suppose $X \sim \mathcal{P}(\lambda)$ avec $\lambda \in \mathbb{R}_+^*$. Alors :

$$\begin{cases} X(\Omega) = \mathbb{N} \\ P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \text{ pour tout } k \in \mathbb{N} \end{cases}$$

D'après le théorème 25, G_x est définie sur $[-1, 1]$ mais on remarque, en posant $a_n = P(X = n) > 0$ pour $n \in \mathbb{N}$, que $\sum a_n t^n$ est de rayon de convergence $+\infty$. En effet, comme $|\frac{a_{n+1}}{a_n}| = \frac{\lambda}{n+1} \xrightarrow[n\to+\infty]{} 0$, d'après la règle de D'Alembert pour les séries entières, $R = +\infty$.
Ainsi, si $X \sim \mathcal{P}(\lambda)$, G_X est définie sur $\mathbb{R}$. De plus, pour tout $t \in \mathbb{R}$, on a :

$$\begin{aligned} G_X(t) &= \sum_{n=1}^{+\infty} t^n P(X = n) \\ &= \sum_{n=0}^{+\infty} t^n e^{-\lambda}\frac{\lambda^n}{n!} \\ &= e^{-\lambda}\sum_{n=0}^{+\infty}\frac{(t\lambda)^n}{n!} \\ &= e^{-\lambda}e^{t\lambda} \\ G_X(t) &= e^{\lambda(t-1)}. \end{aligned}$$

Exercice 23.

Déterminer les fonctions génératrices des lois uniforme sur $[\![1, n]\!]$, de Bernoulli et binomiale.

Théorème 26.

Soit X, Y des variables aléatoires discrètes à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors, pour tout $t \in [-1, 1]$:

$$G_{X+Y}(t) = G_X(t)G_Y(t).$$

Démonstration.

On suppose X et Y indépendantes. Soit $t \in [-1, 1]$. On note $Z = X + Y$; alors Z est une variable aléatoire à valeurs dans $\mathbb{N}$ et on a :

$$G_Z(t) = E(t^Z) = \sum_{n=0}^{+\infty} t^n P(Z = n) = \sum_{n=0}^{+\infty} t^n P(X + Y = n)$$

Or, X et Y étant indépendantes, on a, pour tout $n \in \mathbb{N}$:

$$P(X+Y=n) = P\left(\bigcup_{k=0}^{n} ((X=k) \cap (y=n-k))\right) = \sum_{k=0}^{n} P(X=k)P(Y=n-k).$$

Ainsi :

$$G_Z(t) = \sum_{n=0}^{+\infty} t^n \sum_{k=0}^{n} P(X=k)P(Y=n-k) = \sum_{n=0}^{+\infty}\sum_{k=0}^{n} \underbrace{t^k P(X=k)}_{=a_k} \underbrace{t^{n-k} P(Y=n-k)}_{=b_{n-k}}.$$

où $a_n = t^n P(X=n)$ et $b_n = t^n P(Y=n)$ pour $n \in \mathbb{N}$.
On remarque que $G_X(t) = \sum_{n=0}^{+\infty} a_n$ et $G_Y(t) = \sum_{n=0}^{+\infty} b_n$ et que $\sum a_n, \sum b_n$ convergent absolument (car pour tout $n \in \mathbb{N}$, $|a_n| \leq P(X=n)$). Ainsi, le produit de Cauchy $\sum c_n$ de ces séries converge où $c_n = \sum_{k=0}^{n} a_n b_n$ et on a l'égalité :

$$G_{X+Y}(t) = G_Z(t) = \sum_{n=0}^{+\infty} c_n = \left(\sum_{n=0}^{+\infty} a_n\right)\left(\sum_{n=0}^{+\infty} b_n\right) = G_X(t)G_Y(t).$$

□

Remarque 15.

Plus généralement, si $X_1, ..., X_n$ sont des variables aléatoires discrètes indépendantes, alors $G_{X_1+...+X_n} = G_{X_1} \times \ldots \times G_{X_n}$.

Exercice 24.

Soit $k, m, n \in \mathbb{N}^*$ tels que $k = nm$ et X, Y des variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes. On pose $Z = X + Y$ et on suppose que $X \sim \mathcal{U}([\![0, n-1]\!])$ et $Z \sim \mathcal{U}([\![0, k-1]\!])$.

1. Déterminer la fonction génératrice de Y.
2. En déduire la loi de Y.