Espaces probabilisés 15

En première année, ont été étudiés les espaces probabilisés finis. Les concepts fondamentaux des probabilités (univers, événements, probabilité, variables aléatoires), qui sont les mêmes que dans le cas fini, vont être définis dans des espaces probabilisés quelconques. L'axiomatisation du calcul des probabilités qui est ici présentée est due principalement au mathématicien russe Andreï Kolmogorov (*Fondements de la théorie des probabilités*, 1933).

I Espaces probabilisés

1 Espaces probabilisables

Univers et événement

- Comme dans le cours de première année, il s'agit de modéliser une expérience aléatoire. On associe à une telle expérience un ensemble qui contient tous les résultats possibles de l'expérience, appelé **univers**. L'univers n'est plus supposé fini.

Exemples

1. On lance une pièce de monnaie jusqu'à l'obtention du premier pile. On s'intéresse au nombre de lancers nécessaires. L'univers que l'on peut associer à cette expérience est $\mathbb{N}^* \cup \{+\infty\}$, car *a priori* il n'est pas exclu que l'on n'obtienne jamais pile.
2. On veut modéliser l'expérience consistant à lancer une pièce de monnaie une infinité de fois. L'ensemble des résultats de l'expérience peut être représenté par l'ensemble $\Omega = \{0,1\}^{\mathbb{N}^*}$ des suites à valeurs dans $\{0,1\}$, où 1 représente pile et 0 face ; le terme d'indice n de la suite est le résultat du n-ième lancer. Cet ensemble n'est pas dénombrable (*cf.* exercice 7 de la page 452).

   Il est nécessaire de faire appel à un tel modèle quand le nombre de lancers effectués n'est pas fixé *a priori* : par exemple si on lance la pièce tant que la différence entre le nombre de piles et de faces est inférieure à 2.

- À une propriété que peut vérifier un résultat de l'expérience, on associe une partie de l'univers Ω que l'on appelle un événement. Dans le cas fini, on avait considéré que toute partie de Ω représentait un événement. On verra que dans le cas où Ω est dénombrable, on peut encore prendre $\mathcal{P}(\Omega)$ comme ensemble des événements.

  Dans les où Ω est infini et non dénombrable, ce choix conduit à des difficultés mathématiques dans la définition d'une probabilité qu'on ne surmonte qu'en restreignant l'ensemble des événements, qui sera alors une partie stricte de $\mathcal{P}(\Omega)$.

  Cela dit, l'ensemble des événements ne peut pas être une partie quelconque de $\mathcal{P}(\Omega)$.

  * L'ensemble Ω doit être un événement (l'événement certain).
  * Si A est un événement, alors $\overline{A}$, complémentaire de A dans Ω, doit être un événement (événement contraire).
  * Si A et B sont des événements, alors $A \cup B$ et $A \cap B$ doivent être des événements.
  * On impose une autre condition qui peut sembler moins naturelle : si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements, alors $\bigcup\limits_{n\in\mathbb{N}} A_n$ est un événement. En effet, la stabilité par réunion finie risque d'être insuffisante. Si l'on cherche par exemple à modéliser le jeu de pile ou face infini et qu'on considère pour tout $n \in \mathbb{N}^*$ l'événement A_n : « le n-ième tirage donne pile », on souhaitera s'intéresser à des événements comme « on obtient au moins un pile dans la suite des tirages » qui est $\bigcup\limits_{n=0}^{+\infty} A_n$.

- Enfin, une probabilité est, comme dans le cas d'un univers fini, une application qui à un événement associe un réel appartenant à $[0,1]$. Elle sera donc définie sur une partie de $\mathcal{P}(\Omega)$. On imposera une propriété d'additivité plus riche que dans les univers finis.

Tribu

Définition 1

Soit Ω un ensemble non vide. On appelle **tribu** sur Ω tout sous-ensemble $\mathcal{A}$ de $\mathcal{P}(\Omega)$ tel que :

- Ω appartient à $\mathcal{A}$;
- $\mathcal{A}$ est stable par passage au complémentaire, c'est-à-dire que, pour tout élément A de $\mathcal{A}$, $\overline{A}$ appartient à $\mathcal{A}$;
- $\mathcal{A}$ est stable par réunion dénombrable, c'est-à-dire, pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$, $\bigcup\limits_{n\in\mathbb{N}} A_n$ appartient à $\mathcal{A}$.

Remarque Au lieu de tribu, on dit aussi σ-algèbre, d'où la notation $\mathcal{A}$ adoptée par le programme.

Définition 2

Un **espace probabilisable** est un couple $(\Omega, \mathcal{A})$, où Ω est un ensemble quelconque non vide et $\mathcal{A}$ une tribu sur Ω.
L'ensemble Ω est appelé **l'univers** et un élément de $\mathcal{A}$ est appelé un **événement**.

Exemples

1. Pour tout ensemble non vide Ω, l'ensemble $\mathcal{P}(\Omega)$ est une tribu sur Ω.
2. Pour tout ensemble non vide de Ω et toute partie A de Ω, distincte de $\varnothing$ et de Ω, l'ensemble $\{\Omega, A, \bar{A}, \varnothing\}$ est une tribu sur Ω.

p.861

Exercice 1 Soit Ω un ensemble non vide et $\mathcal{F}$ une partie de $\mathcal{P}(\Omega)$.

1. Montrer qu'il existe des tribus sur Ω contenant $\mathcal{F}$ et que l'intersection de toutes ces tribus est une encore une tribu, que l'on note $\sigma(\mathcal{F})$.
2. Montrer que $\sigma(\mathcal{F})$ est la plus petite tribu sur Ω contenant $\mathcal{F}$.

La tribu $\sigma(\mathcal{F})$ est appelée **tribu engendrée** par $\mathcal{F}$.

Remarque Par exemple, si A est une partie de Ω, non vide et distincte de Ω, la tribu engendrée par $\{A\}$ est $\{\Omega, A, \bar{A}, \varnothing\}$.

p.861

Exercice 2 Montrer que si $(A_i)_{i\in I}$ est une partition d'un ensemble Ω, avec I au plus dénombrable, alors la tribu engendrée par $\{A_i\,;\ i \in I\}$ est l'ensemble $\mathcal{A}$ des $\bigcup\limits_{i\in J} A_i$, où J est une partie quelconque de I.

Opérations sur les événements

Proposition 1

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. Alors :

- $\varnothing$ appartient à $\mathcal{A}$;
- $\mathcal{A}$ est stable par intersection dénombrable, c'est-à-dire que, si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, alors $\bigcap\limits_{n\in\mathbb{N}} A_n$ appartient à $\mathcal{A}$;
- $\mathcal{A}$ est stable par réunion et intersection finie, c'est-à-dire que, si $n \in \mathbb{N}^*$ et si $A_1, \ldots, A_n$ sont des éléments de $\mathcal{A}$, alors $A_1 \cup \cdots \cup A_n$ et $A_1 \cap \cdots \cap A_n$ appartiennent à $\mathcal{A}$;
- pour tous éléments A et B de $\mathcal{A}$, $A \setminus B$ appartient à $\mathcal{A}$.

Démonstration page 862

Remarque On emploie le même vocabulaire que dans les espaces probabilisés finis. En particulier, deux événements A et B tels que $A \cap B = \varnothing$ sont dits incompatibles.

p.862 **Exercice 3** Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements d'un espace probabilisable $(\Omega, \mathcal{A})$. Montrer que l'ensemble B des éléments de Ω qui appartiennent à une infinité d'événements A_n et l'ensemble C des éléments de Ω qui appartiennent à tous les événements A_n sauf un nombre fini sont des événements, en les exprimant à l'aide des symboles $\cup$ et $\cap$.

Systèmes complets d'événements

Définition 3

On appelle **système complet d'événements** d'un espace probabilisable $(\Omega, \mathcal{A})$ toute famille au plus dénombrable $(A_i)_{i\in I}$ d'événements telle que :

- pour tout couple (i, j) d'éléments distincts de I, $A_i \cap A_j = \varnothing$;
- $\bigcup\limits_{i\in I} A_i = \Omega$.

Exemples

1. Pour tout événement A de Ω d'un espace probabilisable $(\Omega, \mathcal{A})$, la famille $(A, \overline{A})$ est un système complet d'événements.
2. Pour tout ensemble Ω au plus dénombrable, muni de la tribu $\mathcal{P}(\Omega)$, la famille $(\{\omega\})_{\omega\in\Omega}$ formée des événements élémentaires est un système complet d'événements.

2 Probabilité

Dans le cas d'un univers fini, on a défini une probabilité comme une application de $\mathcal{P}(\Omega)$ dans $[0, 1]$ telle que $\mathbb{P}(\Omega) = 1$ et additive, c'est-à-dire vérifiant :

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

pour tout couple d'événements incompatibles (A, B). Dans le cas où Ω est infini, on impose une condition d'additivité plus forte.

Définitions

Définition 4

On appelle **probabilité** sur l'espace probabilisable $(\Omega, \mathcal{A})$ toute application $\mathbb{P}$ de $\mathcal{A}$ dans $[0, 1]$ vérifiant :

1. $\mathbb{P}(\Omega) = 1$;
2. pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements **deux à deux incompatibles**, la série de terme général $\mathbb{P}(A_n)$ converge et :
$$\mathbb{P}\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

Pour tout $A \in \mathcal{A}$, $\mathbb{P}(A)$ est appelée **la probabilité de l'événement A**.

Remarques

- La seconde propriété est appelée **σ-additivité**.
- Il résulte de la définition que les événements sont les parties de Ω qui ont une probabilité.

Définition 5

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

Propriétés

Théorème 2

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On a alors :

1. $\mathbb{P}(\varnothing) = 0$;
2. pour toute famille finie $(A_1, A_2, \ldots, A_n)$ d'événements deux à deux incompatibles :
$$\mathbb{P}\left(A_1 \cup \ldots \cup A_n\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$
En particulier, pour tout couple (A, B) d'événements incompatibles :
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \ ;$$
3. pour tout événement A :
$$\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A) \ ;$$
4. pour tout couple (A, B) d'événements vérifiant $A \subset B$:
$$\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$$
et ainsi :
$$\mathbb{P}(A) \leqslant \mathbb{P}(B) \ ;$$
l'application $\mathbb{P}$ est donc croissante ;
5. pour tout couple (A, B) d'événements :
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Démonstration page 862

Remarque La σ-additivité peut se formuler ainsi, en intégrant le cas d'un nombre fini d'événements : si $(A_i)_{i \in I}$ est une famille au plus dénombrable d'événements deux à deux incompatibles, la famille $\left(\mathbb{P}(A_i)\right)_{i \in I}$ est sommable et :

$$\mathbb{P}\Big(\bigcup_{i\in I} A_i\Big) = \sum_{i\in I} \mathbb{P}(A_i).$$

En effet la sommabilité de la suite de réels positifs $\big(\mathbb{P}(A_i)\big)_{i\in\mathbb{N}}$ équivaut à la convergence de la série $\sum \mathbb{P}(A_i)$.

p.863

Exercice 4 On suppose que Ω est fini.

Montrer qu'une application $\mathbb{P} : \mathcal{P}(\Omega) \to [0,1]$ est une probabilité sur l'espace probabilisable $\big(\Omega, \mathcal{P}(\Omega)\big)$ si, et seulement si, $\mathbb{P}(\Omega) = 1$ et, pour tous événements A et B incompatibles :

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

Remarque Une probabilité sur un univers fini Ω telle que nous l'avons définie en première année apparaît donc comme un cas particulier de la définition donnée pour un espace probabilisable quelconque, la tribu des événements étant $\mathcal{P}(\Omega)$.

Définition 6

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

- Tout événement de probabilité nulle est dit **négligeable**.
- Tout événement de probabilité 1 est dit **presque sûr** ou **presque certain**.
- Toute propriété vérifiée sur un ensemble de probabilité 1 est dite **presque sûre**.

Propriétés des systèmes complets d'événements

Proposition 3

Si $(A_i)_{i\in I}$ est un système complet d'événements de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors :

- la famille $\mathbb{P}(A_i)$ est sommable et sa somme vaut 1 ;
- pour tout événement B, la famille $\mathbb{P}(B \cap A_i)$ est sommable et sa somme vaut $\mathbb{P}(B)$.

Démonstration page 863

La propriété suivante est une conséquence très forte de la σ-additivité.

Théorème 4 (Continuité monotone)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

1. Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ croissante, c'est-à-dire vérifiant $A_n \subset A_{n+1}$ pour tout n de $\mathbb{N}$, on a :
$$\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$
2. Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ décroissante, c'est-à-dire vérifiant $A_{n+1} \subset A_n$ pour tout n de $\mathbb{N}$, on a :
$$\mathbb{P}\Big(\bigcap_{n=0}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

Démonstration page 864

Corollaire 5

Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, on a :
$$\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcup_{k=0}^{n} A_k\Big),$$
$$\mathbb{P}\Big(\bigcap_{n=0}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcap_{k=0}^{n} A_k\Big).$$

Démonstration page 865

Proposition 6 (Inégalité de Boole)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

- Si A_0, $A_1, \ldots, A_n$ sont des événements, alors on a :
$$\mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \sum_{k=0}^{n} \mathbb{P}(A_k).$$
- Si $(A_k)_{k\in\mathbb{N}}$ est une suite d'événements, alors on a :
$$\mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right) \leqslant \sum_{k=0}^{+\infty} \mathbb{P}(A_k).$$

Démonstration page 865

Remarque On n'a pas nécessairement $\sum_{k=0}^{+\infty} \mathbb{P}(A_k) < +\infty$, mais si la série diverge, l'inégalité est sans intérêt.

L'inégalité de Boole peut être énoncée en englobant les deux cas.

Corollaire 7

Si $(A_i)_{i\in I}$ est une famille *au plus dénombrable* d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, on a :

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \leqslant \sum_{i\in I} \mathbb{P}(A_i).$$

Démonstration page 865

Corollaire 8

Une réunion au plus dénombrable d'événements négligeables est négligeable.

Démonstration page 865

Attention L'hypothèse « au plus dénombrable » est indispensable.

- Une réunion non dénombrable d'événements n'a aucune raison *a priori* d'être un événement et donc d'avoir une probabilité.
- Même si c'est un événement, il peut avoir une probabilité non nulle.

p.866

Exercice 5 (Premier lemme de Borel-Cantelli)

Soit (A_n) une suite d'événements de l'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On suppose que la série $\sum \mathbb{P}(A_n)$ converge. Montrer que l'événement $\bigcap\limits_{k=0}^{+\infty} \bigcup\limits_{p=k}^{+\infty} A_p$ est négligeable.

Union d'événements non incompatibles

Pour calculer la probabilité d'une réunion d'événements, il existe une formule qui généralise la formule $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ obtenue pour deux événements. Elle n'est pas au programme. Pour une démonstration et une utilisation, *cf.* exercice 16.11 de la page 954.

Formule du crible ou de Poincaré

Si $A_1, A_2, \ldots, A_n$ sont des événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, on a :

$$\begin{aligned}\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) &= \sum_{k=1}^{n} \left((-1)^{k-1} \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} \mathbb{P}(A_{i_1} \cap \ldots \cap A_{i_k})\right)\\ &= \sum_{k=1}^{n} \left((-1)^{k-1} \sum_{\substack{I\subset[\![1,n]\!]\\ \mathrm{card}\ I=k}} \mathbb{P}\big(\bigcap_{i\in I} A_i\big)\right).\end{aligned}$$

3 Probabilité sur un univers au plus dénombrable

Si l'univers Ω est au plus dénombrable, on peut prendre $\mathcal{P}(\Omega)$ comme tribu des événements. Alors comme dans le cas d'un univers fini, une probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ est entièrement déterminée par la connaissance des probabilités des événements élémentaires.

Théorème 9

Soit Ω un ensemble au plus dénombrable.

1. Soit $\mathbb{P}$ est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$. Si l'on pose $p_\omega = \mathbb{P}(\{\omega\})$, pour tout $\omega \in \Omega$, on a alors :
$$(1)\ \ \forall \omega \in \Omega \quad p_\omega \geqslant 0 \qquad \text{et} \qquad (2)\ \ \sum_{\omega \in \Omega} p_\omega = 1.$$
2. Réciproquement, si $(p_\omega)_{\omega \in \Omega}$ est une famille de réels vérifiant les conditions (1) et (2), il existe une unique probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ telle que $\mathbb{P}(\{\omega\}) = p_\omega$ pour tout $\omega \in \Omega$. La probabilité $\mathbb{P}$ est définie, pour tout événement A, par :
$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega.$$

Démonstration page 866

Remarque

Le cas d'un univers fini a été étudié en première année. Dans ce cas, on peut en particulier définir une probabilité pour laquelle les événements élémentaires ont tous même probabilité (probabilité uniforme). Ce n'est pas possible pour un univers dénombrable. En effet si Ω est dénombrable et si p_ω est une constante, soit cette constante n'est pas nulle et la famille $(p_\omega)_{\omega \in \Omega}$ n'est pas sommable, soit elle est nulle et $\sum\limits_{\omega \in \Omega} p_\omega$ vaut 0 et pas 1.

Exemples Considérons le cas $\Omega = \mathbb{N}$.

1. Si $p \in]0,1[$, on définit une probabilité sur $\mathbb{N}$, en posant :
$$\forall n \in \mathbb{N} \quad p_n = (1-p)^n p.$$
En effet, on a $p_n \geqslant 0$ pour tout $n \in \mathbb{N}$ et $\sum\limits_{n=0}^{+\infty} (1-p)^n p = \dfrac{p}{1-(1-p)} = 1$.
2. Si $\lambda > 0$, on définit une probabilité sur $\mathbb{N}$, en posant :
$$\forall n \in \mathbb{N} \quad p_n = e^{-\lambda} \frac{\lambda^n}{n!}.$$
En effet, on a $p_n \geqslant 0$ pour tout $n \in \mathbb{N}$ et $\displaystyle\sum_{n=0}^{+\infty} p_n = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = 1$.

p.866 **Exercice 6**

1. Montrer qu'on définit une probabilité sur $\mathbb{N}^*$ en posant :

$$\forall n \in \mathbb{N}^* \quad \mathbb{P}(\{n\}) = \frac{1}{n(n+1)}\cdot$$

2. Calculer la probabilité qu'un entier soit pair.

   On rappelle que $\sum\limits_{k=1}^{n} \frac{1}{k} = \ln n + \gamma + \mathrm{o}(1)$, quand n tend vers $+\infty$, où γ est la constante d'Euler (*cf.* page 409).

Remarque Définir une probabilité sur un univers non dénombrable est beaucoup plus compliqué. On ne peut pas la définir par les probabilités des événements élémentaires, car une partie non dénombrable ne peut pas être obtenue comme union disjointe dénombrable de singletons. Toutefois, considérer des univers non dénombrables est indispensable, ne serait-ce que pour modéliser un pile ou face infini, dont nous avons vu qu'il conduit à l'univers non dénombrable $\{0,1\}^{\mathbb{N}^*}$ (nous reviendrons sur cette modélisation dans la remarque de la page 859).

II Probabilités conditionnelles

Les définitions qui suivent généralisent simplement les définitions vues en première année dans le cas d'un univers fini.

1 Définition

Théorème 10

Pour tout événement A de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ de probabilité non nulle, l'application :

$$\begin{array}{rccc} \mathbb{P}_A : & \mathcal{A} & \longrightarrow & \mathbb{R} \\ & B & \longmapsto & \dfrac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \end{array}$$

est une probabilité sur $(\Omega, \mathcal{A})$, appelée la **probabilité conditionnelle** sachant A.

Pour tout événement B, $\mathbb{P}_A(B)$ qui est encore noté $\mathbb{P}(B \mid A)$ est appelé la **probabilité conditionnelle de B sachant** A.

Démonstration page 867

2 Formules liées aux probabilités conditionnelles

Formule des probabilités composées

Théorème 11 (Formule des probabilités composées)

Soit n un entier naturel supérieur ou égal à 2.
Pour toute famille finie $(A_1, A_2, \ldots, A_n)$ d'événements de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tels que $\mathbb{P}(A_1 \cap \cdots \cap A_{n-1}) \neq 0$, on a :

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_{n-1} \cap A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1)\,\mathbb{P}(A_3 \mid A_1 \cap A_2) \cdots \mathbb{P}(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Démonstration page 867

Formule des probabilités totales

Théorème 12 (La formule des probabilités totales)

Soit $(A_i)_{i\in I}$ un système complet d'événements de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tel que, pour tout entier i, on ait $\mathbb{P}(A_i) \neq 0$.
Pour tout événement B, on a :

$$\mathbb{P}(B) = \sum_{i\in I} \mathbb{P}(B \mid A_i)\,\mathbb{P}(A_i).$$

Démonstration page 868

Remarque On peut s'affranchir de la condition $\mathbb{P}(A_i) \neq 0$ pour tout $i \in I$ en convenant, si $\mathbb{P}(A_i) = 0$, d'attribuer la valeur 0 au produit $\mathbb{P}(B \mid A_i)\mathbb{P}(A_i)$. En effet, on a alors $\mathbb{P}(B \cap A_i) = 0$ car $B \cap A_i \subset A_i$.

p.868 **Exercice 7** On dit que $(A_i)_{i\in I}$ un système quasi-complet d'événements si I est au plus dénombrable et si les A_i sont des événements deux à deux incompatibles dont la réunion est de probabilité 1. Montrer que la formule des probabilités totales reste vraie avec un système quasi-complet d'événements.

Formule de Bayes

Théorème 13 (Formule de Bayes)

Si A et B sont deux événements de probabilité non nulle d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, on a :

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A)\,\mathbb{P}(A)}{\mathbb{P}(B)}\cdot$$

Démonstration. Cela résulte de la définition des probabilités conditionnelles. □

Point méthode

On combine souvent la formule de Bayes avec la formule des probabilités totales. Soit $(A_i)_{i\in I}$ un système complet d'événements et B un événement. Si A_j $(j \in I)$ et B ont une probabilité non nulle, alors on a :

$$\mathbb{P}(A_j \mid B) = \frac{\mathbb{P}(B \mid A_j)\,\mathbb{P}(A_j)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A_j)\,\mathbb{P}(A_j)}{\sum\limits_{i\in I} \mathbb{P}(B \mid A_i)\,\mathbb{P}(A_i)}.$$

p.868 **Exercice 8** Dans une population donnée, on suppose qu'il existe $p \in \,]0,1[$ et $\alpha > 0$ tels que la probabilité p_n qu'une famille ait n enfants vérifie :

$$\forall n \geqslant 1 \quad p_n = \alpha p^n.$$

On suppose qu'un enfant a la même probabilité d'être une fille ou un garçon.

1. Calculer p_0.
2. Déterminer la probabilité qu'une famille ait deux enfants, sachant qu'elle a deux filles.
3. Déterminer la probabilité qu'une famille ait deux garçons sachant qu'elle a deux filles.

3 Indépendance des événements

Les définitions qui suivent généralisent les définitions vues en première année dans le cas d'un univers fini.

Indépendance entre deux événements

Définition 7

Deux événements A et B de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sont dits indépendants si :

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

Proposition 14

Deux événements A et B de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tels que $\mathbb{P}(A) \neq 0$ sont indépendants si, et seulement si, $\mathbb{P}(B \mid A) = \mathbb{P}(B)$.

Remarques

- Si l'événement A est négligeable, alors il est indépendant de tout événement B. En effet on a $\mathbb{P}(A) = 0$ et *a fortiori* $\mathbb{P}(A \cap B) = 0$.
- L'indépendance est une notion probabiliste. Elle dépend de la probabilité dont est muni $(\Omega, \mathcal{A})$.
- L'indépendance ne doit pas être confondue avec l'incompatibilité : deux événements de probabilité non nulle, incompatibles, ne sont pas indépendants.

Proposition 15

Si A et B sont deux événements indépendants de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, les événements A et $\overline{B}$, les événements $\overline{A}$ et B et les événements $\overline{A}$ et $\overline{B}$ sont indépendants.

Démonstration page 869

Remarque On en déduit que si A est un événement presque sûr, il est indépendant de tout événement B. En effet $\overline{A}$ est de probabilité nulle donc $\overline{A}$ et B sont indépendants, comme nous l'avons déjà vu.

Indépendance d'une famille d'événements.

Définition 8

Soit $(A_i)_{i\in I}$ une famille quelconque d'événements de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On dit que les événements A_i, pour $i \in I$, sont **deux à deux indépendants** si, pour tout couple (i,j) d'éléments distincts de I, on a :

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j).$$

On dit que les événements A_i, pour $i \in I$, sont **mutuellement indépendants** ou que $(A_i)_{i\in I}$ est une **famille d'événements mutuellement indépendants** si, pour toute partie finie J de I, on a :

$$\mathbb{P}\Big(\bigcap_{i\in J} A_i\Big) = \prod_{i\in J} \mathbb{P}(A_i).$$

Remarques

- Souvent, pour une famille d'événements mutuellement indépendants, on dit simplement **indépendants**.
- Si les événements A_i, pour $i \in I$ sont mutuellement indépendants, alors ils sont deux à deux indépendants, comme on le voit en prenant pour J un ensemble à deux éléments.

  Comme dans le cas des univers finis, **la réciproque est fausse**.
- L'indépendance ne dépend pas de l'ordre des événements. Si $(A_i)_{i\in I}$ est une famille d'événements mutuellement indépendants, il en est de même de la famille $(A_{\sigma(i)})_{i\in I}$, où σ est une permutation de I.
- Si $(A_i)_{i\in I}$ est une famille d'événements mutuellement indépendants, alors, pour toute partie I' de I, $(A_i)_{i\in I'}$ est encore une famille d'événements mutuellement indépendants.

p.869 **Exercice 9** Montrer que $(A_i)_{i\in I}$ est une famille d'événements indépendants si, et seulement si, toute sous-famille finie est constituée d'événements indépendants.

Proposition 16

Si $(A_i)_{i\in I}$ est une famille d'événements mutuellement indépendants, et si, pour tout $i \in I$, B_i est égal à A_i ou $\overline{A}_i$, alors $(B_i)_{i\in I}$ est une famille d'événements mutuellement indépendants.

Démonstration page 870

Quand on a une famille d'événements mutuellement indépendants, si on remplace certains d'entre eux par leur réunion ou leur intersection, on obtient encore des événements mutuellement indépendants. C'est l'objet de l'exercice suivant.

p.870 **Exercice 10** Soit $(A_i)_{i\in I}$ une famille d'événements indépendants d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, ainsi que I_1, I_2,..., I_n des parties finies et deux à deux disjointes de I. Pour $1 \leqslant j \leqslant n$, on pose $B_j = \bigcap\limits_{i\in I_j} A_i$.

1. Montrer que $(B_1, B_1, \ldots, B_n)$ est une famille d'événements indépendants.
2. Montrer qu'il en est de même si on remplace $\bigcap\limits_{i\in I_j} A_i$ par $\bigcup\limits_{i\in I_j} A_i$ dans la définition de certains événements B_j.

Modèle de pile ou face infini Nous admettons qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, avec $\Omega = \{0,1\}^{\mathbb{N}^*}$, permettant de modéliser une suite infinie de lancers de pile ou face indépendants. Il vérifie les propriétés suivantes : pour tout $n \in \mathbb{N}^*$, l'ensemble A_n : « le n-ième lancer donne pile » est un événement de probabilité $p \in]0,1[$ et $(A_n)_{n\in\mathbb{N}^*}$ est une famille d'événements mutuellement indépendants.

p.871 **Exercice 11** On garde les mêmes notations. Montrer que, pour tout $\omega \in \Omega$, le singleton $\{\omega\}$ est un événement, de probabilité nulle.

Remarque On peut démontrer que dans un tel modèle, $\mathcal{A}$ ne peut pas être égal à $\mathcal{P}(\Omega)$. En fait, on peut prendre pour $\mathcal{A}$ la tribu engendrée par $\{A_n\,;\, n \in \mathbb{N}^*\}$ (*cf.* exercice 1 de la page 848) .

p.871 **Exercice 12 (Second lemme de Borel-Cantelli)**

1. Soit (A_n) une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, P)$ On suppose que la série $\sum \mathbb{P}(A_n)$ diverge et que les événements A_n sont mutuellement indépendants. Montrer que l'événement $\bigcap\limits_{k=0}^{+\infty} \bigcup\limits_{p=k}^{+\infty} A_p$ est presque sûr.
2. Application. Une pièce de monnaie amène pile avec la probabilité p $(0 < p < 1)$ et face avec la probabilité $q = 1-p$. On la lance une infinité de fois. Montrer que, pour tout $m \geqslant 1$, il apparaît une infinité de séquences de m piles consécutifs de façon presque sûre.

Remarques

- On peut montrer de la même façon qu'on obtient de manière presque sûre une infinité de fois n'importe quelle séquence de piles ou faces.
- En tenant compte du premier lemme de Borel-Cantelli (exercice 5 de la page 853), on voit que si les événements A_n sont mutuellement indépendants, la probabilité de $\bigcap\limits_{k=0}^{+\infty} \bigcup\limits_{p=k}^{+\infty} A_p$, c'est-à-dire la probabilité qu'une infinité d'événements A_n se réalisent, ne peut être que 0 ou 1.

Démonstrations et solutions des exercices du cours

Exercice 1

1. Notons T l'ensemble de toutes les tribus contenant $\mathcal{F}$. Il est non vide, car $\mathcal{P}(\Omega)$ est une tribu sur Ω contenant $\mathcal{F}$. Posons $\sigma(\mathcal{F}) = \bigcap\limits_{\mathcal{A}\in T} \mathcal{A}$. Montrons que $\sigma(\mathcal{F})$ est une tribu.
   - Par définition d'une tribu, $\Omega \in \mathcal{A}$, pour tout $\mathcal{A} \in T$, donc $\Omega \in \sigma(\mathcal{F})$.
   - Soit $A \in \sigma(\mathcal{F})$. Pour tout $\mathcal{A} \in T$, on a $A \in \mathcal{A}$ et donc $\bar{A} \in \mathcal{A}$, par définition d'une tribu. Donc $\bar{A} \in \sigma(\mathcal{F})$.
   - Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\sigma(\mathcal{F})$. Soit $\mathcal{A} \in T$. Alors, pour tout $n \in \mathbb{N}$, $A_n \in \mathcal{A}$, donc par définition d'une tribu, $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{A}$. C'est vrai pour tout $\mathcal{A} \in T$, donc $\bigcup\limits_{n\in\mathbb{N}} A_n \in \sigma(\mathcal{F})$.

   Ainsi $\sigma(\mathcal{F})$ est une tribu sur Ω.
2. - Par définition $\mathcal{F} \subset \mathcal{A}$ pour tout $\mathcal{A} \in T$ donc $\mathcal{F} \subset \sigma(\mathcal{F})$.
   - Si $\mathcal{A}$ est une tribu quelconque contenant $\mathcal{F}$, alors $\mathcal{A} \in T$, donc $\sigma(\mathcal{F}) \subset \mathcal{A}$.

   On a bien montré que $\sigma(\mathcal{F})$ est la plus petite tribu sur Ω contenant $\mathcal{F}$.

Exercice 2

- Par définition, $\mathcal{A}$ contient $\bigcup\limits_{i\in I} A_i = \Omega$.

  Soit $B \in \mathcal{A}$ et J une partie de I telle que $B = \bigcup\limits_{i\in J} A_i$. Comme les A_i sont disjoints, on a :
  $$\overline{B} = \Omega \setminus B = \bigcup_{i\in I} A_i \setminus \bigcup_{i\in J} A_i = \bigcup_{i\in I\setminus J} A_i$$
  donc $\overline{B}$ appartient à $\mathcal{A}$.
  Si $(B_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}$, il existe pour tout $n \in \mathbb{N}$ une partie J_n de I telle que $B_n = \bigcup\limits_{i\in J_n} A_i$. On a alors :
  $$\bigcup_{n\in\mathbb{N}} B_n = \bigcup_{i\in \bigcup\limits_{n\in\mathbb{N}} J_n} A_i,$$
  ce qui montre que $\bigcup\limits_{n\in\mathbb{N}} B_n \in \mathcal{A}$.
  Ainsi, $\mathcal{A}$ est une tribu de Ω qui, de manière évidente, contient $\{A_i\,;\ i \in I\}$.
- Soit $\mathcal{A}'$ une tribu contenant $\{A_i\,;\ i \in I\}$. Pour tout $J \subset I$, comme I est au plus dénombrable, il en est de même de J. Chaque A_i est dans $\mathcal{A}'$ donc $\bigcup\limits_{i\in J} A_i$ appartient aussi à $\mathcal{A}'$. Ainsi $\mathcal{A} \subset \mathcal{A}'$.

Proposition 1

- $\varnothing$ appartient à $\mathcal{A}$ car Ω appartient à $\mathcal{A}$ et $\mathcal{A}$ est stable par passage au complémentaire.
- Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathcal{A}$. Pour tout entier n, $\overline{A_n}$ appartient à $\mathcal{A}$, donc $\bigcup\limits_{n=0}^{+\infty}\overline{A_n}$ appartient à $\mathcal{A}$. Comme $\bigcup\limits_{n=0}^{+\infty}\overline{A_n}=\overline{\bigcap\limits_{n=0}^{+\infty}A_n}$, on en déduit, par passage au complémentaire, que $\bigcap\limits_{n=0}^{+\infty}A_n$ appartient à $\mathcal{A}$.
- Si $A_1,\ldots,A_n$ sont des éléments de $\mathcal{A}$, en posant $A_i=\varnothing$ si $i=0$ ou $i>n$, on obtient, par stabilité par réunion dénombrable $\bigcup\limits_{i=0}^{+\infty}A_i\in\mathcal{A}$, c'est-à-dire $A_1\cup\ldots\cup A_n\in\mathcal{A}$.
  De même, en posant $A_0=\Omega$ et $A_i=\Omega$ pour $i>n$, on obtient, par stabilité par intersection dénombrable, $\bigcap\limits_{i=0}^{+\infty}A_i\in\mathcal{A}$, c'est-à-dire $A_1\cap\ldots\cap A_n\in\mathcal{A}$.
- Si A et B appartiennent à $\mathcal{A}$, il en est de même de $A\setminus B=A\cap\overline{B}$, d'après les propriétés précédentes.

Exercice 3 Considérons $\overline{B}$, ensemble des éléments de Ω appartenant seulement à un nombre fini d'événements A_n. Alors, on a, pour tout $\omega\in\Omega$:

$$\omega\in\overline{B}\iff\exists n\in\mathbb{N}\quad\forall p\geqslant n\quad\omega\notin A_p.$$

On a donc $\overline{B}=\bigcup\limits_{p\in\mathbb{N}}\bigcap\limits_{p\geqslant n}\overline{A_p}$, d'où l'on déduit $B=\bigcap\limits_{n\in\mathbb{N}}\bigcup\limits_{p\geqslant n}A_p$.

De même, on a pour tout $\omega\in\Omega$:

$$\omega\in C\iff\exists n\in\mathbb{N}\quad\forall p\geqslant n\quad\omega\in A_p$$

et donc $C=\bigcup\limits_{n\in\mathbb{N}}\bigcap\limits_{p\geqslant n}A_p$.

Comme une tribu est stable par union et intersection dénombrable, B et C sont des événements.

Théorème 2

1. Considérons la famille d'événements incompatibles $(A_n)_{i\in\mathbb{N}}$, où $A_n=\varnothing$ pour tout entier naturel n. Alors, par définition d'une probabilité, la série de terme général constant $\mathbb{P}(\varnothing)$ converge, d'où $\mathbb{P}(\varnothing)=0$.
2. Soit $(A_1,\ldots,A_n)$ une famille finie d'événements deux à deux incompatibles. La suite $(A_k)_{k\in\mathbb{N}}$, où $A_k=\varnothing$ pour tout entier $k\geqslant n+1$ et pour $k=0$, est une famille d'événements deux à deux incompatibles.
   Alors, par définition d'une probabilité :
   $$\mathbb{P}\left(\bigcup_{k=0}^{+\infty}A_k\right)=\sum_{k=0}^{+\infty}\mathbb{P}(A_k)=\mathbb{P}(A_0)+\sum_{k=0}^{n}\mathbb{P}(A_k)+\sum_{i=n+1}^{+\infty}\mathbb{P}(A_k).$$
   Comme $\mathbb{P}(\varnothing)=0$ et $\bigcup\limits_{k=0}^{+\infty}A_k=A_1\cup\ldots\cup A_n$, on obtient :
   $$\mathbb{P}(A_1\cup\ldots\cup A_n)=\sum_{k=1}^{n}\mathbb{P}(A_k).$$

En particulier, pour deux événements incompatibles, on trouve :
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

3. On a, d'après le point précédent, puisque A et $\overline{A}$ sont incompatibles :
$$\mathbb{P}(A) + \mathbb{P}(\overline{A}) = P\left(A \cup \overline{A}\right) = \mathbb{P}(\Omega) = 1$$
et donc $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
4. Si $A \subset B$, alors B est la réunion des événements incompatibles A et $B \setminus A$ et donc :
$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A).$$
On en déduit $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ et $\mathbb{P}(B) \geqslant \mathbb{P}(A)$.
5. L'événement $A \cup B$ est la réunion des deux événements incompatibles A et $B \setminus A$; d'autre part, B est la réunion des événements incompatibles $A \cap B$ et $B \setminus A$. On a donc :
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \quad \text{et} \quad \mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus A).$$
On en déduit :
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Exercice 4

- Si $\mathbb{P}$ est une probabilité, alors $\mathbb{P}(\Omega) = 1$ et, d'après la proposition 2, pour tous événements A et B incompatibles, on a $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.
- Réciproquement, supposons que $\mathbb{P}(\Omega) = 1$ et que, pour tous événements A et B incompatibles, on a $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

  Par récurrence sur n, on démontre que, pour des événements deux à deux incompatibles $A_0, \ldots, A_n$, on a $\mathbb{P}(A_1 \cup \ldots \cup A_n) = \sum\limits_{k=0}^{n} \mathbb{P}(A_k)$.

  Comme Ω est fini, l'ensemble des événements est fini. Une suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles ne comporte donc qu'un nombre fini d'événements différents de $\varnothing$. De plus, de $\mathbb{P}(\varnothing) = \mathbb{P}(\varnothing) + \mathbb{P}(\varnothing)$, on tire $\mathbb{P}(\varnothing) = 0$. En notant $I = \{n \in \mathbb{N} \mid A_n \neq \varnothing\}$, on obtient :
$$\mathbb{P}\Bigg(\bigcup_{n\in\mathbb{N}} A_n\Bigg) = \mathbb{P}\Bigg(\bigcup_{n\in I} A_n\Bigg) = \sum_{n\in I} \mathbb{P}(A_n) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$
  Donc $\mathbb{P}$ est une probabilité.

Proposition 3 La famille $(A_i)_{i\in I}$ est au plus dénombrable et les événements A_i sont deux à deux incompatibles, donc :
$$1 = \mathbb{P}(\Omega) = \mathbb{P}\Bigg(\bigcup_{i\in I} A_i\Bigg) = \sum_{i\in I} \mathbb{P}(A_i).$$
De même, les événements $B \cap A_i$ sont deux à deux incompatibles et $\bigcup\limits_{i\in I}(B \cap A_i) = B$, donc :
$$\mathbb{P}(B) = \mathbb{P}\Bigg(\bigcup_{i\in I}(B \cap A_i)\Bigg) = \sum_{i\in I} \mathbb{P}\left(B \cap A_i\right).$$

Théorème 4

1. On considère la suite $(B_n)_{n\in\mathbb{N}}$ définie par $B_0 = A_0$ et, pour tout $n \in \mathbb{N}^*$, $B_n = A_n \setminus A_{n-1}$.

   On note que, pour tout $n \in \mathbb{N}$, on a $B_n \subset A_n$ donc $\bigcup\limits_{n=0}^{+\infty} B_n \subset \bigcup\limits_{n=0}^{+\infty} A_n$. Mais d'autre part, si $\omega \in \bigcup\limits_{n=0}^{+\infty} A_n$ et si m est le plus petit des entiers k tels que $\omega \in A_k$, alors on a $\omega \in A_m \setminus A_{m-1} = B_m$ et donc $\omega \in \bigcup\limits_{n=0}^{+\infty} B_n$. On a donc $\bigcup\limits_{n=0}^{+\infty} A_n = \bigcup\limits_{n=0}^{+\infty} B_n$.

   On montre de même que, pour tout $n \in \mathbb{N}$, on a $A_n = \bigcup\limits_{k=0}^{n} B_k$. En effet, on a par définition :
$$\bigcup_{k=0}^{n} B_k \subset \bigcup_{k=0}^{n} A_k \subset A_n$$
   et réciproquement si $\omega \in A_n$ et si m est le plus petit des entiers k tels que $\omega \in A_k$, alors $m \leqslant n$ et $\omega \in B_m$ donc $\omega \in \bigcup\limits_{k=0}^{n} B_k$.
   Les événements B_n sont incompatibles. En effet si $m < n$, on a $B_m \subset A_m \subset A_{n-1}$ et $B_n \subset \overline{A_{n-1}}$ donc $B_m \cap B_n = \varnothing$. On a donc :
$$\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} A_n\Big) = \mathbb{P}\Big(\bigcup_{n=0}^{+\infty} B_n\Big) = \sum_{n=0}^{+\infty} \mathbb{P}(B_n) = \lim_{n\to+\infty} \sum_{k=0}^{n} \mathbb{P}(B_k).$$
   De plus on a, toujours par incompatibilité des événements B_n, l'égalité :
$$\sum_{k=0}^{n} \mathbb{P}(B_k) = \mathbb{P}\Big(\bigcup_{k=0}^{n} B_k\Big) = \mathbb{P}(A_n).$$
   Finalement, on obtient :
$$\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} A_n\Big) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

2. Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements, alors la suite $\big(\overline{A_n}\big)_{n\in\mathbb{N}}$ est une suite croissante d'événements. En appliquant le premier point, on trouve donc :
$$\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} \overline{A_n}\Big) = \lim_{n\to+\infty} \mathbb{P}\big(\overline{A_n}\big).$$
   On en déduit :
$$\mathbb{P}\Big(\bigcap_{n=0}^{+\infty} A_n\Big) = 1-\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} \overline{A_n}\Big) = 1-\lim_{n\to+\infty} \mathbb{P}\big(\overline{A_n}\big) = \lim_{n\to+\infty} \big(1-\mathbb{P}\big(\overline{A_n}\big)\big) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

Corollaire 5

1. La suite $(B_n)_{n\in\mathbb{N}}$ définie par $B_n = \bigcup\limits_{k=0}^{n} A_k$ est croissante et $\bigcup\limits_{n=0}^{+\infty} B_n = \bigcup\limits_{n=0}^{+\infty} A_n$. On a donc :
$$\mathbb{P}\Big(\bigcup_{n=0}^{+\infty} A_n\Big) = \mathbb{P}\Big(\bigcup_{n=0}^{+\infty} B_n\Big) = \lim_{n\to+\infty} \mathbb{P}(B_n) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcup_{k=0}^{n} A_k\Big).$$

2. La suite $(C_n)_{n\in\mathbb{N}}$ définie par $C_n = \bigcap\limits_{k=0}^{n} A_k$ est décroissante et $\bigcap\limits_{n=0}^{+\infty} C_n = \bigcap\limits_{n=0}^{+\infty} A_n$. On a donc :
$$\mathbb{P}\Big(\bigcap_{n=0}^{+\infty} A_n\Big) = \mathbb{P}\Big(\bigcap_{n=0}^{+\infty} C_n\Big) = \lim_{n\to+\infty} \mathbb{P}(C_n) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcap_{k=0}^{n} A_k\Big).$$

Proposition 6

- On démontre la propriété par récurrence sur n. La propriété est évidente pour $n = 0$ et si elle est vérifiée au rang n, alors :
$$\begin{aligned}
\mathbb{P}\left(\bigcup_{k=0}^{n+1} A_k\right) &= \mathbb{P}\left(\left(\bigcup_{k=0}^{n+1} A_k\right) \cup A_{n+1}\right)\\
&= \mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\left(\bigcup_{k=0}^{n+1} A_k\right) \cap A_{n+1}\right)\\
&\leqslant \mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right) + \mathbb{P}(A_{n+1})\\
&\leqslant \sum_{k=0}^{n} \mathbb{P}(A_k) + \mathbb{P}(A_{n+1}) = \sum_{k=0}^{n+1} \mathbb{P}(A_k).
\end{aligned}$$

- On utilise l'inégalité démontrée dans le point précédent. On obtient, par passage à la limite, en utilisant le corollaire 5 :
$$\mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_n\right) \leqslant \sum_{k=0}^{+\infty} \mathbb{P}(A_k).$$

Corollaire 7 Si I est de la forme $[\![0, n]\!]$, avec $n \in \mathbb{N}$, ou est égal à $\mathbb{N}$, l'inégalité énoncée est l'inégalité de Boole. Sinon I est en bijection avec un de ces ensembles et l'inégalité résulte de l'inégalité de Boole par invariance des deux membres de l'inégalité par permutation des termes.

Corollaire 8 Si $(A_i)_{i\in I}$ une famille au plus dénombrable d'événements négligeables, on a $\sum\limits_{i\in I} \mathbb{P}(A_i) = 0$ et d'après l'inégalité de Boole $\mathbb{P}\left(\bigcup\limits_{i\in I} A_i\right) = 0$.

Exercice 5 On pose $B = \bigcap_{k=0}^{+\infty} \bigcup_{p=k}^{+\infty} A_p$. La suite $\left(\bigcup_{p=k}^{+\infty} A_p \right)_{k \in \mathbb{N}}$ est décroissante donc $\mathbb{P}(B) = \lim_{k \to +\infty} \bigcup_{p=k}^{+\infty} A_p$. Pour tout $k \in \mathbb{N}$, on a :

$$0 \leqslant \mathbb{P}\left(\bigcup_{p=k}^{+\infty} A_p \right) \leqslant \sum_{p=k}^{+\infty} \mathbb{P}(A_p).$$

On reconnaît dans $\sum_{p=k}^{+\infty} \mathbb{P}(A_p)$ le reste d'une série convergente, qui tend donc vers 0.

On en déduit $\mathbb{P}(B) = 0$.

Théorème 9

1. Si $\mathbb{P}$ est une probabilité sur Ω, les propriétés (1) et (2) résultent de la définition d'une probabilité et du fait que $(\{\omega\})_{\omega \in \Omega}$ est un système complet d'événements.
2. Réciproquement, supposons que les propriétés (1) et (2) soient vérifiées. Raisonnons par analyse-synthèse.
   - Si la probabilité $\mathbb{P}$ existe, on écrit tout événement A sous la forme $A = \bigcup_{\omega \in A} \{\omega\}$.

     Puisque A est au plus dénombrable, on obtient, par σ-additivité :

     $$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \sum_{\omega \in A} p_\omega.$$
   - Montrons que l'application $\mathbb{P}$ ainsi définie est une probabilité.
     - Pour tout événement A, on a $0 \leqslant \mathbb{P}(A) \leqslant \sum_{\omega \in \Omega} p_\omega \leqslant 1$, car les p_ω sont positifs et $\sum_{\omega \in \Omega} p_\omega = 1$. De plus, $\mathbb{P}(\Omega) = 1$.
     - Si (A_n) est une suite d'événements deux à deux incompatibles, alors en notant $A = \bigcup_{n=0}^{+\infty} A_n$ et en appliquant le théorème de sommation par paquets, on obtient :

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

     Donc $\mathbb{P}$ est une probabilité et, par définition de $\mathbb{P}$, on a $\mathbb{P}(\{\omega\}) = p_\omega$, pour tout $\omega \in \Omega$.

Exercice 6

1. On vérifie les deux conditions du théorème 9.

   On a $\dfrac{1}{n(n+1)} > 0$ pour tout $n \in \mathbb{N}^*$. La série $\sum \dfrac{1}{n(n+1)}$ converge. C'est une série télescopique car $\dfrac{1}{n(n+1)} = \dfrac{1}{n} - \dfrac{1}{n+1}$.

   On a donc $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1 - \lim_{n \to +\infty} \frac{1}{n+1} = 1$.

2. On a $\mathbb{P}(2\mathbb{N}^*) = \sum\limits_{n=1}^{+\infty} \dfrac{1}{2n(2n+1)}\cdot$ Pour $N \in \mathbb{N}^*$, on obtient :

$$\begin{aligned}
\sum_{n=1}^{N} \frac{1}{2n(2n+1)} &= \sum_{n=1}^{N} \left(\frac{1}{2n} - \frac{1}{2n+1}\right) \\
&= 2\sum_{n=1}^{N} \frac{1}{2n} - \sum_{n=1}^{N} \left(\frac{1}{2n} + \frac{1}{2n+1}\right) \\
&= \sum_{n=1}^{N} \frac{1}{n} - \sum_{n=2}^{2N+1} \frac{1}{n} \\
&= \ln N + \gamma + 1 - \ln(2N+1) - \gamma + \mathrm{o}(1) \\
&= 1 - \ln\left(2 + \frac{1}{N}\right) + \mathrm{o}(1),
\end{aligned}$$

On en déduit $\mathbb{P}(2\mathbb{N}^*) = 1 - \ln 2$.

Théorème 10 Montrons que $\mathbb{P}_A$ est une probabilité sur $(\Omega, \mathcal{A})$.

- On a $\mathbb{P}_A(\Omega) = \dfrac{\mathbb{P}(\Omega \cap A)}{\mathbb{P}(A)} = 1$.
- Soit $(B_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux incompatibles. La famille $(B_n \cap A)_{n\in\mathbb{N}}$ est aussi une suite d'événements deux à deux incompatibles. on a alors par additivité :

$$\mathbb{P}\Big(\Big(\bigcup_{n\in\mathbb{N}} B_n\Big) \cap A\Big) = \mathbb{P}\Big(\bigcup_{n\in\mathbb{N}} (B_n \cap A)\Big) = \sum_{n=0}^{+\infty} \mathbb{P}\big(B_n \cap A\big).$$

On a donc :

$$\mathbb{P}_A\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \frac{\mathbb{P}\left(\big(\bigcup_{n\in\mathbb{N}} B_n\big) \cap A\right)}{\mathbb{P}(A)} = \frac{\sum\limits_{n=0}^{+\infty} \mathbb{P}(B_n \cap A)}{\mathbb{P}(A)} = \sum_{n=0}^{+\infty} \mathbb{P}_A(B_n).$$

Théorème 11 Pour tout entier k compris entre 1 et $n-1$, on a :

$$A_1 \cap \ldots \cap A_{n-1} \subset A_1 \cap \ldots \cap A_k,$$

d'où $\mathbb{P}(A_1 \cap \ldots \cap A_k) \geqslant \mathbb{P}(A_1 \cap \ldots \cap A_{n-1}) > 0$. Toutes les probabilités conditionnelles sont donc définies. Il suffit alors d'écrire :

$$\begin{aligned}
&\mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1)\,\mathbb{P}(A_3 \mid A_1 \cap A_2) \ldots \mathbb{P}(A_n \mid A_1 \cap \ldots \cap A_{n-1}) \\
&\qquad = \mathbb{P}(A_1) \prod_{k=2}^{n} \mathbb{P}\left(A_k \mid A_1 \cap \ldots \cap A_{k-1}\right) \\
&\qquad = \mathbb{P}(A_1) \prod_{k=2}^{n} \frac{\mathbb{P}\big(A_1 \cap \ldots A_{k-1} \cap A_k\big)}{\mathbb{P}\big(A_1 \cap \ldots \cap A_{k-1}\big)} \\
&\qquad = \mathbb{P}\big(A_1 \cap \ldots \cap A_{n-1} \cap A_n\big),
\end{aligned}$$

car le produit de droite est télescopique : tous les termes se simplifient deux à deux sauf $\mathbb{P}\big(A_1 \cap \ldots \cap A_{n-1} \cap A_n\big)$.

Théorème 12 On a :

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

Exercice 7 On note $C = \overline{\bigcup_{i \in I} A_i}$, événement de probabilité nulle. En adjoignant C au système précédent, on obtient un système complet d'événements. On a donc, avec les notations du théorème :

$$\mathbb{P}(B) = \mathbb{P}(B \cap C) + \sum_{i \in I} \mathbb{P}(B \cap A_i) = \sum_{i \in I} \mathbb{P}(B \cap A_i) = \sum_{i \in I} \mathbb{P}(B \mid A_i)\,\mathbb{P}(A_i),$$

car $B \cap C$ est aussi de probabilité nulle.

Exercice 8

1. On a $p_0 = 1 - \sum\limits_{n=1}^{+\infty} p_n = 1 - \alpha \sum\limits_{p=1}^{+\infty} p^n = 1 - \dfrac{\alpha p}{1-p}\cdot$ On remarque qu'on a donc $\dfrac{\alpha p}{1-p} \leqslant 1$, c'est-à-dire $\alpha \leqslant \dfrac{1}{p} - 1$.
2. On considère, pour $k \in \mathbb{N}$, les événements E_k « la famille a k enfants », F_k « la famille a k filles » et G_k « la famille a k garçons ». On cherche :

$$\mathbb{P}(E_2 \mid F_2) = \frac{\mathbb{P}(F_2 \mid E_2)\,\mathbb{P}(E_2)}{\mathbb{P}(F_2)}\cdot$$

On connait $\mathbb{P}(F_2 \mid E_2) = \dfrac{1}{4}$ et $\mathbb{P}(E_2) = \alpha p^2$. Pour calculer $\mathbb{P}(F_2)$, on applique la formule des probabilités totales avec le système complet d'événements $(E_n)_{n \in \mathbb{N}}$. On obtient :

$$\mathbb{P}(F_2) = \sum_{n=0}^{+\infty} \mathbb{P}(F_2 \mid E_n)\mathbb{P}(E_n).$$

Pour $n \geqslant 2$, on a $\mathbb{P}(F_2 \mid E_n) = \dbinom{n}{2}\left(\dfrac{1}{2}\right)^n$, car il faut choisir 2 filles parmi n enfants et, pour chaque enfant la probabilité d'être une fille ou un garçon est $\dfrac{1}{2}\cdot$ Cette probabilité est évidemment nulle si $n < 2$. On obtient :

$$\begin{aligned}\mathbb{P}(F_2) = \sum_{n=2}^{+\infty} \binom{n}{2}\left(\frac{1}{2}\right)^n \alpha p^n &= \frac{\alpha}{2}\sum_{n=2}^{+\infty} n(n-1)\left(\frac{p}{2}\right)^n \\ &= \frac{\alpha p^2}{8}\sum_{n=2}^{+\infty} n(n-1)\left(\frac{p}{2}\right)^{n-2}.\end{aligned}$$

Par le théorème de dérivation des séries entières, on a, pour tout $x \in \left]-1, 1\right[$:

$$\sum_{n=2}^{+\infty} n(n-1)x^{n-2} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{1}{1-x}\right) = \frac{2}{(1-x)^3}\cdot$$

On en déduit :

$$\mathbb{P}(F_2) = \frac{\alpha p^2}{8} \frac{2}{\left(1-\frac{p}{2}\right)^3} = \frac{2\alpha p^2}{(2-p)^3},$$

puis :

$$\mathbb{P}(E_2 \mid F_2) = \frac{\frac{\alpha p^2}{4}}{\frac{2\alpha p^2}{(2-p)^3}} = \frac{(2-p)^3}{8}\cdot$$

3. En appliquant de nouveau la formule des probabilités totales, on obtient :

$$\mathbb{P}(G_2 \cap F_2) = \sum_{n=0}^{+\infty} \mathbb{P}(F_2 \cap G_2 \mid E_n)\mathbb{P}(E_n) = \mathbb{P}(F_2 \cap G_2 \mid E_4)\mathbb{P}(E_4),$$

car la probabilité conditionnelle est nulle si $n \neq 4$. On a donc :

$$\mathbb{P}(G_2 \cap F_2) = \binom{4}{2}\left(\frac{1}{2}\right)^4 \alpha p^4 = 6\alpha\left(\frac{p}{2}\right)^4.$$

On en déduit :

$$\mathbb{P}(G_2 \mid F_2) = \frac{\mathbb{P}(G_2 \cap F_2)}{\mathbb{P}(F_2)} = \frac{6\alpha\left(\frac{p}{2}\right)^4}{\frac{2\alpha p^2}{(2-p)^3}} = \frac{3p^2(2-p)^3}{16}\cdot$$

Proposition 15 Comme A et B sont indépendants, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
De l'égalité $\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$, on déduit :

$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\big(1 - \mathbb{P}(B)\big) = \mathbb{P}(A)\mathbb{P}(\overline{B}),$$

ce qui montre que A et $\overline{B}$ sont indépendants.
Comme B et A sont indépendants, le raisonnement précédent montre que B et $\overline{A}$ sont indépendants. Enfin, en appliquant ce qui précède à $\overline{A}$ et B, on obtient l'indépendance de $\overline{A}$ et $\overline{B}$.

Exercice 9

- Supposons que $(A_i)_{i\in I}$ soit une famille d'événements indépendants. Soit J une partie finie de I. Si $J' \subset J$, alors J' est une partie finie de I et, par définition de l'indépendance de la famille $(A_i)_{i\in I}$, on a :

$$\mathbb{P}\Big(\bigcap_{i\in J'} A_i\Big) = \prod_{i\in J'} \mathbb{P}(A_i),$$

donc $(A_i)_{i\in J}$ est une famille d'événements indépendants.

- Supposons que, pour toute partie finie J de I, $(A_i)_{i\in J}$ soit une famille d'événements mutuellement indépendants. On a en particulier, pour pour toute partie finie J de I :

$$\mathbb{P}\Big(\bigcap_{i\in J} A_i\Big) = \prod_{i\in J} \mathbb{P}(A_i),$$

donc $(A_i)_{i\in I}$ est une famille d'événements indépendants.

Proposition 16 Soit J une partie finie et non vide de I. Alors $(A_i)_{i\in J}$ est une famille d'événements mutuellement indépendants. Montrons que $(B_i)_{i\in J}$ est également une famille d'événements mutuellement indépendants.

- On étudie d'abord le cas où il existe $k \in J$ tel que $B_k = \overline{A}_k$ et $B_i = A_i$ si $i \in J\backslash\{k\}$. Soit K une partie non vide de J.
  - Si $k \notin K$, alors $\mathbb{P}(\bigcap_{i\in K} B_i) = \mathbb{P}(\bigcap_{i\in K} A_i) = \prod_{i\in K} \mathbb{P}(A_i) = \prod_{i\in K} \mathbb{P}(B_i)$.
  - Si $k \in K$, on note que :
  $$\left(\bigcap_{i\in K} A_i\right) \cup \left(\bigcap_{i\in K} B_i\right) = \left(\bigcap_{i\in K\backslash\{k\}} A_i\right) \cap (A_k \cup B_k) = \bigcap_{i\in K\backslash\{k\}} A_i$$
  et comme cette union est disjointe;
  $$\mathbb{P}\left(\bigcap_{i\in K} B_i\right) = \mathbb{P}\left(\bigcap_{i\in K\backslash\{k\}} A_i\right) - \mathbb{P}\left(\bigcap_{i\in K} A_i\right).$$
  Comme $(A_i)_{i\in J}$ est une famille d'événements mutuellement indépendants, on a :
  $$\begin{aligned}\mathbb{P}\left(\bigcap_{i\in K} B_i\right) &= \prod_{i\in K\backslash\{k\}} \mathbb{P}(A_i) - \prod_{i\in K} \mathbb{P}(A_i) = \big(1 - \mathbb{P}(A_k)\big) \prod_{i\in K\backslash\{k\}} \mathbb{P}(A_i) \\ &= \mathbb{P}(\overline{A}_k) \prod_{i\in K\backslash\{k\}} \mathbb{P}(A_i) = \prod_{i\in K} \mathbb{P}(B_i).\end{aligned}$$

  On a donc démontré, dans ce cas, que $(B_i)_{i\in J}$ est une famille d'événements mutuellement indépendants.
- On passe au cas général. En utilisant ce premier résultat autant de fois qu'il y a de contraires dans les événements B_i pour $i \in J$, nous pouvons en déduire que $(B_i)_{i\in J}$ est une famille d'événements mutuellement indépendants.

On a donc en particulier :
$$\mathbb{P}\left(\bigcap_{i\in J} B_i\right) = \prod_{i\in J} \mathbb{P}(B_i).$$
Comme cela est vrai pour toute partie finie non vide J de I, on en déduit que $(B_i)_{i\in I}$ est une famille d'événements mutuellement indépendants.

Exercice 10

1. Soit $J \subset [\![1, n]\!]$. On a, par indépendance de la famille $(A_i)_{i\in I}$:
$$\begin{aligned}\mathbb{P}\left(\bigcap_{j\in J} B_j\right) = \mathbb{P}\left(\bigcap_{i\in \bigcup_{j\in J} I_j} A_i\right) &= \prod_{i\in \bigcup_{j\in J} I_j} \mathbb{P}(A_i) \\ &= \prod_{j\in J}\left(\prod_{i\in I_j} \mathbb{P}(A_i)\right) = \prod_{j\in J} \mathbb{P}(B_j).\end{aligned}$$
Donc $(B_1, B_1, \ldots, B_n)$ est une famille d'événements indépendants.

2. Posons $C_j = B_j$ si $B_j = \bigcap\limits_{i \in I_j} A_i$ et $C_j = \overline{B_j} = \bigcap\limits_{i \in I_j} \overline{A_i}$ si $B_j = \bigcup\limits_{i \in I_j} A_i$. Alors l'indépendance de $(C_1, C_2, \ldots, C_n)$ résulte de la proposition 16 et de la première question de l'exercice. En appliquant de nouveau la proposition 16, on en déduit que $(B_1, B_1, \ldots, B_n)$ est une famille d'événements indépendants.

Exercice 11 Soit $\omega = (\omega_n)_{n \in \mathbb{N}^*} \in \Omega$. Pour tout $n \in \mathbb{N}^*$, on pose $B_n = A_n$ si $\omega_n = 1$ et $B_n = \overline{A_n}$ sinon. On a alors :

$$\{\omega\} = \bigcap_{n \in \mathbb{N}^*} B_n.$$

Les A_n, pour $n \in \mathbb{N}^*$, sont des événements, donc les B_n sont également des événements, et $\{\omega\}$ qui est l'intersection d'un nombre dénombrable d'événements est un événement.

Comme $(A_n)_{n \in \mathbb{N}^*}$ est une famille d'événements indépendants, il en est de même de la famille $(B_n)_{n \in \mathbb{N}^*}$ d'après la proposition 16 de la page 859. On a donc, d'après le corollaire 5 de la page 852 :

$$\mathbb{P}(\{\omega\}) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}^*} B_n\right) = \lim_{N \to +\infty} \mathbb{P}\left(\bigcap_{n=1}^{N} B_n\right) = \lim_{N \to +\infty} \prod_{n=1}^{N} \mathbb{P}(B_n).$$

Pour tout $n \in \mathbb{N}^*$, on a $\mathbb{P}(B_n) = p$ ou $1 - p$. On en déduit que, pour tout $N \in \mathbb{N}^*$, on a $\prod\limits_{n=1}^{N} \mathbb{P}(B_n) \leqslant \big(\max(p, 1-p)\big)^N$. Par passage à la limite, on obtient $\mathbb{P}(\{\omega\}) = 0$.

Exercice 12

1. On pose $B = \bigcap\limits_{k=0}^{+\infty} \bigcup\limits_{p=k}^{+\infty} A_p$. On a donc $\overline{B} = \bigcup\limits_{k=0}^{+\infty} \bigcap\limits_{p=k}^{+\infty} \overline{A_p}$.

   Les événements $\overline{A_n}$ sont aussi mutuellement indépendants donc, pour tout $k \in \mathbb{N}$:

$$\mathbb{P}\left(\bigcap_{p=k}^{+\infty} \overline{A_p}\right) = \lim_{N \to +\infty} \mathbb{P}\left(\bigcap_{p=k}^{N} \overline{A_p}\right) = \lim_{N \to +\infty} \prod_{p=k}^{N} \mathbb{P}(\overline{A_p})$$
$$= \lim_{N \to +\infty} \prod_{p=k}^{N} \big(1 - \mathbb{P}(A_p)\big).$$

   Sachant que $0 \leqslant 1 - x \leqslant e^{-x}$ pour tout réel $x \in [0, 1]$, on a :

$$0 \leqslant \prod_{p=k}^{N} (1 - \mathbb{P}(A_p)) \leqslant \prod_{p=k}^{N} \exp\big(-\mathbb{P}(A_p)\big) \leqslant \exp\left(-\sum_{p=k}^{N} \mathbb{P}(A_p)\right).$$

   La série $\sum \mathbb{P}(A_p)$ diverge donc $\lim\limits_{N \to +\infty} \sum\limits_{p=k}^{N} \mathbb{P}(A_p) = +\infty$. On en déduit :

$$\mathbb{P}\left(\bigcap_{p=k}^{+\infty} \overline{A_p}\right) = \lim_{N \to +\infty} \prod_{p=k}^{N} \big(1 - \mathbb{P}(A_p)\big) = 0.$$

Ainsi, $\overline{B}$ est négligeable, car réunion dénombrable d'événements négligeables, et l'événement contraire B est presque sûr.

2. Pour tout $n \in \mathbb{N}^*$, on note π_n l'événement « le n-ième lancer donne pile » et A_n l'événement $\pi_{nm+1} \cap \pi_{nm+2} \cap \ldots \cap \pi_{nm+m}$, c'est-à-dire « les lancers entre les rangs $nm+1$ et $nm+m$ donnent pile ». Pour tout $n \in \mathbb{N}$, les événements π_1, π_2, $\ldots$, π_{nm+m} sont indépendants. Il résulte de l'exercice 10 de la page 859 que les événements A_1, A_2, $\ldots$, A_n sont indépendants. La suite $(A_n)_{n\in\mathbb{N}^*}$ est donc constituée d'événements indépendants.

   De plus, pour tout $n \in \mathbb{N}^*$, on a, par indépendance des événements π_i :
   $$\mathbb{P}(A_n) = \prod_{i=nm+1}^{nm+m} \mathbb{P}(\pi_i) = p^m.$$
   La série $\sum \mathbb{P}(A_n)$ est donc grossièrement divergente. En gardant les notations des questions précédentes, on en déduit que $\mathbb{P}(B) = 1$. Il est donc presque sûr qu'une infinité d'événements A_n se réalisent. Comme l'événement C « il apparaît une infinité de séquences de m piles consécutifs » contient l'événement B, il est *a fortiori* presque sûr.

S'entraîner et approfondir

15.1 (Probabilité image)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et f une application de Ω dans un ensemble Ω'.

1. Montrer que $\mathcal{A}' = \{B \in \mathcal{P}(\Omega') \mid f^{-1}(B) \in \mathcal{A}\}$ est une tribu sur Ω'. On l'appelle **tribu image** de $\mathcal{A}$ par f.
2. Monter que $\mathbb{P}_f : B \in \mathcal{A}' \mapsto \mathbb{P}\big(f^{-1}(B)\big)$ est une probabilité sur $(\Omega', \mathcal{A}')$. On l'appelle **probabilité image** de $\mathbb{P}$ par f.

15.2 Soit A_1, $A_2, \ldots, A_n$ des événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Montrer que :

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \min_{1 \leqslant k \leqslant n} \left(\sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{\substack{1 \leqslant i \leqslant n \\ i \neq k}} \mathbb{P}(A_i \cap A_k)\right).$$

15.3 On souhaite démontrer qu'il n'existe pas de probabilité $\mathbb{P}$ sur l'espace probabilisé $(\mathbb{N}^*, \mathcal{P}(\mathbb{N}^*))$ telle que, pour tout entier $n \geqslant 1$, la probabilité de l'ensemble A_n des multiples de n soit $\dfrac{1}{n}\cdot$

On raisonne par l'absurde et on suppose qu'une telle probabilité existe.

1. Montrer que $(A_p)_{p \in \mathcal{P}}$, où $\mathcal{P}$ désigne l'ensemble des nombres premiers, est une famille d'événements indépendants.
2. Soit $(p_n)_{n \in \mathbb{N}^*}$ la suite croissante des nombres premiers.

   On rappelle que la série $\sum \dfrac{1}{p_n}$ diverge (*cf.* par exemple l'exercice 9.15 de la page 479).

   Soit A l'ensemble des entiers naturels ayant une infinité de diviseurs premiers. Exprimer A en fonction des événements A_{p_i} (pour $i \in \mathbb{N}^*$). En déduire la probabilité de A.
3. Conclure.

15.4 Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, mutuellement indépendants.

1. Montrer que :

$$\mathbb{P}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = 1 - \lim_{n \to +\infty} \prod_{i=0}^{n} \mathbb{P}\left(\overline{A_i}\right).$$

2. On suppose que, pour tout $n \in \mathbb{N}$, $\mathbb{P}(A_n) \neq 1$.

   Montrer que $\mathbb{P}\left(\bigcup\limits_{n \in \mathbb{N}} A_n\right) = 1$ si, et seulement si, la série de terme général $\mathbb{P}(A_n)$ diverge.
3. On suppose que, pour tout $n \in \mathbb{N}$, $\mathbb{P}(A_n) = \dfrac{1}{(n+2)^2}\cdot$

   (a) Calculer $\mathbb{P}\left(\bigcup\limits_{n \in \mathbb{N}} A_n\right)$.

   (b) Calculer la probabilité qu'un seul des événements A_n se réalise.

15.5 Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On pose, pour tout $n \in \mathbb{N}$:

$$B_n = \bigcup_{k\geqslant n} A_k \text{ et } C_n = \bigcap_{k\geqslant n} A_k, \quad \text{puis } B = \bigcap_{n\in\mathbb{N}} B_n \text{ et } C = \bigcup_{n\in\mathbb{N}} C_n.$$

1. On dit que la suite $(A_n)_{n\in\mathbb{N}}$ converge si $B = C$ et, par définition, sa limite A est alors $A = B = C$.
   Montrer que si la suite $(A_n)_{n\in\mathbb{N}}$ converge, alors A est un événement et la suite $(\mathbb{P}(A_n))_{n\in\mathbb{N}}$ converge vers $\mathbb{P}(A)$.
2. On suppose que, pour tout $n \in \mathbb{N}$, les événements B_n et C_n sont indépendants.
   (a) Montrer que B et C sont indépendants.
   (b) Montrer que si, de plus, la suite (A_n) converge vers A, alors $\mathbb{P}(A) \in \{0, 1\}$.

⋆⋆ **15.6** *(ENS Ulm 2015)*

Soit A_1, A_2,..., A_n des événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Pour $k \in [\![1, n]\!]$, on considère l'événement C_k « appartenir à A_i pour au moins k valeurs de l'indice i entre 1 et n ». Montrer que :

$$\prod_{k=1}^{n} \mathbb{P}(C_k) \leqslant \prod_{k=1}^{n} \mathbb{P}(A_k).$$

⋆ **15.7** Un avion comporte n sièges ($n \geqslant 2$). Chacun des n passagers a une place qui est réservée.

- Le premier passager arrive. Il est distrait et s'installe à une place choisie au hasard.
- Les passagers suivants quand ils arrivent s'installent à leur place sauf si celle-ci est déjà occupée, auquel cas ils choisissent une place au hasard parmi les places restantes.

Déterminer la probabilité que le n-ième passager s'installe à sa place.

Indication : on pourra raisonner par récurrence sur n.

⋆ **15.8** Soit b et n deux entiers naturels non nuls. Une urne contient b boules blanches et n boules noires. On tire au hasard et sans remise les boules jusqu'à ce qu'on obtienne pour la première fois une boule d'une couleur différente des précédentes. Celle-ci est alors remise dans l'urne. On reprend alors la procédure au début. On continue ainsi jusqu'au tirage de la dernière boule (si à la fin, il ne reste plus que des boules d'une seule couleur, on les tire toutes).

Montrer que la probabilité que la dernière boule tirée soit blanche est $\frac{1}{2}$.

Indication : on pourra raisonner par récurrence forte sur $n + b$.

15.9 Un fumeur a dans ses poches deux boîtes contenant respectivement n et $n + 1$ allumettes. Chaque fois qu'il a besoin d'une allumette, il prend au hasard dans l'une des boîtes.

Quelle est la probabilité pour que, lorsqu'il prend la dernière allumette d'une boîte, l'autre boîte contienne encore k allumettes ?

15.10 (Le tournoi)

Des joueurs notés $j_1, j_2, \dots$ (il y a une infinité de joueurs) s'affrontent à pile ou face, avec une pièce honnête, de la façon suivante : j_1 et j_2 commencent, le perdant est éliminé et le gagnant rencontre j_3, le perdant est éliminé et le gagnant rencontre j_4, ... Est déclaré vainqueur le joueur qui gagne trois parties consécutives et le jeu s'arrête alors.

Pour $n \geqslant 1$, on note p_n la probabilité que j_n gagne le tournoi et q_n la probabilité qu'il joue.

1. Montrer que $p_n = \dfrac{1}{8}q_n$.
2. (a) Que vaut q_n pour $1 \leqslant n \leqslant 4$.
   (b) Montrer que, pour $n \geqslant 5$, $q_n = \dfrac{1}{2}q_{n-1} + \dfrac{1}{4}q_{n-2}$.
3. Calculer p_n pour tout n.

15.11 Un gardien doit ouvrir un porte ; il possède un trousseau de six clés différentes. Quand il est ivre, il remélange les clés après chaque essai ; sinon, il retire la mauvaise clé du lot. Sachant qu'il est ivre un jour sur quatre et que ce soir-là il a essayé au moins cinq clés, quelle est la probabilité qu'il soit ivre ?

15.12 Plusieurs joueurs, nommés j_1, j_2, ..., j_k, ... jouent l'un après l'autre, dans l'ordre des indices. À chaque joueur j_k (pour $k \geqslant 1$) est imparti un événement précis A_k, de probabilité $p_k \in \,]0,1[$. Le jeu se termine lorsque l'un des joueurs réalise l'événement qui lui est imparti.

On pose $q_k = 1 - p_k$. On suppose l'indépendance mutuelle de toutes les suites de résultats des coups joués par les concurrents.

L'événement « j_k gagne la partie » sera noté G_k.

1. On suppose que le nombre de joueurs est fini, égal à un entier $c \geqslant 2$. Lorsqu'aucun joueur n'a gagné, on recommence un tour à partir du joueur j_1 et on continue ainsi jusqu'à ce que l'un des joueurs gagne.
   (a) Calculer $\mathbb{P}(G_k)$ pour $k \in [\![1, c]\!]$.
   (b) Démontrer que la partie se termine presque sûrement au bout d'un nombre fini de coups.
   (c) On dit que le jeu est équitable si $P(G_1) = \ldots = P(G_c)$. Montrer que le jeu est équitable si et seulement si, pour tout $k \in [\![1, c-1]\!]$, on a :
   $$p_{k+1} = \frac{p_k}{1 - p_k}.$$
   (d) En supposant le jeu équitable, exprimer p_k en fonction de p_1.
2. On garde les mêmes notations. On suppose qu'il y a une infinité de joueurs. Ainsi un joueur qui n'a pas gagné à son tour est définitivement éliminé.
   (a) Montrer que le jeu ne peut pas être équitable.
   (b) Montrer que la suite de terme général $Q_n = q_1 \dots q_n$ converge. On note a sa limite.
   Montrer que la probabilité que le jeu se termine en un nombre fini de coups est $1 - a$.

15.13 Soit $p \in]0,1[$. Une personne joue à pile ou face, la probabilité d'obtenir pile étant p. Elle gagne dès qu'elle a obtenu deux piles de plus que de faces. Elle perd dès qu'elle a obtenu deux faces de plus que de piles.

1. Pour $n \in \mathbb{N}^*$, quelle est la probabilité que la partie dure plus de $2n$ coups.
2. Quelle est la probabilité que la personne gagne ? La partie peut-elle durer indéfiniment ?

15.14 On considère une suite infinie de lancers d'une pièce de monnaie, numérotés par $n \in \mathbb{N}^*$, la probabilité d'obtenir pile (noté P) étant $p \in]0,1[$ et la probabilité d'obtenir face (noté F) étant $q = 1 - p$.

1. Calculer la probabilité de l'événement « la première séquence PP apparaît avant la première séquence FP ».
2. Pour tout $n \geqslant 2$, calculer la probabilité x_n de l'événement « la séquence PF apparaît pour la première fois aux lancers numérotés $n-1$ et n et il n'y a pas eu avant de séquence FP ». En déduire la probabilité x de l'événement « la première séquence PF apparaît avant la première séquence FP ».
3. Déterminer de même la probabilité des événements :
   (a) « la première séquence PP apparaît avant la première séquence PF » ;
   (b) « la première séquence PF apparaît avant la première séquence PP » ;
   (c) « la première séquence PF apparaît avant la première séquence FF » ;
   (d) « la première séquence PP apparaît avant la première séquence FF ».

15.15 **(La ruine des joueurs)**

Deux joueurs A et B ayant pour capitaux initiaux respectivement a et b euros ($(a,b) \in \mathbb{N}^2$) s'affrontent dans un jeu constitué d'une succession de parties indépendantes. À chaque partie, A a la probabilité $p \in]0,1[$ de gagner et B la probabilité $q = 1 - p$. À l'issue de chaque partie, le perdant donne 1 euro au gagnant. Le jeu s'arrête lorsque l'un des joueurs est ruiné. On note $R(a,b)$ la probabilité que A finisse ruiné.

1. Montrer que $R(a,b) = pR(a+1,b-1) + qR(a-1,b+1)$.
2. Calculer $R(a,b)$. On pourra remarquer que la capital total $a+b$ reste fixe au cours de la partie.
3. La partie peut-elle durer indéfiniment ?
4. Calculer $\lim\limits_{b \to +\infty} R(a,b)$. Commenter.

15.16 Un banquier se rend chaque jour de son domicile à sa banque, puis de sa banque à son domicile. Il possède un unique parapluie. À chaque fois qu'il part, s'il pleut et si le parapluie est à sa disposition, alors il le prend.

On suppose que la probabilité qu'il pleuve vaut constamment $p \in]0,1[$ (on pose $q = 1 - p$). Pour $n \in \mathbb{N}^*$ note p_n la probabilité que le parapluie soit disponible là où se trouve le banquier (au domicile ou au bureau) au bout de n trajets et $q_n = 1 - p_n$. La probabilité qu'il soit disponible initialement est p_0, quelconque.

1. Montrer qu'il existe $S \in \mathcal{M}_2(\mathbb{R})$ tel que, pour tout $n \in \mathbb{N}$, on ait $\begin{pmatrix} p_{n+1} \\ q_{n+1} \end{pmatrix} = S \begin{pmatrix} p_n \\ q_n \end{pmatrix}$.
2. En déduire que la suite $(p_n)_{n \in \mathbb{N}}$ converge. Calculer $p_\infty = \lim\limits_{n \to +\infty} p_n$.

Solution des exercices

15.1 1. • On a $f^{-1}(\Omega') = \Omega \in \mathcal{A}$ donc $\Omega' \in \mathcal{A}'$.

• Si $B \in \mathcal{A}'$, alors $f^{-1}(B) \in \mathcal{A}$. On écrit $f^{-1}(\Omega' \setminus B) = \Omega \setminus f^{-1}(B)$. De $f^{-1}(B) \in \mathcal{A}$, on déduit $\Omega \setminus f^{-1}(B) \in \mathcal{A}$, car $\mathcal{A}$ est stable par passage au complémentaire, et donc $\Omega' \setminus B \in \mathcal{A}'$.

• Enfin si $(B_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{A}'$ alors, pour tout $n \in \mathbb{N}$, on a $f^{-1}(B_n) \in \mathcal{A}$. On en déduit :

$$f^{-1}\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \bigcup_{n\in\mathbb{N}} f^{-1}(B_n) \in \mathcal{A}$$

et donc $\bigcup\limits_{n\in\mathbb{N}} B_n \in \mathcal{A}'$.

On conclut que $\mathcal{A}'$ est une tribu sur Ω'.

2. On note que $\mathbb{P}_f$ est bien une application de $\mathcal{A}'$ dans $[0,1]$.

• On a $\mathbb{P}_f(\Omega') = \mathbb{P}\Big(f^{-1}(\Omega')\Big) = \mathbb{P}(\Omega) = 1$.

• Si $(B_n)_{n\in\mathbb{N}}$ est une suite d'éléments disjoints de $\mathcal{A}'$, les ensembles $f^{-1}(B_n)$ sont de éléments disjoints de $\mathcal{A}$ et on a donc :

$$\begin{aligned}\mathbb{P}_f\left(\bigcup_{n\in\mathbb{N}} B_n\right) &= \mathbb{P}\left(f^{-1}\left(\bigcup_{n\in\mathbb{N}} B_n\right)\right) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} f^{-1}(B_n)\right)\\ &= \sum_{n=0}^{+\infty} \mathbb{P}\left(f^{-1}(B_n)\right) = \sum_{n=0}^{+\infty} \mathbb{P}_f(B_n).\end{aligned}$$

Donc $\mathbb{P}_f$ est une probabilité sur $(\Omega', \mathcal{A}')$.

15.2 Il faut démontrer que, pour tout $k \in [\![1,n]\!]$, on a

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{\substack{1\leqslant i\leqslant n\\ i\neq k}} \mathbb{P}(A_i \cap A_k).$$

Comme A_1, ,..., A_n jouent des rôles symétriques, on peut supposer $k = n$. On transforme le second membre de l'inégalité :

$$\begin{aligned}\sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i=1}^{n-1} \mathbb{P}(A_i \cap A_n) &= \mathbb{P}(A_n) + \sum_{i=1}^{n-1} \left(\mathbb{P}(A_i) - \mathbb{P}(A_i \cap A_n)\right)\\ &= \mathbb{P}(A_n) + \sum_{i=1}^{n-1} \mathbb{P}(A_i \setminus A_n).\end{aligned}$$

En utilisant l'inégalité de Boole, on obtient :

$$\mathbb{P}(A_n) + \sum_{i=1}^{n-1} \mathbb{P}(A_i \setminus A_n) \geqslant \mathbb{P}\left(A_n \cup \left(\bigcup_{i=1}^{n-1} (A_i \setminus A_n)\right)\right) = \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right),$$

ce qui donne l'inégalité voulue.

15.3 1. Soit $p_1, p_2, \ldots, p_r$ des nombres premiers distincts. Ces entiers sont premiers entre eux, donc un entier est divisible par chaque p_i $(1 \leqslant i \leqslant r)$ si, et seulement s'il est divisible par $\prod\limits_{i=1}^{r} p_i$. On a donc $\bigcap\limits_{i\in[\![1,r]\!]} A_{p_i} = A_{\prod\limits_{i=1}^{r} p_i}$. On en déduit :

$$\mathbb{P}\Big(\bigcap_{i\in[\![1,r]\!]} A_{p_i}\Big) = \mathbb{P}\left(A_{\prod\limits_{i=1}^{r} p_i}\right) = \frac{1}{\prod\limits_{i=1}^{r} p_i} = \prod_{i=1}^{r} \frac{1}{p_i} = \prod_{i=1}^{r} \mathbb{P}(A_{p_i}).$$

Donc la famille $(A_p)_{p\in\mathcal{P}}$ est constituée d'événements indépendants.

2. On a $A = \bigcap\limits_{r\geqslant 1} \bigcup\limits_{i\geqslant r} A_{p_i}$ (*cf.* exercice 3 de la page 849) et donc $\overline{A} = \bigcup\limits_{r\geqslant 1} \bigcap\limits_{i\geqslant r} \overline{A_{p_i}}$.
Soit $r \in \mathbb{N}^*$. Comme $(A_p)_{p\in\mathcal{P}}$ est une famille d'événements indépendants, il en est de même de $\left(\overline{A_p}\right)_{p\in\mathcal{P}}$. On a donc, pour $N \geqslant r$:

$$\mathbb{P}\Big(\bigcap_{r\leqslant i\leqslant N} \overline{A_{p_i}}\Big) = \prod_{r\leqslant i\leqslant N} \left(1 - \frac{1}{p_i}\right) \leqslant \exp\left(-\sum_{i=r}^{N} \frac{1}{p_i}\right).$$

Puisque la série $\sum \frac{1}{p_n}$ diverge, on en déduit :

$$\mathbb{P}\left(\bigcap_{i=r}^{+\infty} \overline{A_{p_i}}\right) = \lim_{N\to+\infty} \mathbb{P}\left(\bigcap_{r\leqslant i\leqslant N} \overline{A_{p_i}}\right) \leqslant \lim_{N\to+\infty} \exp\left(-\sum_{i=r}^{N} \frac{1}{p_i}\right) = 0\cdot$$

Ainsi $\overline{A}$ est une réunion dénombrable d'événements négligeables, donc est négligeable. En prenant l'événement contraire, on obtient $\mathbb{P}(A) = 1$.

3. Comme A est de probabilité 1, il n'est pas vide. Il y a donc des entiers ayant une infinité de diviseurs premiers, ce qui donne la contradiction voulue.

15.4 1. On a :

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 1 - \mathbb{P}\left(\overline{\bigcup_{n\in\mathbb{N}} A_n}\right) = 1 - \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} \overline{A_n}\right)$$
$$= 1 - \lim_{n\to+\infty} \mathbb{P}\left(\bigcap_{i=0}^{n} \overline{A_i}\right) = 1 - \lim_{n\to+\infty} \prod_{i=0}^{n} \mathbb{P}\left(\overline{A_i}\right).$$

2. D'après la question précédente, $\mathbb{P}\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) = 1$ équivaut à

$$\lim_{n\to+\infty} \prod_{i=0}^{n} \mathbb{P}\left(\overline{A_i}\right) = 0.$$

Comme aucun terme du produit n'est nul, cela équivaut à

$$\lim_{n\to+\infty} \sum_{i=0}^{n} \ln\left(\mathbb{P}\left(\overline{A_i}\right)\right) = -\infty,$$

c'est-à-dire à la divergence de la série $\sum \ln\left(\mathbb{P}\left(\overline{A_n}\right)\right)$.

Si $\mathbb{P}(A_n)$ ne tend pas vers 0, alors les séries de terme général $\ln\left(\mathbb{P}\left(\overline{A_n}\right)\right)$ et $\mathbb{P}(A_n)$ sont grossièrement divergentes.

Si $\lim\limits_{n\to+\infty} \mathbb{P}(A_n) = 0$, alors $\ln \mathbb{P}\left(\overline{A_n}\right) = \ln(1-\mathbb{P}(A_n)) \sim -\mathbb{P}(A_n)$. Les séries de terme général $\ln \mathbb{P}\left(\overline{A_n}\right)$ et $\mathbb{P}(A_n)$ ont même nature.

Ainsi, on a $\mathbb{P}\left(\bigcup\limits_{n\in\mathbb{N}} A_n\right) = 1$ si, et seulement si, la série de terme général $\mathbb{P}(A_n)$ diverge.

3. La condition $\mathbb{P}(A_n) \neq 1$ pour tout $n \in \mathbb{N}$ est réalisée.

(a) Pour tout $n \in \mathbb{N}$, on a :

$$\prod_{i=0}^{n} \mathbb{P}\left(\overline{A_i}\right) = \prod_{i=0}^{n}\left(1-\frac{1}{(i+2)^2}\right) = \prod_{i=0}^{n}\frac{(i+1)(i+3)}{(i+2)^2}$$
$$= \frac{(n+1)!(n+3)!}{2((n+2)!)^2} = \frac{n+3}{2(n+2)}.$$

On en déduit :

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 1 - \lim_{n\to+\infty}\frac{n+3}{2(n+2)} = \frac{1}{2}.$$

(b) Notons B_n l'événement « seul l'événement A_n se réalise ». On a :

$$\mathbb{P}(B_n) = \mathbb{P}\left(A_n \cap \bigcap_{i\neq n}\overline{A_i}\right) = \lim_{N\to+\infty}\mathbb{P}\left(A_n \cap \bigcap_{\substack{0\leqslant i\leqslant N\\ i\neq n}}\overline{A_i}\right)$$
$$= \lim_{N\to+\infty}\mathbb{P}(A_n)\prod_{\substack{0\leqslant i\leqslant N\\ i\neq n}}\mathbb{P}(\overline{A_i}).$$

Pour $N > n$, on a, d'après le calcul mené dans la question précédente :

$$\mathbb{P}(A_n)\prod_{\substack{0\leqslant i\leqslant N\\ i\neq n}}\mathbb{P}\left(\overline{A_i}\right) = \frac{\mathbb{P}(A_n)}{\mathbb{P}(\overline{A_n})}\prod_{0\leqslant i\leqslant N}\mathbb{P}(\overline{A_i})$$
$$= \frac{1}{(n+1)(n+3)} \times \frac{N+3}{2(N+2)}.$$

On en déduit : $\mathbb{P}(B_n) = \dfrac{1}{2(n+1)(n+3)}$.

L'événement B « un seul événement A_n se réalise » est $\bigcup\limits_{n\in\mathbb{N}} B_n$.

C'est une réunion dénombrable d'événements incompatibles donc :

$$\mathbb{P}(B) = \sum_{i=0}^{+\infty}\frac{1}{2(n+1)(n+3)} = \frac{1}{4}\sum_{i=0}^{+\infty}\left(\frac{1}{n+1}-\frac{1}{n+3}\right)$$
$$= \frac{1}{4}\lim_{N\to+\infty}\left(1+\frac{1}{2}-\frac{1}{N+1}-\frac{1}{N+2}\right) = \frac{3}{8}.$$

15.5 1. Pour tout $n \in \mathbb{N}$, B_n et C_n sont des événements donc B et C sont des événements (car une réunion ou une intersection d'un nombre dénombrable d'événements est un événement). Comme $A = B = C$, on en déduit que A est un événement.

La suite (B_n) (resp. (C_n)) est décroissante (resp. croissante) donc :

$$\mathbb{P}(A) = \mathbb{P}(B) = \lim_{n\to+\infty} \mathbb{P}(B_n) \quad \text{et} \quad \mathbb{P}(A) = \mathbb{P}(C) = \lim_{n\to+\infty} \mathbb{P}(C_n).$$

On a de plus, $C_n \subset A_n \subset B_n$ et donc $\mathbb{P}(C_n) \leqslant \mathbb{P}(A_n) \leqslant \mathbb{P}(B_n)$ donc, par encadrement :

$$\lim_{n\to+\infty} \mathbb{P}(A_n) = \mathbb{P}(A).$$

2. (a) Les suites $(B_n)_{n\in\mathbb{N}}$ et $(\overline{C_n})_{n\in\mathbb{N}}$ sont décroissantes. Il en est donc de même de la suite $\left(B_n \cap \overline{C_n}\right)_{n\in\mathbb{N}}$. Comme :

$$B \cap \overline{C} = \left(\bigcap_{n\in\mathbb{N}} B_n\right) \bigcap \left(\bigcap_{n\in\mathbb{N}} \overline{C_n}\right) = \bigcap_{n\in\mathbb{N}} \left(B_n \cap \overline{C_n}\right),$$

on a, par indépendance des événements B_n et $\overline{C_n}$:

$$\begin{aligned}\mathbb{P}(B \cap \overline{C}) &= \lim_{n\to+\infty} \mathbb{P}\left(B_n \cap \overline{C_n}\right) = \lim_{n\to+\infty} \mathbb{P}\left(B_n\right) \mathbb{P}\left(\overline{C_n}\right)\\ &= \lim_{n\to+\infty} \mathbb{P}\left(B_n\right) \lim_{n\to+\infty} \mathbb{P}\left(\overline{C_n}\right) = \mathbb{P}(B)\,\mathbb{P}\left(\overline{C}\right).\end{aligned}$$

Les événements B et $\overline{C}$ sont donc indépendants. On sait qu'alors il en est de même de B et C.

(b) On a $B = C = A$. On en déduit :

$$0 = \mathbb{P}(A \cap \overline{A}) = \mathbb{P}(A)\,\mathbb{P}\left(\overline{A}\right),$$

donc $\mathbb{P}(A) \in \{0, 1\}$.

15.6 On remarque pour commencer que, pour tout $k \in [\![1, n]\!]$, on a :

$$C_k = \bigcup_{1\leqslant i_1<i_2<\ldots<i_k\leqslant n} A_{i_1} \cap A_{i_2} \cap \ldots A_{i_k}.$$

On démontre la propriété par récurrence sur n.

- La propriété est évidente pour $n = 1$. On la démontre pour $n = 2$, ce qui sera utile pour passer du rang n au rang $n+1$. Il s'agit de démontrer :

  $$\mathbb{P}(A_1 \cup A_2)\,\mathbb{P}(A_1 \cap A_2) \leqslant \mathbb{P}(A_1)\mathbb{P}(A_2),$$

  ce qui équivaut à :

  $$\big(\mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)\big)\,\mathbb{P}(A_1 \cap A_2) \leqslant \mathbb{P}(A_1)\mathbb{P}(A_2).$$

  En posant $x = \mathbb{P}(A_1)$, $y = \mathbb{P}(A_2)$ et $z = \mathbb{P}(A_1 \cap A_2)$, il faut démontrer que $xy - (x+y-z)z \geqslant 0$, c'est-à-dire $(x-z)(y-z) \geqslant 0$. Cette inégalité est vérifiée car $z \leqslant x$ et $z \leqslant y$, ce qui résulte de $A_1 \cap A_2 \subset A_1$ et $A_1 \cap A_2 \subset A_2$.

- On suppose que la propriété est vérifiée au rang n. On la démontre au rang $n+1$. On pose, pour tout $k \in [\![1, n+1]\!]$:

$$C'_k = \bigcup_{1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n+1} A_{i_1} \cap A_{i_2} \cap \ldots A_{i_k}.$$

Il faut donc démontrer que :

$$\prod_{k=1}^{n+1} \mathbb{P}(C'_k) \leqslant \prod_{k=1}^{n+1} \mathbb{P}(A_k).$$

On suppose que $\mathbb{P}(C'_k) \neq 0$ pour tout $k \in [\![1, n+1]\!]$, sinon l'inégalité est évidente. On remarque que, pour $1 \leqslant k \leqslant n-1$, on a :

$$C'_{k+1} = (C_k \cap A_{n+1}) \cup C_{k+1}.$$

En appliquant la propriété au rang 2, on obtient :

$$\mathbb{P}(C'_{k+1})\mathbb{P}\big((C_k \cap A_{n+1}) \cap C_{k+1}\big) \leqslant \mathbb{P}(C_k \cap A_{n+1})\mathbb{P}(C_{k+1}).$$

Il est clair que $C_{k+1} \subset C_k$. On a donc $\mathbb{P}\big((C_k \cap A_{n+1}) \cap C_{k+1}\big) = \mathbb{P}(C_{k+1} \cap A_{n+1})$. On obtient, pour tout $k \in [\![1, n-1]\!]$:

$$\mathbb{P}(C'_{k+1})\,\mathbb{P}(C_{k+1} \cap A_{n+1}) \leqslant \mathbb{P}(C_k \cap A_{n+1})\mathbb{P}(C_{k+1}).$$

En multipliant ces inégalités pour $1 \leqslant k \leqslant n-1$, on obtient :

$$\prod_{k=2}^{n} \mathbb{P}(C'_k) \prod_{k=2}^{n} \mathbb{P}(C_k \cap A_{n+1}) \leqslant \prod_{k=1}^{n-1} \mathbb{P}(C_k \cap A_{n+1}) \prod_{k=2}^{n} \mathbb{P}(C_k).$$

La suite $(\mathbb{P}(C_k \cap A_{n+1}))_{1 \leqslant k \leqslant n}$ décroît donc s'il existe $k \in [\![2, n]\!]$ tel que l'on ait $\mathbb{P}(C_k \cap A_{n+1}) = 0$, alors $\mathbb{P}(C_n \cap A_{n+1}) = 0$, c'est-à-dire $\mathbb{P}(C'_{n+1}) = 0$, ce qui contredit notre hypothèse. On obtient donc en simplifiant :

$$\prod_{k=2}^{n} \mathbb{P}(C'_k) \times \mathbb{P}(C_n \cap A_{n+1}) \leqslant \mathbb{P}(C_1 \cap A_{n+1}) \times \prod_{k=2}^{n} \mathbb{P}(C_k).$$

Comme $C'_{n+1} = C_n \cap A_{n+1}$ et $C'_1 = C_1 \cup A_{n+1}$, on a :

$$\prod_{k=1}^{n+1} \mathbb{P}(C'_k) \leqslant \mathbb{P}(C_1 \cup A_{n+1})\,\mathbb{P}(C_1 \cap A_{n+1}) \times \prod_{k=2}^{n} \mathbb{P}(C_k).$$

En utilisant de nouveau la propriété au rang 2, puis l'hypothèse de récurrence, on obtient :

$$\prod_{k=1}^{n+1} \mathbb{P}(C'_k) \leqslant \mathbb{P}(A_{n+1}) \prod_{k=1}^{n} \mathbb{P}(C_k) \leqslant \prod_{k=1}^{n+1} \mathbb{P}(A_k).$$

15.7 Notons p_n la probabilité cherchée. Soit $n \geqslant 2$. On considère un avion à n sièges. On numérote les passagers par ordre d'arrivée, et on attribue le numéro k à la place du passager numéro k. On considère les événements G : « le n-ième passager s'installe à sa place » et A_k :« le premier passager s'installe à la k-ième place », pour tout $1 \leqslant k \leqslant n$. On applique la formule des probabilités totales. On obtient

$$p_n = \mathbb{P}(G) = \sum_{k=1}^{n} \mathbb{P}(A_k)\,\mathbb{P}(G \mid A_k).$$

Le premier passager choisit au hasard donc, pour $1 \leqslant k \leqslant n$, on a $\mathbb{P}(A_k) = \dfrac{1}{n}\cdot$

- Si le premier passager s'assoit à sa place, tous les autres en font de même ; on a donc $\mathbb{P}(G \mid A_1) = 1$.
- Si le premier passager s'assoit la n-ième place, le n-ième passager ne pourra pas s'assoir à sa place ; on a donc $\mathbb{P}(G \mid A_n) = 0$.
- Si le premier passager s'assoit à la k-ième place, pour $2 \leqslant k \leqslant n-1$, les passagers jusqu'au $(k-1)$-ième s'installent à leur place. Il reste ensuite $n-k+1$ passagers (numérotés de k à n) et $n-k+1$ sièges (le siège numéro 1 et les sièges du $(k+1)$-ième au n-ième). Le passager numéro k choisit son siège au hasard et les suivants s'installent à leur place si elle n'est pas occupée et sinon au hasard. Autrement dit, on est ramené à la situation initiale. On a donc $\mathbb{P}(G \mid A_k) = p_{n-k+1}$.

On obtient :

$$p_n = \frac{1}{n}\left(1 + \sum_{k=2}^{n-1} p_{n-k+1}\right) = \frac{1}{n}\left(1 + \sum_{j=2}^{n-1} p_j\right).$$

On a, de même, $p_{n+1} = \dfrac{1}{n+1}\left(1 + \sum\limits_{j=2}^{n} p_j\right)$. On en déduit $(n+1)p_{n+1} - np_n = p_n$ et donc $p_{n+1} = p_n$. La suite $(p_n)_{n\geqslant 2}$ est constante. Il est clair que $p_2 = \dfrac{1}{2}$, donc $p_n = \dfrac{1}{2}$ pour tout $n \geqslant 2$.

15.8 On remarque que l'expérience se termine au bout d'un nombre fini de tirages, car à chaque séquence se terminant par un changement de couleur, le nombre total de boules dans l'urne diminue d'au moins une unité. On note $p(b,n)$ la probabilité cherchée et on montre que, pour tous b et n dans $\mathbb{N}^*$, on a $p(b,n) = \dfrac{1}{2}$, par récurrence forte sur $b+n \geqslant 2$.

- Pour $b+n = 2$, on a $b = n = 1$; l'expérience se termine au bout de deux tirages et la probabilité de terminer par une boule blanche est $\dfrac{1}{2}\cdot$
- Soit $N > 2$. On suppose la propriété vérifiée pour $b+n < N$ et on la démontre pour $b+n = N$. Pour $1 \leqslant i \leqslant b$, on note A_i l'événement « la première séquence comprend i boules blanches, suivies d'une boule noire » et, pour $1 \leqslant i \leqslant n$, on note B_i l'événement « la première séquence comprend i boules blanches, suivies d'une boule noire ». Par la formule des probabilités totales, on a :

$$p(b,n) = \sum_{i=1}^{b} \mathbb{P}(A_i)\,p(b-i,n) + \sum_{i=1}^{n} \mathbb{P}(B_i)\,p(b,n-i).$$

Par hypothèse de récurrence, on a $p(b-i,n)=\dfrac{1}{2}$ si $i<b$ et $p(0,n)=0$ car s'il ne reste que des boules noires, on les tire toutes successivement. De même, $p(b,n-i)=\dfrac{1}{2}$ si $i<n$ et $p(b,0)=1$. On obtient :

$$p(b,n)=\frac{1}{2}\left(\sum_{i=1}^{b-1}\mathbb{P}(A_i)+\sum_{i=1}^{n-1}\mathbb{P}(B_i)\right)+\mathbb{P}(B_n).$$

On calcule :

$$\mathbb{P}(B_n)=\frac{n}{n+b}\times\frac{n-1}{n+b-1}\times\cdots\times\frac{1}{1+b}=\frac{n!\,b!}{(n+b)!}\cdot$$

On trouve de même $\mathbb{P}(A_b)=\dfrac{n!\,b!}{(n+b)!}\cdot$ On peut donc écrire :

$$\begin{aligned}p(b,n)&=\frac{1}{2}\left(\sum_{i=1}^{b-1}\mathbb{P}(A_i)+\sum_{i=1}^{n-1}\mathbb{P}(B_i)\right)+\frac{1}{2}\Big(\mathbb{P}(A_b)+\mathbb{P}(B_n)\Big)\\&=\frac{1}{2}\left(\sum_{i=1}^{b}\mathbb{P}(A_i)+\sum_{i=1}^{n}\mathbb{P}(B_i)\right)=\frac{1}{2}\cdot\end{aligned}$$

15.9 L'expérience aléatoire peut être vue comme une suite d'épreuve de Bernoulli indépendantes (chaque fois qu'il a besoin d'allumette, il choisit au hasard la boîte dans laquelle il la prend ; ce choix entre les deux boîtes se fait de manière équiprobable et indépendamment de ce qui précède). L'événements A_k « lorsqu'il prend la dernière allumette d'une boîte, l'autre boîte contient encore k allumettes » est réalisé s'il a, lors des $2n+1-k$ premiers tirages :

- soit choisi n fois la première boîte et $n+1-k$ la seconde, en terminant par la première boîte, ce qui est réalisé avec la probabilité $\dfrac{\binom{2n-k}{n-1}}{2^{2n+1-k}}$, car il faut choisir les $n-1$ tirages effectués dans la première boîte, parmi les $2n-k$ premiers tirages ;
- soit choisi $n+1$ fois la seconde boîte et $n-k$ la première, en terminant par la seconde boîte, ce qui est réalisé avec la probabilité $\dfrac{\binom{2n-k}{n}}{2^{2n+1-k}}$, car il faut choisir les n tirages effectués dans la seconde boîte, parmi les $2n-k$ premiers tirages.

La probabilité cherchée est donc :

$$\frac{\binom{2n-k}{n-1}+\binom{2n-k}{n}}{2^{2n+1-k}}=\frac{\binom{2n-k+1}{n}}{2^{2n+1-k}}\cdot$$

15.10 1. Pour $n\in\mathbb{N}^*$, on note J_n l'événement « le joueur j_n joue » et G_n l'événement « le joueur j_n gagne ». Pour tout $n\in\mathbb{N}^*$, on a :

$$p_n=\mathbb{P}(G_n)=\mathbb{P}(G_n\cap J_n)=\mathbb{P}(G_n\mid J_n)\,\mathbb{P}(J_n)=\left(\frac{1}{2}\right)^3 q_n=\frac{1}{8}q_n,$$

car si A_n joue, pour qu'il gagne, il faut qu'il soit vainqueur de trois parties successives.

2. (a) Pour $n \leqslant 4$, le joueur j_n est sûr de jouer car on joue en tout au moins trois parties. On a donc $q_n = 1$ si $1 \leqslant n \leqslant 4$.

 (b) Pour $n \geqslant 5$, le joueur j_n entre en jeu contre j_{n-1} ou j_{n-2}. On note J_n^1 (resp. J_n^2) l'événement « j_n joue et rencontre j_{n-1} » (resp. « j_n joue et rencontre j_{n-2} »). Ces événements sont incompatibles. On obtient donc :
$$\begin{aligned} q_n = \mathbb{P}(J_n) &= \mathbb{P}(J_n^1) + \mathbb{P}(J_n^2) = \mathbb{P}(J_n^1 \cap J_{n-1}) + \mathbb{P}(J_n^2 \cap J_{n-2}) \\ &= \mathbb{P}(J_n^1 \mid J_{n-1})\,\mathbb{P}(J_{n-1}) + \mathbb{P}(J_n^2 \mid J_{n-2})\,\mathbb{P}(J_{n-2}) \\ &= \frac{1}{2}q_{n-1} + \frac{1}{4}q_{n-2}. \end{aligned}$$
 En effet, si j_{n-1} est rentré dans la partie, pour qu'il rencontre j_n, il faut et il suffit qu'il gagne sa première partie contre un adversaire rentré avant lui ; de même, une fois que j_{n-2} est rentré dans la partie, pour qu'il rencontre j_n, il faut et il suffit qu'il gagne sa première partie contre un adversaire rentré avant lui et qu'il gagne contre j_{n-1}.

3. Pour pouvoir étendre la relation à $n \geqslant 2$ et ainsi gérer plus facilement les conditions initiales, on considère la suite $(q'_n)_{n\in\mathbb{N}}$ vérifiant la relation de récurrence $q'_n = \frac{1}{2}q'_{n-1} + \frac{1}{4}q'_{n-2}$, avec $q'_3 = q_3 = 1$ et $q'_4 = q_4 = 1$. On a alors $q'_n = q_n$ pour $n \geqslant 3$.
 On obtient $q'_2 = 2$, $q'_1 = 0$ et $q'_0 = 8$. La relation de récurrence linéaire sur deux termes a pour équation caractéristique $x^2 - \frac{1}{2}x - \frac{1}{4} = 0$, de solutions $\frac{1 \pm \sqrt{5}}{4}\cdot$ Il existe $(\alpha, \beta) \in \mathbb{R}^2$ tel que, pour tout $n \in \mathbb{N}$:
$$q'_n = \alpha\left(\frac{1-\sqrt{5}}{4}\right)^n + \beta\left(\frac{1+\sqrt{5}}{4}\right)^n.$$
 Avec $q'_0 = 8$ et $q'_1 = 0$, on trouve $\alpha = \frac{4(1+\sqrt{5})}{\sqrt{5}}$ et $\beta = \frac{4(\sqrt{5}-1)}{\sqrt{5}}\cdot$ On obtient, pour tout $n \geqslant 3$:
$$q_n = \frac{4}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{4}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{4}\right)^{n-1}\right)$$
$$p_n = \frac{1}{8}q_n = \frac{1}{2\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{4}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{4}\right)^{n-1}\right).$$

15.11 On note I l'événement « le gardien est ivre » et A l'événement « le gardien fait plus de quatre tentatives ». On applique la formule de Bayes et la formule des probabilités totales :
$$\mathbb{P}(I \mid A) = \frac{\mathbb{P}(I)\,\mathbb{P}(A \mid I)}{\mathbb{P}(I)\,\mathbb{P}(A \mid I) + \mathbb{P}(\overline{I})\,\mathbb{P}(A \mid \overline{I})}.$$

- On a, d'après la formule des probabilités composées :
$$\mathbb{P}(A \mid \overline{I}) = \frac{5}{6}\,\frac{4}{5}\,\frac{3}{4}\,\frac{2}{3} = \frac{1}{3}.$$

- S'il est ivre, quel que soit l'essai la probabilité de réussir à ouvrir la porte est $\dfrac{1}{6}$ et la probabilité contraire $\dfrac{5}{6}\cdot$ On en déduit $\mathbb{P}(A \mid I) = \left(\dfrac{5}{6}\right)^4$.

On obtient :

$$\mathbb{P}(I \mid A) = \frac{\frac{1}{4}\frac{5^4}{6^4}}{\frac{1}{4}\frac{5^4}{6^4} + \frac{3}{4}\frac{1}{3}} \approx 0,32.$$

15.12 1. (a) Pour $k \in [\![1, c]\!]$, j_k ne peut jouer et donc gagner qu'en un nombre de coups de la forme $k + cn$, avec $n \in \mathbb{N}$. La probabilité qu'il gagne au bout de $k + cn$ coups exactement est

$$(q_1 q_2 \dots q_c)^n (q_1 \dots q_{k-1} p_k),$$

car cela est réalisé si, et seulement si, tous les joueurs ont perdus n fois et les joueurs avant lui ont perdu une fois de plus. On en déduit :

$$\mathbb{P}(G_k) = \sum_{n=0}^{+\infty} (q_1 \dots q_c)^n (q_1 \dots q_{k-1} p_k) = \frac{q_1 \dots q_{k-1} p_k}{1 - q_1 \dots q_c}\cdot$$

(b) L'événement G « la partie se termine au bout d'un nombre fini de coups » est la réunion disjointes des G_k. On a donc :

$$\mathbb{P}(G) = \sum_{k=1}^{c} \mathbb{P}(G_k) = \frac{1}{1 - q_1 \dots q_c} \sum_{k=1}^{c} q_1 \dots q_{k-1} p_k$$
$$= \frac{1}{1 - q_1 \dots q_c} \sum_{k=1}^{c} q_1 \dots q_{k-1}(1 - q_k) = 1,$$

car la somme est télescopique. La partie se termine presque sûrement au bout d'un nombre fini de coups.

(c) Le jeu est équitable si, et seulement si :

$$\forall k \in [\![1, c-1]\!] \quad \mathbb{P}(G_{k+1}) = \mathbb{P}(G_k),$$

ce qui équivaut, pour tout $k \in [\![1, c-1]\!]$ à $p_k = q_k p_{k+1}$, c'est-à-dire à $p_{k+1} = \dfrac{p_k}{1 - p_k}\cdot$

(d) Pour tout $k \in [\![1, c-1]\!]$, on $\dfrac{1}{p_{k+1}} = \dfrac{1}{p_k} - 1$ et donc $\dfrac{1}{p_k} = \dfrac{1}{p_1} - (k-1)$, puis $p_k = \dfrac{p_1}{1 - p_1(k-1)}\cdot$

2. (a) Pour tout $k \in \mathbb{N}^*$, on a $\mathbb{P}(G_k) = q_1 \dots q_{k-1} p_k > 0$. Comme les événements G_k sont incompatibles, la série $\sum \mathbb{P}(G_k)$ converge. C'est impossible si la suite $(\mathbb{P}(G_k))$ est constante, non nulle. Le jeu ne peut pas être équitable.

(b) La suite de terme général $Q_n = q_1 \dots q_n$ est décroissante et minorée par 0, donc elle converge. On pose $Q_0 = 1$. La probabilité que le jeu se termine en un nombre fini de coups est :

$$\sum_{k=1}^{+\infty} \mathbb{P}(G_k) = \sum_{k=1}^{+\infty} q_1 \dots q_{k-1}(1 - q_k) = \sum_{k=1}^{+\infty} (Q_{k-1} - Q_k) = 1 - a.$$

15.13 Pour tout $k \in \mathbb{N}^*$, on note Π_k l'événement « le k-ième tirage donne pile » et F_k l'événement « le k-ième tirage donne face ». Pour $n \in \mathbb{N}^*$, on note A_{2n} l'événement « la partie dure plus de $2n$ coups ».

1. Pour $n \geqslant 1$, l'événement A_{2n} est réalisé si après chacun des lancers on n'a jamais deux piles de plus que de faces ou le contraire. Il faut donc que dans chaque groupe de deux lancers successifs commençant par un lancer d'indice impair, on obtienne exactement un pile et un face. On en déduit :
$$A_{2n} = \bigcap_{i=1}^{n} (\Pi_{2i-1} \cap F_{2i}) \cup (F_{2i-1} \cap \Pi_{2i}).$$
puis, par indépendance des résultats des différents lancers :
$$\mathbb{P}(A_{2n}) = \prod_{i=1}^{n} \mathbb{P}\left((\Pi_{2i-1} \cap F_{2i}) \cup (F_{2i-1} \cap \Pi_{2i})\right) = \big(2p(1-p)\big)^n.$$

2. On note G l'événement « la personne gagne ». La personne gagne nécessairement au bout d'un nombre pair de lancers. Pour $n \geqslant 1$, on note G_{2n} l'événement « la personne gagne au $2n$-ième lancer ». L'événement G_{2n} est réalisé à condition que la partie dure plus de $2n-2$ coups et que les deux derniers lancers donnent pile. On a donc, pour $n \geqslant 2$:
$$G_{2n} = A_{2n-2} \cap \Pi_{2n-1} \cap \Pi_{2n}.$$
Par indépendance, on a :
$$\mathbb{P}(G_{2n}) = (2p(1-p))^{n-1}\, p^2.$$
On note que la formule reste valable pour $n = 1$. On a $0 < 2p(1-p) < 1$. On en déduit :
$$\mathbb{P}(G) = \sum_{n=1}^{+\infty} (2p(1-p))^{n-1} p^2 = \frac{p^2}{1-2p(1-p)} = \frac{p^2}{1-2p+2p^2}.$$
La probabilité qu'il perde s'obtient en échangeant p et $1-p$. On trouve $\dfrac{(1-p)^2}{1-2p+2p^2}$. Comme $\dfrac{p^2}{1-2p+2p^2} + \dfrac{(1-p)^2}{1-2p+2p^2} = 1$, il est presque sûr que la partie se terminera au bout d'un nombre fini de lancers.

15.14

1. La première séquence PP apparaît nécessairement aux deux premiers lancers, sinon au lancer précédent on a P et donc déjà une séquence PP ou F et une séquence FP. La probabilité cherchée est p^2.
2. Si la séquence PF apparaît pour la première fois aux lancers $n-1$ et n, les lancers précédents sont nécessairement des P, car sinon en considérant le rang k du dernier F avant celui qui est numéroté n, on aurait une séquence FP aux lancers k et $k+1$. Il faut donc avoir une suite de $(n-1)$ P suivi d'un F. Ainsi $x_n = p^{n-1}q$. L'événement « la première séquence PP apparaît avant la première séquence PF » est la réunion des événements précédents pour $n \geqslant 2$. On a donc :
$$x = \sum_{n=2}^{+\infty} x_n = q \sum_{n=2}^{+\infty} p^{n-1} = \frac{pq}{1-p} = p.$$

3. (a) On note y la probabilité cherchée et y_n la probabilité de l'événement « la séquence PP apparaît pour la première fois aux lancers $n-1$ et n et il n'y a pas eu avant de séquence PF ».
Si $n > 2$ et si la séquence PP apparaît pour la première fois aux lancers $n-1$ et n, on a nécessairement des F aux rangs précédents, car sinon on obtiendrait une séquence PF. Ainsi y_n est la probabilité d'obtenir $(n-2)$ F suivis de deux P. On a donc $y_n = q^{n-2}p^2$ et :
$$y = \sum_{n=2}^{+\infty} q^{n-2}p^2 = \frac{p^2}{1-q} = p.$$

(b) On note z la probabilité cherchée et on défini z_n comme précédemment. Pour obtenir PF pour la première fois aux lancers $n-1$ et n sans avoir de séquence PP avant, il faut avoir, si $n > 2$, un F au rang $n-2$, sinon on a une séquence PP et des F dans les lancers précédents, sinon on obtient une séquence PF. Ainsi z_n est la probabilité d'avoir $(n-2)$ F puis P et F : $z_n = q^{n-2}pq$ et :
$$z = \sum_{n=2}^{+\infty} q^{n-2}pq = \frac{pq}{1-q} = q.$$

(c) On note t la probabilité cherchée et on défini t_n comme précédemment. Si la première séquence PF apparaît aux lancers $n-1$ et n soit il n'y a pas de F avant (suite de $(n-1)$ P et d'un F), soit en considérant le dernier F avant le rang $n-1$, on voit qu'il ne peut être précédé ni d'un P ni d'un F : c'est le premier lancer. On a alors un F, puis $(n-2)$ P et un F (cela nécessite $n \geqslant 3$). On obtient $t_n = p^{n-1}q + p^{n-2}q^2$ si $n \geqslant 3$ et $t_2 = pq$. On en déduit :
$$t_n = \sum_{n=2}^{+\infty} p^{n-1}q + \sum_{n=3}^{+\infty} p^{n-2}q^2 = \frac{pq}{1-p} + \frac{pq^2}{1-p} = p + pq.$$

(d) On note u la probabilité cherchée et on défini u_n comme précédemment. Si la première séquence PP apparaît aux lancers n et $n-1$ sans qu'il y ait avant de séquence FF, c'est que dans les lancers précédents, on a une alternance de P et F (des F au lancer $n-2$ et à tous ceux qui ont la même parité). Le résultat dépend de la parité de n. On obtient :
$$t_n = \begin{cases} (pq)^{\frac{n-2}{2}}p^2 & \text{si } n \text{ est pair} \\ (pq)^{\frac{n-3}{2}}qp^2 & \text{si } n \text{ est impair.} \end{cases}$$
On a donc :
$$t = \sum_{k=0}^{+\infty}(pq)^k p^2 + \sum_{k=0}^{+\infty}(pq)^k qp^2 = \frac{p^2}{1-pq} + \frac{qp^2}{1-pq} = \frac{p^2(1+q)}{1-pq}.$$

15.15 L'espace probabilisé qui modélise cette situation est celui d'un jeu infini de pile ou face. En effet, on a une suite de parties indépendantes, à deux issues.

1. Pour $(a,b) \in \mathbb{N}^2$, on note $A_{a,b}$ l'événement « si les joueurs débutent la partie avec des capitaux respectifs de a et b euros, alors A finira ruiné ». En appliquant la

formule des probabilités totales, avec le système complet d'événements $(G_1, \overline{G_1})$, où G_1 est l'événement « A gagne la première partie », on trouve :

$$\begin{aligned} R(a,b) &= \mathbb{P}(A_{a,b}) = \mathbb{P}(G_1)\,\mathbb{P}(A_{a,b} \mid G_1) + \mathbb{P}(\overline{G_1})\,\mathbb{P}(A_{a,b} \mid \overline{G_1}) \\ &= pR(a+1,b) + qR(a,b+1). \end{aligned}$$

En effet, si A gagne la première partie, il dispose avant d'entreprendre les parties suivantes de $a+1$ euros et B de $b-1$ euros. Le raisonnement est le même dans l'autre cas.

2. Tout au long de la partie, le capital total des deux joueurs reste constant. On le note N. En fait $R(a,b)$ ne dépend donc que de a : on le note u_a. On a donc $u_a = pu_{a+1} + qu_{a-1}$, c'est-à-dire :

$$u_{a+1} = \frac{1}{p}u_a - \frac{q}{p}u_{a-1}.$$

On a donc une suite récurrente linéaire d'ordre 2 dont l'équation caractéristique est $x^2 - \frac{1}{p}x + \frac{q}{p}$. Une des racines est 1, l'autre est $\frac{q}{p}\cdot$

• Premier cas : $p \neq q$, c'est-à-dire $p \neq \frac{1}{2}\cdot$

Il existe $(\alpha,\beta) \in \mathrm{IR}^2$ tel que, pour tout $a \in [\![0,N]\!]$, $u_a = \alpha + \beta\left(\frac{q}{p}\right)^a$.

On remarque que $u_0 = 1$, car si le capital de A est nul, il est ruiné d'emblée. De même, $u_N = 0$, car alors c'est B qui est ruiné. On en déduit $\alpha = -\dfrac{\left(\frac{q}{p}\right)^N}{1-\left(\frac{q}{p}\right)^N}$ et $\beta = \dfrac{1}{1-\left(\frac{q}{p}\right)^N}$, puis $u_a = \dfrac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^N}{1-\left(\frac{q}{p}\right)^N}\cdot$

En reprenant la notation initiale, on a :

$$R(a,b) = \frac{\left(\frac{q}{p}\right)^a - \left(\frac{q}{p}\right)^{a+b}}{1-\left(\frac{q}{p}\right)^{a+b}}\cdot$$

• Second cas : $p = q$, c'est-à-dire $p = \frac{1}{2}\cdot$.

Il existe $(\alpha,\beta) \in \mathrm{IR}^2$ tel que, pour tout $a \in [\![0,N]\!]$, $u_a = \alpha + \beta a$.

De $u_0 = 1$ et $u_N = 0$, on tire $\alpha = 1$ et $\beta = -\frac{1}{N}$ et donc $u_a = 1 - \frac{a}{N}\cdot$ On conclut :

$$R(a,b) = \frac{b}{a+b}\cdot$$

3. La probabilité de ruine $S(a,b)$ du joueur B quand les capitaux initiaux sont a et b s'obtient à partir de $R(a,b)$ en échangeant les rôles, c'est-à-dire en échangeant

p et q et a et b. On obtient :

$$S(a,b) = \frac{\left(\frac{p}{q}\right)^b - \left(\frac{p}{q}\right)^{a+b}}{1 - \left(\frac{p}{q}\right)^{a+b}} \quad \text{si} \quad p \neq q \quad \text{et} \quad S(a,b) = \frac{a}{a+b} \quad \text{si} \quad p = q.$$

Si $p \neq q$, on obtient :

$$R(a,b) + S(a,b) = \frac{q^a p^b - q^{a+b}}{p^{a+b} - q^{a+b}} + \frac{q^a p^b - p^{a+b}}{q^{a+b} - p^{a+b}} = 1.$$

Ce résultat est évident pour $p = q$. Il est donc presque sûr que le jeu s'arrêtera au bout d'un nombre fini de parties.

4. Si $p \leqslant q$, on trouve $\lim\limits_{b \to +\infty} R(a,b) = 1$.

   Si $p > q$, on trouve $\lim\limits_{b \to +\infty} R(a,b) = \left(\dfrac{p}{q}\right)^a$. Dans ce cas, c'est-à-dire si le joueur A joue mieux que le joueur B, la ruine du joueur A n'est pas certaine, même si le capital initial du joueur B est très important.

15.16 1. On note A_n (resp. B_n) l'événement « le parapluie est disponible (resp. indisponible) ». Pour $n \in \mathbb{N}$ applique la formule des probabilités totales, avec le système complet d'événements (A_n, B_n). On obtient :

$$p_{n+1} = \mathbb{P}(A_{n+1}) = \mathbb{P}(A_n)\,\mathbb{P}(A_{n+1} \mid A_n) + \mathbb{P}(B_n)\,\mathbb{P}(A_{n+1} \mid B_n) = p\,p_n + q_n.$$

En effet, $\mathbb{P}(A_{n+1} \mid A_n) = p$, car si le parapluie est disponible après n déplacements, il faut que le banquier l'emporte, c'est-à-dire qu'il pleuve, pour qu'il reste disponible après $n+1$ déplacements ; $\mathbb{P}(A_{n+1} \mid B_n) = 1$, car si après n déplacements le parapluie n'est pas disponible, il le sera après $n+1$.

On trouve de même :

$$q_{n+1} = \mathbb{P}(B_{n+1}) = \mathbb{P}(A_n)\,\mathbb{P}(B_{n+1} \mid A_n) + \mathbb{P}(B_n)\,\mathbb{P}(B_{n+1} \mid B_n) = q\,p_n + 0\,q_n$$

et donc $S = \begin{pmatrix} p & 1 \\ q & 0 \end{pmatrix}$.

2. On diagonalise S : les valeurs propres sont 1 et $-q$. Une base de vecteurs propres est $((1,q),(-1,1))$. On obtient $S = P \begin{pmatrix} 1 & 0 \\ 0 & -q \end{pmatrix} P^{-1}$, où $P = \begin{pmatrix} 1 & -1 \\ q & 1 \end{pmatrix}$ et $P^{-1} = \dfrac{1}{q+1} \begin{pmatrix} 1 & 1 \\ -q & 1 \end{pmatrix}$. On a, pour tout $n \in \mathbb{N}$:

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = S^n \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = P \begin{pmatrix} 1 & 0 \\ 0 & (-q)^n \end{pmatrix} P^{-1} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}.$$

On en déduit que les suites (p_n) et (q_n) convergent. On note p_∞ et q_∞ leurs limites. On a :

$$\begin{pmatrix} p_\infty \\ q_\infty \end{pmatrix} = P \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \frac{1}{1+q} \begin{pmatrix} 1 & 1 \\ q & q \end{pmatrix} \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \frac{1}{1+q} \begin{pmatrix} 1 \\ q \end{pmatrix},$$

car $p_0 + q_0 = 1$. On a donc $p_\infty = \dfrac{1}{1+q}$.

Chapitre 16 : Variables aléatoires discrètes

Variables aléatoires discrètes

16

I Variables aléatoires discrètes

1 Définition

Définition 1

Une **variable aléatoire ! discrète** sur l'espace probabilisable $(\Omega, \mathcal{A})$ est une application de Ω dans un ensemble quelconque E telle que :

1. l'ensemble $X(\Omega)$ est au plus dénombrable ;
2. pour tout $x \in X(\Omega)$, on a $X^{-1}(\{x\}) \in \mathcal{A}$.

Si $E \subset \mathbb{R}$, on dit que X est une variable aléatoire discrète réelle.
Si $X(\Omega)$ est fini, on dit que X est une variable aléatoire finie.

Remarque L'appartenance de $X^{-1}(\{x\})$ à $\mathcal{A}$ signifie que c'est un événement.

Proposition 1

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A})$, à valeurs dans E. Alors pour tout partie A de E, l'ensemble $X^{-1}(A)$ est un événement.

Démonstration page 924

Notations

- Comme pour les variables aléatoires finies, l'événement $X^{-1}(A)$ est noté
$$\{X \in A\}.$$
- En particulier, pour tout $x \in E$, l'événement $\{\omega \in \Omega \mid X(\omega) = x\}$ est noté :
$$\{X = x\}.$$
On remarque que, pour tout $x \in E$, on a $\{X = x\} = \varnothing$ si $x \notin X(\Omega)$; ainsi $\{X = x\} = \varnothing$ sauf pour un nombre au plus dénombrable de valeurs de x.

Exemples Soit $(\Omega, \mathcal{A})$ un espace probabilisable.

1. Si A est un événement, alors la fonction $\mathbb{1}_A$ est une variable aléatoire finie. En effet, on a $X(\Omega) \subset \{0,1\}$, qui est fini. De plus $\{X=1\} = A$ et $\{X=0\} = \bar{A}$ sont des événements.
2. Une fonction constante sur Ω est une variable aléatoire finie. En effet, il existe a tel que $X(\Omega) = \{a\}$, donc $\{X=a\} = \Omega \in \mathcal{A}$.

2 Loi d'une variable aléatoire discrète

Soit X une variable aléatoire discrète sur l'espace probabilisable $(\Omega, \mathcal{A})$. Si $(\Omega, \mathcal{A})$ est muni d'une probabilité, alors on peut définir la loi de X.

Théorème 2

Si X est une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors l'application :

$$\begin{array}{rcl} \mathcal{P}(X(\Omega)) & \longrightarrow & [0,1] \\ A & \longmapsto & \mathbb{P}(X \in A) \end{array}$$

est une probabilité sur $\big(X(\Omega), \mathcal{P}(X(\Omega))\big)$, appelée **loi de X** et notée $\mathbb{P}_X$.

Démonstration page 924

Remarques

- L'ensemble $X(\Omega)$ étant dénombrable, il n'y a aucun inconvénient à le munir de la tribu $\mathcal{P}(X(\Omega))$.
- Cette probabilité est la probabilité image de $\mathbb{P}$ par l'application X (voir l'exercice 15.1 de la page 873).
- Il peut arriver qu'on ne connaisse pas exactement $X(\Omega)$, mais qu'on sache seulement *a priori* qu'il existe un ensemble E dénombrable tel que $X(\Omega) \subset E$. Alors la loi de X apparaîtra comme une probabilité sur $\big(E, \mathcal{P}(E)\big)$ telle que $\mathbb{P}_X(\{\omega\}) = 0$ si $\omega \in E \setminus X(\Omega)$.

Définition 2

Une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ est dite **presque sûrement constante** s'il existe a tel que $\mathbb{P}(X=a) = 1$.

Proposition 3

Si X est une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, la famille $\big(\{X=x\}\big)_{x \in X(\Omega)}$ est un système complet d'événements, appelé **système complet d'événements associé à X**. On a, en particulier :

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

Démonstration page 924

Exemple Le système complet d'événements associé à la variable $\mathbb{1}_A$ est $(A, \bar{A})$.

Proposition 4

Si X est une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, la loi de X est déterminée de manière unique par la donnée, pour tout $x \in X(\Omega)$, de :

$$\mathbb{P}_X(\{x\}) = \mathbb{P}(X = x).$$

Plus précisément, on a pour tout $A \subset X(\Omega)$:

$$\mathbb{P}_X(A) = \sum_{x \in A} \mathbb{P}(X = x).$$

Démonstration. Compte tenu de la proposition 3 de la page précédente, il s'agit d'un cas particulier de la définition d'une probabilité sur un univers au plus dénombrable par la probabilité des événements élémentaires (théorème 9 de la page 854). □

Remarque Il résulte de la définition que la loi d'une variable aléatoire discrète n'est rien d'autre qu'une probabilité sur un ensemble dénombrable E muni de la tribu $\mathcal{P}(E)$. La proposition suivante montre que toute probabilité sur un ensemble dénombrable peut être considérée comme la loi d'une variable aléatoire discrète. Une telle probabilité est encore appelée **loi discrète**.

Proposition 5

Soit E un ensemble au plus dénombrable et $(p_x)_{x \in E}$ une famille de réels positifs tels que $\sum\limits_{x \in E} p_x = 1$. Alors il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et une variable aléatoire X sur cet espace probabilisé tels que :

$$X(\Omega) = E \qquad \text{et} \qquad \forall x \in E \quad \mathbb{P}(X = x) = p_x.$$

Principe de démonstration. On pourra choisir $\big(\Omega, \mathcal{A}\big) = \big(E, \mathcal{P}(E)\big)$ et $X = \mathrm{Id}_E$.

Démonstration page 924

Remarque Les conditions énoncées en première année pour définir la loi d'une variable aléatoire finie apparaissent comme un cas particulier des conditions énoncées dans la proposition 5. Les lois étudiées en première année peuvent être utilisées dans des espaces probabilisés quelconques.

Exemple La loi d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est donc une probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Elle est définie par la donnée d'une série $\sum p_n$ à termes positifs convergente de somme 1.

Définition 3

On dit que deux variables aléatoires discrètes X et Y ont **même loi** si $\mathbb{P}_X = \mathbb{P}_Y$. On note alors $X \sim Y$.

Si X est une variable aléatoire et $\mathcal{L}$ est une loi discrète, on note alors $X \sim \mathcal{L}$ le fait que la loi de X est $\mathcal{L}$.

Remarques

- On dit aussi de variables aléatoires de même loi qu'elles sont **équidistribuées**.
- Deux variables aléatoires discrètes X et Y définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ ont même loi si, et seulement si, $X(\Omega) = Y(\Omega)$ et, pour tout $x \in X(\Omega)$, $\mathbb{P}(X = x) = \mathbb{P}(Y = x)$.

  Cela ne signifie nullement que $X = Y$.
- On considérera encore que X et Y ont même loi si $X(\Omega)$ et $Y(\Omega)$ ne diffèrent que par des valeurs de probabilité nulle, c'est-à-dire si, pour tout $x \in X(\Omega) \cup Y(\Omega)$, on a $\mathbb{P}(X = x) = \mathbb{P}(Y = x)$

p.925 **Exercice 1** Soit X une variable aléatoire réelle sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1. Calculer $\lim\limits_{n \to +\infty} \mathbb{P}(X \leqslant n)$. En déduire $\lim\limits_{x \to +\infty} \mathbb{P}(X \leqslant x)$.
2. Calculer $\lim\limits_{n \to +\infty} \mathbb{P}(X \leqslant -n)$. En déduire $\lim\limits_{x \to -\infty} \mathbb{P}(X \leqslant x)$.

Image d'une variable aléatoire par une application

Proposition 6

Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$ et f une application définie sur $X(\Omega)$. Alors $f \circ X$ est une variable aléatoire discrète, notée $f(X)$, dont la loi est donnée par :

$$\forall y \in f(X)(\Omega) \quad \mathbb{P}(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} \mathbb{P}(X = x).$$

Démonstration page 925

II Lois usuelles

Les lois discrètes définies ici s'ajoutent aux lois finies étudiées en première année (loi uniforme, loi de Bernoulli, loi binomiale).

1 Loi géométrique

Définition 4

Soit $p \in]0, 1[$.
On dit qu'une variable aléatoire discrète X suit la loi géométrique de paramètre p si $X(\Omega) = \mathbb{N}^*$ et :

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

où $q = 1 - p$.

La proposition « X suit la loi géométrique de paramètre p » se note $X \sim \mathcal{G}(p)$.

Remarque On définit bien ainsi la loi d'une variable aléatoire discrète car on a $p\,q^{k-1} \geqslant 0$ pour tout $k \in \mathbb{N}^*$ et :

$$\sum_{k=1}^{+\infty} p\,q^{k-1} = \frac{p}{1-q} = 1.$$

Situation type

On considère une suite d'épreuves de Bernoulli indépendantes, la probabilité de succès à chaque épreuve étant $p \in]0,1[$ (on pose $q = 1-p$), modélisée par un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On considère le temps d'attente du premier succès.

Plus précisément, pour $k \in \mathbb{N}^*$, on note A_k l'événement « la k-ième épreuve est un succès » et l'on pose :

- $T = n$ (pour $n \in \mathbb{N}^*$) si l'événement $\overline{A_1} \cap \ldots \cap \overline{A_{n-1}} \cap A_n$ est réalisé ; alors $\{T = n\}$ est un événement, car intersection finie d'événements, et $\mathbb{P}(T = n) = (1-p)^{n-1}p = q^{n-1}p$;
- $T = +\infty$ si aucun des A_n, pour $n \in \mathbb{N}^*$, n'est réalisé ; on a donc :

  $$\{T = +\infty\} = \bigcap_{n \in \mathbb{N}^*} \overline{A_n} \in \mathcal{A},$$

  et $\mathbb{P}(T = +\infty) = \lim\limits_{n \to +\infty} \mathbb{P}\left(\bigcap\limits_{k=1}^{n} \overline{A_k}\right) = \lim\limits_{n \to +\infty} q^n = 0.$

La variable aléatoire discrète T prend la valeur $+\infty$ avec une probabilité nulle et, pour tout $k \in \mathbb{N}^*$, on a $\mathbb{P}(T = k) = \mathbb{P}(X = k)$, où X suit la loi géométrique de paramètre p. Dans un tel cas, on dira encore que T suit la loi géométrique $\mathcal{G}(p)$.

Proposition 7

Si la variable aléatoire X suit la loi géométrique de paramètre $p \in]0,1[$ à valeurs dans $\mathbb{N}^*$, on a, pour tout entier naturel k,

$$\mathbb{P}(X > k) = q^k.$$

Démonstration page 925

Interprétation Ce résultat est tout à fait évident si l'on interprète l'on interprète une loi géométrique comme le temps d'attente d'un premier succès. L'événement $\{X > k\}$ est « les k premières épreuves ont été des échecs », de probabilité q^k.

Corollaire 8

Si la variable aléatoire X suit la loi géométrique de paramètre $p \in]0,1[$ à valeurs dans $\mathbb{N}^*$, on a, pour tout couple d'entiers naturels (k, ℓ) :

$$\mathbb{P}(X > k + \ell \mid X > k) = \mathbb{P}(X > \ell).$$

Démonstration page 925

Remarque La probabilité que X prenne une valeur supérieure à $k+\ell$ sachant qu'elle est supérieure à k, donc qu'elle s'accroisse au moins de ℓ est égale à la probabilité qu'elle prenne une valeur supérieure à ℓ. Si l'on pense à X comme à une durée, on peut dire que X ne tient pas compte du passé. On dit que la variable aléatoire X est **sans mémoire**.

Proposition 9

Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$. On suppose que X est sans mémoire, c'est-à-dire que, pour tout couple d'entiers naturels (k, ℓ), on a :

$$\mathbb{P}(X > k + \ell \mid X > k) = \mathbb{P}(X > \ell).$$

Montrer que X suit une loi géométrique.

Principe de démonstration. On étudie la suite de terme général $u_n = \mathbb{P}(X > n)$.

Démonstration page 925

p.926 **Exercice 2 (Temps d'attente du r-ième succès)**

On considère une suite d'épreuves de Bernoulli indépendantes, de probabilité de succès $p \in]0, 1[$. On pose $q = 1 - p$. On définit la variable aléatoire réelle Y_r égale au nombre de tirages nécessaires pour obtenir r succès pour la première fois et à ∞ si on n'obtient jamais r succèss.

Déterminer la loi de Y_r.

2 Loi de Poisson

Définition 5

Soit $\lambda > 0$. On dit qu'une variable aléatoire réelle discrète X suit **la loi de Poisson de paramètre $\boldsymbol{\lambda}$**, si $X(\Omega) = \mathbb{N}$ et :

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

La proposition « X suit la loi de Poisson de paramètre λ » se note $X \sim \mathcal{P}(\lambda)$.

Remarque

On définit bien ainsi la loi d'une variable aléatoire discrète car on a $\dfrac{\lambda^k}{k!} e^{-\lambda} \geqslant 0$ pour tout $k \in \mathbb{N}$ et :

$$\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

Il n'est pas possible de donner un modèle simple pour la loi de Poisson. Celle-ci apparaît comme une loi limite.

Proposition 10

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires discrètes telles que X_n suive la loi binomiale de paramètre (n, p_n). On suppose que la suite (np_n) converge vers un réel $\lambda > 0$. Alors, pour tout entier k de $\mathbb{N}$, on a :

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

Démonstration page 926

Remarque

Si la variable aléatoire X suit une loi binomiale avec n grand et p proche de 0, elle suit approximativement une loi de Poisson de paramètre $\lambda = np$. On dit encore que la loi de Poisson est la **loi des événements « rares »**.

Dans la pratique, on peut décrire par une loi de Poisson le nombre d'événements d'un certain type se produisant dans une période de temps donnée, par exemple :

- le nombre de clients se présentant dans un magasin ;
- le nombre de véhicules franchissant un poste de péage ;
- le nombre d'appels reçus par un standard téléphonique.

Voir l'exercice 11 de la page 905.

III Couples de variables aléatoires

1 Définitions

Définition 6

Si X et Y sont deux variables aléatoires discrètes sur l'espace probabilisable $(\Omega, \mathcal{A})$, à valeurs respectivement dans E et E', l'application :

$$\begin{array}{ccc} \Omega & \longrightarrow & E \times E' \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array}$$

est appelée **couple de variables aléatoires discrètes** sur $(\Omega, \mathcal{A})$. On note (X, Y) ce couple de variables aléatoires.

Si $E = E' = \mathbb{R}$, alors (X, Y) est appelé un **couple de variables aléatoires réelles discrètes**.

Les couples de variables aléatoires sont simplement les variables aléatoires à valeurs dans $E \times E'$, comme le prouve la proposition suivante.

Proposition 11

- Si X et Y sont deux variables aléatoires discrètes sur l'espace probabilisable $(\Omega, \mathcal{A})$, à valeurs respectivement dans E et E', le couple (X, Y) est une variable aléatoire discrète à valeurs dans $E \times E'$.
- Réciproquement, toute variable aléatoire discrète sur l'espace probabilisable $(\Omega, \mathcal{A})$ à valeurs dans $E \times E'$ peut s'écrire (X, Y) où X et Y sont des variables aléatoires discrètes à valeurs respectivement dans E et E'.

Démonstration page 927

Proposition 12

Soit (X, Y) un couple de variables aléatoires discrètes sur l'espace probabilisable $(\Omega, \mathcal{A})$. Alors la famille d'événements :

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

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

Démonstration page 927

Remarque Il découle immédiatement de cette proposition que si (X, Y) est un couple de variables aléatoires sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors on a :

$$\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big) = 1.$$

Proposition 13

L'ensemble des variables aléatoires réelles discrètes sur l'espace probabilisable $(\Omega, \mathcal{A})$ est un espace vectoriel réel. De plus, le produit de deux variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A})$ est une variable aléatoire réelle discrète.

Démonstration page 927

2 Loi conjointe

Définition 7

Soit X et Y deux variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans E et E' respectivement. La **loi conjointe** de X et Y est la loi du couple (X, Y).

Remarques

- La loi de (X, Y) est déterminée par la famille :

$$\Big(\mathbb{P}\big(\{X = x\} \cap \{Y = y\}\big)\Big)_{(x,y)\in X(\Omega)\times Y(\Omega)}.$$

Cela résulte de la proposition 4 de la page 894 appliquée à la variable (X, Y).

- Soit E et E' deux ensembles au plus dénombrables, $(p_{x,y})_{(x,y)\in E\times E'}$ une famille de réels. Il résulte de la proposition 5 de la page 894 qu'il existe un couple (X,Y) de variables aléatoires à valeurs dans $E\times E'$ tel que, pour tout $(x,y)\in E\times E'$, on ait $\mathbb{P}\big(\{X=x\}\cap\{Y=y\}\big)=p_{x,y}$ si, et seulement si, les $p_{x,y}$ sont positifs et vérifient $\sum\limits_{(x,y)\in E\times E'} p_{x,y}=1$.

p.927 **Exercice 3** Soit a et λ des réels strictement positifs. Soit X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. On suppose que la loi conjointe de X et Y vérifie :

$$\forall (i,j)\in\mathbb{N}^2 \quad \mathbb{P}\left(\{X=i\}\cap\{Y=j\}\right)=a\frac{(i+j)\lambda^{i+j}}{i!\,j!}\cdot$$

Déterminer la valeur de a en fonction de λ.

3 Lois marginales

Définition 8

Pour tout couple (X,Y) de variables aléatoires discrètes, la loi de X est appelée **première loi marginale du couple** et celle de Y est appelée **deuxième loi marginale du couple**.

Le théorème suivant exprime le fait que l'on peut déduire les loi marginales de la loi du couple. Pour obtenir une loi marginale, on somme par rapport à l'autre variable.

Théorème 14

Soit (X,Y) un couple de variables aléatoires discrètes sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$. On dispose des égalités suivantes :

$$\forall x\in X(\Omega) \quad \mathbb{P}(X=x)=\sum_{y\in Y(\Omega)}\mathbb{P}\big(\{X=x\}\cap\{Y=y\}\big);$$

$$\forall y\in Y(\Omega) \quad \mathbb{P}(Y=y)=\sum_{x\in X(\Omega)}\mathbb{P}\big(\{X=x\}\cap\{Y=y\}\big).$$

Démonstration. Ces égalités résultent de ce que $(\{Y=y\})_{y\in Y(\Omega)}$ et $\big(\{X=x\}\big)_{y\in X(\Omega)}$ sont des systèmes complets d'événements (on applique la proposition 3 de la page 851). □

p.928 **Exercice 4** On reprend l'exemple de l'exercice 3 de la page 900. Déterminer les lois marginales de X et Y.

p.928 **Exercice 5** On considère une suite d'épreuves de Bernoulli indépendantes de probabilité de succès $p\in\,]0,1[$. On note X (respectivement Y) le rang du premier (respectivement du second) succès.

1. Déterminer la loi du couple (X,Y).
2. En déduire les lois de X et Y.

4 Lois conditionnelles

Définition 9

Soit (X, Y) un couple de variables aléatoires discrètes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Pour tout y de $Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$, la **loi conditionnelle de X sachant $\{Y = y\}$** est la loi de X dans l'espace probabilisé $\left(\Omega, \mathcal{A}, \mathbb{P}_{\{Y=y\}}\right)$. Elle est donc déterminée par la donnée, pour tout $x \in X(\Omega)$, de :

$$\mathbb{P}_{\{Y=y\}}(X = x) = \mathbb{P}\left(X = x \mid Y = y\right).$$

De même, pour tout x de $X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$, la **loi conditionnelle de Y sachant $\{X = x\}$** est la loi de Y dans l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P}_{\{X=x\}})$.

Remarques

- On rappelle que $\mathbb{P}_{\{Y=y\}}$ est la probabilité conditionnelle sachant $\{Y = y\}$.
- Les lois conditionnelles sont des lois de variables aléatoires ; elles en ont les propriétés.

p.928 **Exercice 6** On reprend l'exemple du couple (X, Y) des deux premiers succès dans une suite d'épreuves de Bernoulli.

Déterminer pour $j \geqslant 2$, la loi conditionnelle de X sachant $\{Y = j\}$ et pour $i \geqslant 1$, la loi conditionnelle de $Y - i$ sachant $\{X = i\}$. Commenter.

Proposition 15

Soit (X, Y) un couple de variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$.
On suppose que, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, on a $\mathbb{P}(X = x) \neq 0$ et $\mathbb{P}(Y = y) \neq 0$. Alors, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$:

$$\begin{aligned}
\mathbb{P}(\{X = x\} \cap \{Y = y\}) &= \mathbb{P}(Y = y)\mathbb{P}(X = x \mid Y = y) \\
&= \mathbb{P}(X = x)\mathbb{P}(Y = y \mid X = x),
\end{aligned}$$

$$\mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(Y = y)\,\mathbb{P}(X = x \mid Y = y),$$

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

Démonstration. Les premières égalités résultent de la définition des probabilités conditionnelles, les deux dernières de la formule des probabilités totales appliquée aux systèmes complets d'événements $(\{Y = y\})_{y \in Y(\Omega)}$ et $(\{X = x\})_{x \in X(\Omega)}$ associés aux variables aléatoires Y et X. □

Remarque Ainsi, connaissant une des lois marginales et la probabilité conditionnelle de l'autre variable par rapport à celle-ci, on obtient la loi conjointe et la loi marginale de la deuxième variable.

p.929 **Exercice 7** Soit X et Y deux variables aléatoires sur le même espace probabilisé, à valeurs dans $\mathbb{N}$, $p \in]0,1[$ et $\lambda > 0$. On suppose que X suit la loi de Poisson de paramètre λ et que, pour $n \in \mathbb{N}$, la loi conditionnelle de Y sachant $\{X = n\}$ est la loi binomiale de paramètre (n,p). Déterminer la loi de Y.

5 Généralisation aux n-uplets de variables aléatoires

Définition 10

Soit $n \in \mathbb{N}^*$. Si $X_1, \ldots, X_n$ sont des variables aléatoires discrètes sur l'espace probabilisable $(\Omega, \mathcal{A})$, à valeurs respectivement dans $E_1, \ldots, E_n$, alors l'application :

$$\begin{array}{ccc} \Omega & \longrightarrow & E_1 \times \cdots \times E_n \\ \omega & \longmapsto & (X_1(\omega), \ldots, X_n(\omega)) \end{array}$$

est appelée **n-uplet de variables aléatoires** sur Ω.
On le noté $(X_1, \ldots, X_n)$.

Si $E_1 = \cdots = E_n = \mathbb{R}$, $(X_1, \ldots, X_n)$ est appelé un **vecteur aléatoire discret**.

On peut démontrer, comme dans le cas des couples de variables aléatoires discrètes, la proposition suivante.

Proposition 16

Les n-uplets de variables aléatoires discrètes, à valeurs respectivement dans $E_1, \ldots, E_n$, sont les variables aléatoires discrètes à valeurs dans $E_1 \times \cdots \times E_n$.

Remarque Un vecteur aléatoire discret est donc une variable aléatoire à valeurs dans un espace vectoriel $\mathbb{R}^n$.

Définition 11

Soit $X_1, \ldots, X_n$ des variables aléatoires sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

- La **loi conjointe** de $X_1, \ldots, X_n$ est la loi du n-uplet $(X_1, \ldots, X_n)$.
- Les **lois marginales** du n-uplet $(X_1, \ldots, X_n)$ sont les lois des variables aléatoires $X_1, X_2, \ldots, X_n$.

Remarque
Si X_1, ..., X_n sont des variables aléatoires discrètes sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$, alors la loi conjointe de X_1, ..., X_n est déterminée par la donnée, pour tout $(x_1,\dots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)$, de :

$$\mathbb{P}(\{X_1=x_1\}\cap\cdots\cap\{X_n=x_n\}).$$

L'événement $\{X_1=x_1\}\cap\cdots\cap\{X_n=x_n\}$ est noté $\{X_1=x_1,\dots,X_n=x_n\}$ également et sa probabilité $\mathbb{P}(X_1=x_1,\dots,X_n=x_n)$.

Comme dans le cas des couples de variables aléatoires, les lois marginales des n-uplets s'obtiennent à partir de la loi conjointe. Pour obtenir une loi marginale, il suffit donc sommer par rapport aux autres variables.

Théorème 17

Soit $(X_1,\dots,X_n)$ un n-uplet de variables aléatoires sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$. Pour tout $k\in[\![1,n]\!]$ et tout $x_k\in X_k(\Omega)$, on a :

$$\mathbb{P}(X_k=x_k)=\sum_{\substack{x_1\in X_1(\Omega),\dots,x_{k-1}\in X_{k-1}(\Omega),\\ x_{k+1}\in X_{k+1}(\Omega),\dots,x_n\in X_n(\Omega)}}\mathbb{P}(X_1=x_1,\dots,X_n=x_n).$$

Démonstration page 929

p.930 **Exercice 8** Montrer que si X_1, ..., X_n sont des variables aléatoires discrètes sur le même espace probabilisable et f est une fonction quelconque définie sur $E_1\times\cdots\times E_n$, alors $f(X_1,\dots,X_n)$ est une variable aléatoire discrète.

Exemple En particulier si $X_1,\dots,X_n$ sont des variables aléatoires discrètes, alors $X_1+\cdots+X_n$, $X_1\times\cdots\times X_n$, $\min(X_1,\dots,X_n)$ et $\max(X_1,\dots,X_n)$ sont des variables aléatoires.

IV Indépendance de variables aléatoires

1 Indépendance de deux variables aléatoires

Définition 12

Deux variables aléatoires discrètes X et Y sur $(\Omega,\mathcal{A},\mathbb{P})$ sont dites **indépendantes** si, pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$, les événements $\{X=x\}$ et $\{Y=y\}$ sont indépendants, c'est-à-dire vérifient :

$$\mathbb{P}(\{X=x\}\cap\{Y=y\})=\mathbb{P}(X=x)\,\mathbb{P}(Y=y).$$

Attention L'indépendance est une notion qui dépend de la probabilité choisie.

Exemple Une variable presque sûrement constante est indépendante de toute variable aléatoire discrète, car un événement de probabilité 0 ou 1 est indépendant de tout événement.

Remarques

- Dans le cas de variables aléatoires indépendantes, la donnée des lois marginales permet donc de connaître la loi conjointe.
- L'indépendante de deux variables aléatoires se lit sur la forme de la loi conjointe. En effet, si X et Y sont indépendantes, on obtient en posant $\varphi(x) = \mathbb{P}(X = x)$ et $\psi(y) = \mathbb{P}(Y = y)$:
$$\forall (x, y) \in X(\Omega) \times Y(\Omega) \quad \mathbb{P}(\{X = x\} \cap \{Y = y\}) = \varphi(x)\, \psi(y).$$
La réciproque fait l'objet de l'exercice suivant.

p.930 **Exercice 9** On suppose que la loi conjointe du couple (X, Y) s'écrit :
$$\forall (x, y) \in X(\Omega) \times Y(\Omega) \quad \mathbb{P}\left(\{X = x\} \cap \{Y = y\}\right) = \varphi(x)\, \psi(y).$$
où φ et ψ sont des fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$. Montrer que X et Y sont indépendantes et qu'il existe des constantes C et C' telle que :
$$\forall (x, y) \in X(\Omega) \times Y(\Omega) \quad \mathbb{P}(X = x) = C\varphi(x) \qquad \text{et} \qquad \mathbb{P}(Y = y) = C'\psi(y).$$

Proposition 18

Si X et Y sont des variables aléatoires discrètes $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans E et F respectivement, indépendantes, alors, pour toute partie A de E et toute partie B de F, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, c'est-à-dire vérifient :
$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A)\, \mathbb{P}(Y \in B).$$

Démonstration page 930

Proposition 19

Soit (X, Y) un couple de variables aléatoires discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Il y a équivalence entre :

(i) les variables aléatoires X et Y sont indépendantes ;

(ii) pour tout y de $Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$, la loi conditionnelle de X sachant $\{Y = y\}$ est égale à la loi de X ;

(iii) pour tout x de $Y(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$, la loi conditionnelle de Y sachant $\{X = x\}$ est égale à la loi de Y.

Démonstration page 931

Proposition 20

Si X et Y sont des variables aléatoires discrètes, indépendantes, sur $(\Omega, \mathcal{A}, \mathbb{P})$ alors, pour toute fonction f définie sur $X(\Omega)$ et toute fonction g définie sur $Y(\Omega)$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Démonstration page 931

Loi de la somme de deux variables indépendantes

Proposition 21

Si X et Y sont deux variables aléatoires indépendantes, à valeurs dans $\mathbb{N}$, on a :

$$\forall n \in \mathbb{N} \quad \mathbb{P}(X+Y=n) = \sum_{k=0}^{n} \mathbb{P}(X=k)\,\mathbb{P}(Y=n-k).$$

Si X et Y sont deux variables aléatoires indépendantes, à valeurs dans $\mathbb{Z}$, on a :

$$\forall n \in \mathbb{Z} \quad \mathbb{P}(X+Y=n) = \sum_{k=-\infty}^{+\infty} \mathbb{P}(X=k)\,\mathbb{P}(Y=n-k).$$

Démonstration page 931

L'exercice suivant montre que la loi de Poisson, comme la loi binomiale est stable.

p.931

Exercice 10 Montrer que, si X et Y sont des variables aléatoires indépendantes, suivant des lois de Poisson de paramètres respectifs λ et μ, alors la somme $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

Exemple Supposons que le nombre de clients se présentant dans un magasin dans un intervalle de temps donné suive une loi de Poisson. Notons λ le paramètre de cette loi pour l'intervalle $[a,b[$ et μ pour l'intervalle $[b,c[$, avec $a<b<c$. Alors, la loi correspondant à l'intervalle $[a,c[$ est $\mathcal{P}(\lambda+\mu)$.
On en déduit la modélisation suivante : la nombre de clients se présentant dans le magasin dans un intervalle de temps $[a,b[$ suit une loi de Poisson de paramètre $\alpha\,(b-a)$, où α est un réel strictement positif fixé.

Réciproquement, l'exercice suivant montre que c'est la seule modélisation envisageable.

p.932

Exercice 11 Pour $a<b$, notons $N_{a,b}$ le nombre de clients se présentant dans un magasin dans l'intervalle de temps $[a,b[$. Soit $a<b<c$; on suppose que les variables $N_{a,b}$ et $N_{b,c}$ sont indépendantes et que, pour tout $n \in \mathbb{N}$, la loi de $N_{a,b}$ sachant $\{N_{a,c}=n\}$ est $\mathcal{B}(n,p)$, avec $p=(b-a)/(c-a)$.

1. Montrer que pour tout $(k,\ell) \in \mathbb{N}^2$:
$$\frac{\mathbb{P}(N_{a,b}=k)\,\mathbb{P}(N_{b,c}=\ell)}{\mathbb{P}(N_{a,c}=k+\ell)} = \binom{k+\ell}{k} p^k\,(1-p)^\ell.$$
2. En déduire qu'il existe une constante λ telle que :
$$\forall n \in \mathbb{N}^* \quad \mathbb{P}(N_{a,b}=n) = \frac{\lambda}{n}\,\mathbb{P}(N_{a,b}=n-1).$$
3. En déduire que $N_{a,b}$ suit une loi de Poisson de paramètre λ.

2 Indépendance de n variables aléatoires

Définition 13

Les variables aléatoires discrètes $X_1,\dots,\ X_n$ sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$ sont dites **indépendantes deux à deux** si, pour tous entiers i et j distincts de $[\![1,n]\!]$, les variables aléatoires X_i et X_j sont indépendantes.

Définition 14

Les variables aléatoires discrètes $X_1,\dots,\ X_n$ sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$ sont dites **mutuellement indépendantes** si, pour tout $(x_1,x_2,\dots,x_n)\in X_1(\Omega)\times X_2(\Omega)\times\cdots\times X_n(\Omega)$, on a :

$$\mathbb{P}\left(\{X_1=x_1\}\cap\ldots\cap\{X_n=x_n\}\right)=\prod_{i=1}^{n}\mathbb{P}(X_i=x_i).$$

Remarques

- On dit souvent indépendantes au lieu de mutuellement indépendantes.
- L'indépendance d'un n-uplet de variables aléatoires ne dépend pas de l'ordre de ces variables.

Proposition 22

Si les variables aléatoires discrètes $X_1,\dots,\ X_n$ sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$, à valeurs dans $E_1,\ E_2,\dots,\ E_n$ sont mutuellement indépendantes, alors on a, pour toutes parties A_1 de E_1, A_2 de $E_2,\dots,\ A_n$ de E_n :

$$\mathbb{P}\left(\{X_1\in A_1\}\cap\ldots\cap\{X_n\in A_n\}\right)=\prod_{i=1}^{n}\mathbb{P}(X_i\in A_i).$$

Démonstration page 933

p.933 **Exercice 12** Soit $X_1,\dots,X_n$ des variables aléatoires indépendantes de loi géométrique de paramètres respectifs $p_1,\dots,p_n$.

1. Montrer que $Y=\min(X_1,\dots,X_n)$ suit aussi une loi géométrique.
2. Déterminer la loi de $Z=\max(X_1,\dots,X_n)$.

Proposition 23

Si $(X_1,\dots,X_n)$ est une famille de variables aléatoires discrètes, mutuellement indépendantes, de l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$, alors toute sous-famille est formée de variables mutuellement indépendantes.

Démonstration page 934

Corollaire 24

Si les variables aléatoires discrètes $X_1, \dots, X_n$ de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sont mutuellement indépendantes, alors elles sont indépendantes deux à deux.

Attention Comme dans le cas fini, la réciproque est fausse.

p.934 **Exercice 13** Montrer que des événements $A_1,\dots,\ A_n$ d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sont mutuellement indépendants si, et seulement si, les variables aléatoires $\mathbb{1}_{A_1},\dots,\ \mathbb{1}_{A_n}$ sont mutuellement indépendantes.

Remarque

- Si les variables $X_1,\dots,\ X_n$ sont mutuellement indépendantes, il existe des fonctions $\varphi_1,\dots,\ \varphi_n$ définies respectivement sur $X_1(\Omega),\dots,\ X_n(\Omega)$ telles que, pour tout $(x_1,\dots,x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$,
$$\mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i = x_i\}\right) = \prod_{i=1}^{n} \varphi_i(x_i).$$
En effet, il suffit de poser, pour $1 \leqslant i \leqslant n$, $\varphi_i(x_i) = \mathbb{P}(X_i = x_i)$.
- Comme dans le cas de deux variables aléatoires (cf exercice 9 page 904), on peut démontrer que si la loi conjointe peut s'écrire sous cette forme, alors les variables $X_1,\dots,\ X_n$ sont mutuellement indépendantes.

Les propriétés démontrées pour deux variables aléatoires se généralisent à n variables aléatoires.

Proposition 25

Soit $(X_1, X_2, \dots, X_n)$ un n-uplet de variables aléatoires discrètes, mutuellement indépendantes, de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Si pour tout $i \in [\![1, n]\!]$, la fonction f_i est définie sur $X_i(\Omega)$, les variables aléatoires $f_1(X_1)$, $f_2(X_2)$, $\dots$, $f_n(X_n)$ sont mutuellement indépendantes.

Démonstration page 935

Remarque La proposition précédente se généralise à des variables aléatoires discrètes fonctions de plusieurs variables aléatoires X_k (pour $1 \leqslant k \leqslant n$) par la proposition suivante.

Proposition 26

Soit $X_1, X_2, \dots, X_n$ des variables aléatoires discrètes indépendantes à valeurs respectivement dans des ensembles $E_1, \dots, E_n$, et $1 \leqslant p < n$.
Si f est une fonction définie sur $E_1 \times \cdots \times E_p$ et g une fonction définie sur $E_{p+1} \times \cdots \times E_n$, alors $f(X_1, \dots, X_p)$ et $g(X_{p+1}, \dots, X_n)$ sont indépendantes.

Démonstration page 935

Remarque En particulier, les variables $X_1 + \cdots + X_p$ et $X_{p+1} + \cdots + X_n$ sont indépendantes.

p.935

Exercice 14 Montrer que, si X_1, X_2, ..., X_n sont des variables aléatoires mutuellement indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$, telles que X_k suive la loi de Poisson de paramètre λ_k pour tout entier k de $[\![1, n]\!]$, alors la variable aléatoire $X_1 + \cdots + X_n$ suit la loi de Poisson de paramètre $\sum_{k=1}^{n} \lambda_k$.

Ce qui précède peut se généraliser à plus de deux fonctions.

Proposition 27 (Lemme des coalitions)

Soit $(X_1, X_2, \ldots, X_n)$ un n-uplet de variables aléatoires, mutuellement indépendantes, k un entier naturel non nul, I_1, $I_2, \ldots$, I_k des sous-ensembles non vides et disjoints de $[\![1, n]\!]$. Pour tout j entre 1 et k, on considère une variable aléatoire Y_j qui est une fonction des variables X_i pour $i \in I_j$. Alors les variables $Y_1, \ldots$, Y_k sont indépendantes.

Démonstration. On montre que les k variables $(X_i)_{i\in I_1}, (X_i)_{i\in I_2}, \ldots, (X_i)_{i\in I_k}$ sont mutuellement indépendantes, comme dans la proposition 26. On en déduit que Y_1, $Y_2, \ldots$, Y_n sont mutuellement indépendantes, d'après la proposition 25 de la page précédente. □

3 Suite de variables aléatoires indépendantes

Définition 15

Une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires réelles discrètes de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ telle que pour toute partie finie I de $\mathbb{N}$, les variables aléatoires réelles discrètes X_i où i décrit I soient mutuellement indépendantes, est appelée **suite de variables aléatoires mutuellement indépendantes**.

Remarque Il suffit de vérifier que, pour tout entier n, les variables $X_0, \ldots, X_n$ sont mutuellement indépendantes. En effet, toute sous-famille finie de $(X_n)_{n\in\mathbb{N}}$ est contenue dans une sous-famille de la forme $(X_0, \ldots, X_n)$. D'après la proposition 23 de la page 906, elle est constituée de variables aléatoires mutuellement indépendantes si les variables X_0, ..., X_n les variables sont mutuellement indépendantes.

Suite de variable aléatoires de lois prescrites

Nous admettrons le théorème suivant.

Théorème 28

Pour toute suite (P_n) de lois de probabilité discrètes, il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et une suite (X_n) de variables aléatoires discrètes, indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$ telle que, pour tout $n \in \mathbb{N}$, la loi $\mathbb{P}_{X_n}$ soit P_n.

Remarques

- En particulier, si P est une loi de probabilité discrète, il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et une suite (X_n) de variables aléatoires discrètes, indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$ telle que, pour tout $n \in \mathbb{N}$, la loi de X_n soit P. On parle alors de suites de variables indépendantes identiquement distribuées (avec parfois l'abréviation i.i.d.) Une telle suite modélise une suite d'épreuves identiques aux résultats indépendants.

  Exemples

  * Si P est la loi de Bernoulli de paramètre p, on obtient une modélisation du jeu de pile ou face. Pour tout $n \in \mathbb{N}$, $\{X_n = 1\}$ et $\{X_n = 0\}$ sont des événements, qu'on appelle respectivement obtenir pile et face ou succès et échec à la n-ième épreuve ; les variables X_n étant indépendantes, les événements $\{X_n = 1\}$, pour $n \in \mathbb{N}$, sont indépendants et de probabilité p.
  * De la même façon, si P est la loi uniforme sur $[\![1,6]\!]$, on modélise l'expérience aléatoire consistant à lancer un dé une infinité de fois.

- Ce théorème est d'une grande importance pratique, car comme nous l'avons signalé dans le chapitre précédent, la construction d'un espace probabilisé modélisant le jeu de pile ou face est très difficile.

p.935 **Exercice 15** Soit (X_n) une suite de variables aléatoires indépendantes sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, suivant toutes la loi de Bernoulli de paramètre $p \in \left]0,1\right[$. Pour tout $k \in \mathbb{N}^*$, on note Y_k le temps d'attente du k-ième 1. On pose $Z_1 = Y_1$ et pour $k \geqslant 2$, $Z_k = Y_k - Y_{k-1}$.

Démontrer que $(Z_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes, de même loi.

V Espérance, variance, covariance

1 Espérance

Définition 16

Soit X une variable aléatoire discrète, à valeurs dans $\mathbb{R}_+$.
L'**espérance** de X, notée $\mathbb{E}(X)$, est la somme dans $[0, +\infty]$ de la famille de réels positifs $\big(\mathbb{P}(X = x)\, x\big)_{x \in X(\Omega)}$:

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

Remarque Si X est une variable aléatoire à valeurs dans $\mathbb{N}$, la sommabilité de la famille $\big(\mathbb{P}(X = x)\, x\big)_{x \in \mathbb{N}}$ équivaut à la convergence de la série $\sum \mathbb{P}(X = n)\, n$ et l'on a $\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)\, n$.

Définition 17

Soit X une variable aléatoire discrète réelle. On dit que X **est d'espérance finie** si la famille $\big(\mathbb{P}(X=x)\,x\big)_{x\in X(\Omega)}$ est sommable. Dans ce cas, l'espérance de X, notée $\mathbb{E}(X)$, est la somme de cette famille :

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

Remarques

- Soit X est une variable aléatoire discrète. Si X est positive, elle possède une espérance finie ou infinie. Si X est de signe quelconque et si la famille $\big(\mathbb{P}(X=x)\,x\big)_{x\in X(\Omega)}$ n'est pas sommable, X ne possède pas d'espérance.
- Si $X(\Omega)$ est fini, la variable aléatoire X est d'espérance finie et la définition de l'espérance coïncide avec celle qui a été donnée en première année.
- La différence essentielle entre le cas discret et le cas fini est que dans le cas discret l'espérance n'est pas toujours définie.

Exemples

1. Soit $a\in\mathbb{R}$.
   Si X est une variable aléatoire presque sûrement égale à a, alors $\mathbb{E}(X)=a$.
2. Pour tout événement A, on a $\mathbb{E}\big(\mathbb{1}_A\big)=\mathbb{P}(A)$.
3. En thermodynamique statistique, l'énergie moyenne d'une particule à spectre discret est égale à l'espérance de son énergie :
   $$\overline{E}=\sum_{n\in\mathbb{N}}E_n\mathcal{P}(E_n),$$
   où (E_n) est le spectre des niveaux d'énergie et $\mathcal{P}(E_n)$ est la probabilité d'occupation par la particule de l'état d'énergie E_n.

p.936 **Exercice 16** Donner des exemples de variables aléatoires discrètes :

1. ayant une espérance infinie ;
2. n'ayant pas d'espérance.

p.937 **Exercice 17** Montrer qu'une variable aléatoire discrète bornée est d'espérance finie.

Proposition 29

Soit X une variable aléatoire discrète.

1. Si X suit une loi géométrique de paramètre p, alors son espérance est finie et $\mathbb{E}(X)=\frac{1}{p}\cdot$
2. Si X suit une loi de Poisson de paramètre λ, alors son espérance est finie et $\mathbb{E}(X)=\lambda$.

Démonstration page 937

p.937

Exercice 18

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. Montrer que

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n).$$

Retrouver ainsi le premier résultat de la proposition 29 de la page précédente.

Propriétés de l'espérance

Théorème 30 (Théorème de transfert)

Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et f une fonction à valeurs réelles définie sur $X(\Omega)$. Alors la variable aléatoire $f(X)$ est d'espérance finie si, et seulement si, la famille $\big(\mathbb{P}(X = x)\, f(x)\big)_{x \in X(\Omega)}$ est sommable. On a alors :

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

Démonstration. On sait que $f(X)(\Omega) = f\left(X(\Omega)\right)$. Pour tout $y \in f(X)(\Omega)$, on pose

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

On a alors $\{f(X) = y\} = \{X \in I_y\}$ et comme I_y est au plus dénombrable, car inclus dans $X(\Omega)$:

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

L'ensemble $X(\Omega)$ est la réunion disjointe des I_y, pour $y \in f(X)(\Omega)$.

- Supposons que la famille $(\mathbb{P}(X = x)\, f(x))_{x \in X(\Omega)}$ soit sommable. On applique le théorème de sommation par paquets. Pour tout $y \in f(X)(\Omega)$, la famille $(\mathbb{P}(X = x)\, f(x))_{x \in I_y}$ est sommable, de somme :

  $$\sum_{x \in I_y} f(x)\, \mathbb{P}(X = x) = y\, \mathbb{P}(f(X) = y).$$

  On sait qu'alors la famille $\big(y\mathbb{P}(f(X) = y)\big)_{y \in f(X)(\Omega)}$ est sommable et que :

  $$\sum_{y \in f(X)(\Omega)} y\, \mathbb{P}(f(X) = y) = \sum_{x \in X(\Omega)} f(x)\, \mathbb{P}(X = x).$$

  Autrement dit, $f(X)$ est d'espérance finie et :

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

- Supposons réciproquement que la variable aléatoire $f(X)$ soit d'espérance finie, c'est-à-dire que la famille $\big(y\mathbb{P}(f(X) = y)\big)_{y \in f(X)(\Omega)}$ soit sommable. Cela équivaut à la sommabilité de la famille $\big(|y|\, \mathbb{P}(f(X) = y)\big)_{y \in f(X)(\Omega)}$.

Montrons la sommabilité de la famille $(\mathbb{P}(X = x)\, f(x))_{x\in X(\Omega)}$. Soit I une partie finie de $X(\Omega)$. Alors $J = f(I)$ est une partie finie de $f(X)(\Omega)$ et :

$$\sum_{y\in J} |y|\mathbb{P}(f(X) = y) \leqslant \sum_{y\in f(X)(\Omega)} |y|\,\mathbb{P}(f(X) = y).$$

Mais, par ailleurs, I est inclus dans $f^{-1}(J)$ donc :

$$\begin{aligned}\sum_{y\in J} |y|\,\mathbb{P}(f(X) = y) = \sum_{y\in J} |y| \sum_{x\in I_y} \mathbb{P}(X = x) &= \sum_{y\in J}\sum_{x\in I_y} |f(x)|\,\mathbb{P}(X = x)\\ &= \sum_{x\in f^{-1}(J)} |f(x)|\mathbb{P}(X = x)\\ &\geqslant \sum_{x\in I} |f(x)|\mathbb{P}(X = x).\end{aligned}$$

On a donc, pour toute partie I finie de $X(\Omega)$:

$$\sum_{x\in I} |f(x)|\mathbb{P}(X = x) \leqslant \sum_{y\in f(X)(\Omega)} |y|\mathbb{P}(f(X) = y).$$

Cela montre que la famille $(\mathbb{P}(X = x)\,|f(x)|)_{x\in X(\Omega}$ est sommable et donc que la famille $(\mathbb{P}(X = x)\, f(x))_{x\in X(\Omega}$ est sommable et, en utilisant le premier point, on retrouve l'égalité souhaitée. □

Remarque Si f est définie plus généralement sur un ensemble dénombrable E contenant $X(\Omega)$, on peut remplacer la famille $(\mathbb{P}(X = x)\, f(x))_{x\in X(\Omega)}$ par la famille $(\mathbb{P}(X = x)\, f(x))_{x\in E}$.

Théorème 31 (Linéarité de l'espérance)

Si X et Y sont deux variables aléatoires réelles discrètes, d'espérance finie, sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et λ un réel, alors les variables aléatoires $X + Y$ et λX sont d'espérance finie. De plus, on a :

$$\mathbb{E}(\lambda X) = \lambda\mathbb{E}(X) \qquad \text{et} \qquad \mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

Autrement dit, l'ensemble des variables aléatoires réelles discrètes sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ d'espérance finie est un $\mathbb{R}$-espace vectoriel, noté $L_1(\Omega, \mathcal{A}, \mathbb{P})$ et l'espérance est une forme linéaire sur cet espace vectoriel.

Principe de démonstration. On applique la formule de transfert à X et $f : x \mapsto \lambda x$, puis à (X, Y) et $f : (x, y) \mapsto x + y$.

Démonstration page 938

Définition 18

Toute variable aléatoire réelle discrète d'espérance nulle est dite **centrée**.

Corollaire 32

Si X est une variable aléatoire discrète d'espérance finie, alors $X - \mathbb{E}(X)$ est centrée.

Démonstration. La variable aléatoire constante $\mathbb{E}(X)$ a pour espérance $\mathbb{E}(X)$. Donc d'après le théorème 31, la variable aléatoire $X - \mathbb{E}(X)$ possède une espérance et :

$$\mathbb{E}\big(X - \mathbb{E}(X)\big) = \mathbb{E}(X) - \mathbb{E}(X) = 0. \qquad \square$$

Point méthode

Pour déterminer l'espérance d'une variable aléatoire discrète dont on ne sait pas déterminer la loi, on peut la décomposer en somme de variables aléatoires discrètes dont l'espérance est connue.

p.939 **Exercice 19** Soit $n \in \mathbb{N}^*$. On considère une suite $(X_k)_{k\in\mathbb{N}^*}$ de variables aléatoires indépendantes sur le même espace probabilisé, toutes de loi uniforme sur $[\![1, n]\!]$. On cherche à calculer le nombre moyen de tirages nécessaires pour obtenir tous les numéros de 1 à n.

1. Pour $1 \leqslant k \leqslant n$, on note T_k le nombre de tirages nécessaires pour obtenir k numéros distincts et l'on pose $Z_1 = T_1$ et $Z_k = T_k - T_{k-1}$ pour $2 \leqslant k \leqslant n$.

   Montrer que les variables aléatoires Z_k pour $2 \leqslant k \leqslant n$ suivent une loi géométrique dont on précisera le paramètre.
2. En déduire l'espérance de T_n et un équivalent de $\mathbb{E}(T_n)$ quand n tend vers $+\infty$.

Proposition 33 (Positivité de l'espérance)

Si X une variable aléatoire discrète, à valeurs dans $\mathbb{R}_+$, alors on a :

$$\mathbb{E}(X) \geqslant 0.$$

De plus $\mathbb{E}(X) = 0$ équivaut à $X = 0$ presque sûrement.

Démonstration page 940

Corollaire 34 (Croissance de l'espérance)

Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, d'espérance finie, telles que $X \leqslant Y$, alors on a :

$$\mathbb{E}(X) \leqslant \mathbb{E}(Y).$$

De plus $\mathbb{E}(X) = \mathbb{E}(Y)$ équivaut à $X = Y$ presque sûrement.

Démonstration. On applique la proposition 33 à la variable aléatoire $X - Y$ et on utilise la linéarité de l'espérance. $\square$

Corollaire 35

Si X est une variable aléatoire discrète, alors X est d'espérance finie si, et seulement si, $|X|$ est d'espérance finie. Si les espérances sont finies, on a :

$$|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|).$$

Démonstration page 940

Proposition 36

Soit X et Y sont deux variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ telles que $|X| \leqslant Y$. Si Y est d'espérance finie, alors X est d'espérance finie.

Principe de démonstration. Considérer $Z = (X, Y)$, ainsi que les projections $\pi_1 : (x, y) \mapsto x$ et $\pi_2 : (x, y) \mapsto y$ et appliquer le théorème de transfert. Démonstration page 940

Proposition 37

Soit X et Y sont deux variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$. Alors XY est d'espérance finie si, et seulement si, la famille $\left(xy\,\mathbb{P}\left((X,Y) = (x,y)\right)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable et l'on a alors :

$$\mathbb{E}(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy\,\mathbb{P}\left((X,Y) = (x,y)\right).$$

Démonstration page 941

Théorème 38

Si X et Y sont deux variables aléatoires discrètes réelles indépendantes sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, d'espérance finie, alors XY est d'espérance finie et :

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

Démonstration page 941

p.941 **Exercice 20**

Soit X et Y deux variables aléatoires discrètes sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Démontrer que X et Y sont indépendantes si, et seulement si, pour toutes fonctions numériques f et g bornées, définies respectivement sur $X(\Omega)$ et $Y(\Omega)$, on a :

$$\mathbb{E}\big(f(X)g(Y)\big) = \mathbb{E}\big(f(X)\big)\mathbb{E}\big(g(Y)\big).$$

Théorème 39 (Inégalité de Markov)

Toute variable aléatoire discrète positive, d'espérance finie, vérifie l'inégalité :

$$\forall a > 0 \quad \mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}.$$

Démonstration page 942

2 Variance

Définition 19

Soit X une variable aléatoire réelle discrète et $r \in \mathbb{N}$. Si X^r est d'espérance finie, on dit que X admet un **moment d'ordre r**. Le moment d'ordre r de X est le réel $\mathbb{E}(X^r)$.

Remarques

- D'après la formule de transfert, si le moment d'ordre r existe, on a :
$$\mathbb{E}(X^r) = \sum_{x \in X(\Omega)} x^r \, \mathbb{P}(X = x).$$
- Une variable aléatoire bornée admet des moments de tout ordre.

p.942

Exercice 21 Soit X une variable aléatoire réelle discrète, positive, possédant un moment d'ordre $r \geqslant 1$.
Montrer que, pour tout $a > 0$, on a $\mathbb{P}(X \geqslant a) \leqslant \dfrac{\mathbb{E}(X^r)}{a^r}$.

Proposition 40

Soit X une variable aléatoire réelle discrète. Si X admet un moment d'ordre 2, alors X est d'espérance finie.

Démonstration page 942

p.942

Exercice 22 Démontrer que la variable aléatoire réelle discrète X admet un moment d'ordre r, elle admet un moment d'ordre k pour tout $k \in [\![0, r]\!]$.

Lemme 41

Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, admettant un moment d'ordre 2, alors XY est d'espérance finie.

Démonstration page 942

Proposition 42

Les variables aléatoires réelles discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$ admettant un moment d'ordre 2 forment un $\mathbb{R}$-espace vectoriel, noté $L_2(\Omega, \mathcal{A}, \mathbb{P})$.

Démonstration page 942

Théorème 43 (Inégalité de Cauchy-Schwarz)

Si les variables aléatoires réelles discrètes X et Y appartiennent à $L_2(\Omega, \mathcal{A}, \mathbb{P})$, alors XY est d'espérance finie et :
$$\mathbb{E}(XY)^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2).$$

Démonstration page 942

p.943

Exercice 23 Étudier les cas d'égalité dans l'inégalité de Cauchy-Schwarz.

Définition 20

Si la variable aléatoire réelle discrète X est d'espérance finie et si la variable aléatoire $X-\mathbb{E}(X)$ admet une moment d'ordre 2, on appelle **variance** de X le réel $\mathbb{V}(X)$ défini par :

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

Remarques

- La variance est la moyenne du carré de la distance entre les valeurs de X et $\mathbb{E}(X)$. Elle mesure donc la dispersion de X par rapport à sa moyenne.
- Comme la variable $X - \mathbb{E}(X)$ est centrée, la variance est aussi appelée moment centré d'ordre 2.
- Une variable aléatoire réelle finie possède toujours une variance.

p.943

Exercice 24 Soit X une variable aléatoire réelle discrète possédant un moment d'ordre 2. À quelle condition la variance de X est-elle nulle ?

Théorème 44 (Formule de Kœnig-Huygens)

Soit X une variable aléatoire réelle discrète. La variable X admet une variance si, et seulement si, X admet un moment d'ordre 2 et l'on a alors :

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

Démonstration page 943

Remarque Dans la quasi-totalité des cas, on utilise la formule de Kœnig-Huygens pour calculer une variance et non la définition.

Proposition 45

Soit X une variable aléatoire discrète.

1. Si X suit une loi géométrique de paramètre p, alors elle possède une variance et $\mathbb{V}(X) = \dfrac{1-p}{p^2}\cdot$
2. Si X suit une loi de Poisson de paramètre λ, alors elle possède une variance et $\mathbb{V}(X) = \lambda$.

Principe de démonstration. Dans les deux cas, commencer par calculer $\mathbb{E}\big(X(X-1)\big)$.

Démonstration page 944

Proposition 46

Si (a, b) est un couple de réels et X une variable aléatoire réelle discrète admettant un moment d'ordre 2, alors $aX+b$ admet un moment d'ordre 2 et :

$$\mathbb{V}(aX+b) = a^2\mathbb{V}(X).$$

Démonstration page 944

Théorème 47 (Inégalité de Bienaymé-Tchebychev)

Toute variable aléatoire réelle discrète X admettant un moment d'ordre 2 vérifie l'inégalité :

$$\forall \varepsilon > 0 \quad \mathbb{P}\Big(|X - \mathbb{E}(X)| \geqslant \varepsilon\Big) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Démonstration page 944

Définition 21

Si X est une variable aléatoire réelle discrète admettant un moment d'ordre 2, **l'écart-type de X** est le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.

Exemple En thermodynamique statistique, l'écart quadratique moyen en énergie d'une particule à spectre discret d'énergie est l'écart-type de son énergie :

$$\Delta E = \sqrt{\sum_{n\in\mathbb{N}} (E_n - \overline{E})^2 \mathcal{P}(E_n)},$$

où (E_n) est le spectre des niveaux d'énergie, $\mathcal{P}(E_n)$ est la probabilité d'occupation par la particule de l'état d'énergie E_n et $\overline{E}$ l'énergie moyenne de la particule.

Définition 22

Soit X variable aléatoire réelle discrète admettant un moment d'ordre 2. Si $\mathbb{E}(X) = 0$ et $\sigma(X) = 1$, la variable aléatoire X est dite **centrée réduite**.

Proposition 48

Si X est une variable aléatoire réelle discrète X admettant une variance non nulle, la variable aléatoire réelle discrète $X^* = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$ est une variable aléatoire réelle discrète centrée réduite, appelée **la variable aléatoire réelle centrée réduite associée à X**.

Démonstration page 945

3 Covariance

Définition 23

Soit X et Y deux variables aléatoires réelles discrètes d'espérance finie. Si la variable $(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))$ est d'espérance finie, on appelle **covariance** de X et Y (ou du couple (X,Y)) le réel noté $\mathrm{Cov}(X,Y)$ défini par :

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

Théorème 49 (Formule de Kœnig-Huygens)

Soit X et Y deux variables aléatoires discrètes sur le même espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$. Si X, Y et XY admettent une espérance, alors le couple (X,Y) admet une covariance donnée par la formule :

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

La covariance de (X,Y) existe en particulier si X et Y ont un moment d'ordre 2.

Démonstration page 945

Corollaire 50

Si X et Y sont deux variables aléatoires réelles discrètes, indépendantes, possédant une espérance finie, alors le couple (X,Y) possède une covariance et $\mathrm{Cov}(X,Y)=0$.

Démonstration. Cela résulte directement du théorème 38 de la page 914. □

Définition 24

Si un couple (X,Y) de variables aléatoires réelles discrètes possède une covariance et vérifie $\mathrm{Cov}(X,Y)=0$, on dit que les variables aléatoires X et Y sont **non corrélées**.

Remarque Le corollaire 50 montre que deux variables aléatoires réelles discrètes, indépendantes, sont non corrélées.
Comme dans le cas des variables finies, la réciproque est fausse. Deux variables non corrélées ne sont pas nécessairement indépendantes.

p.945

Exercice 25 Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables de Bernoulli indépendantes, de paramètre $p\in\,]0,1[$. On pose $q=1-p$. On dit que la première série est de longueur n si les premières épreuves ont donné le même résultat et la $(n+1)$-ième un résultat différent. De même la deuxième série commence à l'épreuve qui suit la dernière épreuve de la première série et s'arrête avant le changement suivant. On note L_1 et L_2 respectivement la longueur de la première et de la seconde série.

Démontrer que $\mathrm{Cov}(L_1,L_2)=-\dfrac{(p-q)^2}{pq}$.

Proposition 51

Pour tout $X \in L_2(\Omega, \mathcal{A}, \mathbb{P})$, on a $\mathrm{Cov}(X,X) = \mathbb{V}(X)$.
L'application $(X,Y) \longmapsto \mathrm{Cov}(X,Y)$ est une forme bilinéaire symétrique positive sur $L_2(\Omega, \mathcal{A}, \mathbb{P})$.

Démonstration page 946

p.947 **Exercice 26** Soit X_1, $X_2, \ldots,$ X_n des variables aléatoires discrètes sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, possédant toutes un moment d'ordre 2. On considère la matrice $M = \big(\mathrm{Cov}(X_i, X_j)\big)_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$.

Montrer que M est diagonalisable et a des valeurs propres positives ou nulles.

À quelle condition 0 est-il valeur propre de M ?

Variance d'une somme

Théorème 52

Pour toute famille $(X_1, \ldots, X_n)$ de variables aléatoires réelles discrètes appartenant à $L_2(\Omega, \mathcal{A}, \mathbb{P})$, la variable aléatoire réelle $X_1 + X_2 + \cdots + X_n$ admet une variance :

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{k=1}^{n} \mathbb{V}(X_k) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{Cov}(X_i, X_j).$$

Démonstration page 947

La covariance de deux variables aléatoires indépendantes étant nulle, on en déduit le corollaire suivant.

Corollaire 53

Pour toute famille $(X_1, \ldots, X_n)$ de variables aléatoires réelles discrètes appartenant à $L_2(\Omega, \mathcal{A}, \mathbb{P})$, deux à deux indépendantes, on a :

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{k=1}^{n} \mathbb{V}(X_k).$$

Remarques

- La propriété est vérifiée en particulier si les variables sont mutuellement indépendantes.
- Cette formule montre l'intérêt de décomposer une variable aléatoire en somme de variables aléatoires indépendantes plus simples.

Exemple Soit (X_n) une suite de variables aléatoires mutuellement indépendantes, suivant toutes la loi de Bernoulli de paramètre $p \in \,]0,1[$. Pour tout $k \in \mathbb{N}^*$, on note Y_k le temps d'attente du k-ième 1. On pose $Z_1 = Y_1$ et $Z_k = Y_k - Y_{k-1}$

pour $k \geqslant 2$. Il a été démontré dans l'exercice 15 de la page 909 que $(Z_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes, suivant la loi $\mathcal{G}(p)$ et donc d'espérance $\dfrac{1}{p}$ et de variance $\dfrac{1-p}{p^2}\cdot$

De l'égalité $Y_k = \sum\limits_{i=1}^{k} Z_i$, on déduit $\mathbb{E}(Y_k) = \dfrac{k}{p}$ et $\mathbb{V}(Y_k) = \dfrac{k(1-p)}{p^2}\cdot$

4 Loi faible des grands nombres

Théorème 54 (Loi faible des grands nombres)

Soit (X_n) une suite de variables aléatoires discrètes sur le même espace probabilisé, deux à deux indépendantes, de même loi, admettant un moment d'ordre 2. On pose $m = \mathbb{E}(X_1)$ et pour tout $n \in \mathbb{N}^*$, $S_n = \sum\limits_{k=1}^{n} X_k$. On a alors :

$$\forall \varepsilon > 0 \qquad \lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = 0.$$

Démonstration page 947

Remarque La loi faible des grands nombres exprime le fait qu'en un certain sens la suite $\left(\dfrac{S_n}{n}\right)$ converge vers m.

Exemple On considère une suite d'épreuves indépendantes. Soit A un événement fixé, de probabilité p. Pour $n \in \mathbb{N}^*$, on considère X_n la variable de Bernoulli qui vaut 1 si l'événement A est réalisé lors de la n-ième épreuve ; elle est de paramètre $p = \mathbb{P}(A)$. Notons Y_n la fréquence de réalisation de l'événement A, c'est-à-dire $Y_n = \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k$. On a, pour tout $\varepsilon > 0$, $\lim\limits_{n\to+\infty} \mathbb{P}\big(|Y_n - p| \geqslant \varepsilon\big) = 0$. La loi faible des grands nombres s'accorde avec l'intuition qui voit la probabilité d'un événement comme la limite d'une fréquence.

Approfondissement

- On peut obtenir le même résultat en supposant simplement que les variables X_n ont une espérance finie, et pas un moment d'ordre 2 (*cf.* exercice 16.20 de la page 959).
- Si ce théorème s'appelle loi *faible* des grands nombres, c'est qu'il existe des lois *fortes* des grands nombres. On peut démontrer par exemple que, si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires, mutuellement indépendantes, de même loi, d'espérance finie m, alors il existe un événement

négligeable A, tel que :

$$\forall \omega \in \Omega \setminus A \qquad \lim_{n \to +\infty} \frac{S_n(\omega)}{n} = m.$$

On dit que la suite $\left(\dfrac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge presque sûrement vers m (*cf.* exercice 16.26 de la page 963 pour une démonstration dans un cas particulier).

VI Fonctions génératrices

Définition 25

Pour tout variable aléatoire X à valeurs dans $\mathbb{N}$, on appelle **fonction génératrice** de X la fonction définie G_X définie par $G_X(t) = \mathbb{E}(t^X)$. D'après la formule de transfert, $G_X(t)$ est défini si, et seulement si, la série $\sum \mathbb{P}(X = n)t^n$ converge absolument et on a alors :

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

Proposition 55

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.
La série entière $\sum \mathbb{P}(X = n)t^n$ est de rayon de convergence supérieur ou égal à 1 ; elle converge normalement sur le disque fermé de centre 0 et de rayon 1. La fonction G_X est définie et continue sur $[-1, 1]$.

Démonstration page 948

Remarques

- D'après les propriétés des séries entières, la fonction G_X est aussi définie et $\mathcal{C}^\infty$ sur $]-R, R[$, où R est le rayon de convergence de la série entière.
- Si X est une variable finie, alors G_X est définie sur $\mathbb{R}$ et c'est une fonction polynomiale.

Proposition 56

La loi d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est déterminée de manière unique par G_X. Plus précisément, on a :

$$\forall n \in \mathbb{N} \quad \mathbb{P}(X = n) = n!\, G_X^{(n)}(0).$$

Deux variables aléatoires à valeurs dans $\mathbb{N}$ ont donc même loi si, et seulement si, elles ont même fonction génératrice.

Démonstration page 948

Lemme 57

Soit (a_n) une suite à termes positifs telle la série $\sum a_n$ converge. On note f la somme de la série entière $\sum a_n t^n$.

La fonction f est dérivable en 1 si, et seulement si, la série de terme général na_n converge et l'on a alors :

$$f'(1) = \sum_{n=1}^{+\infty} na_n.$$

Démonstration page 948

Théorème 58

Une variable aléatoire X à valeurs dans $\mathbb{N}$ est d'espérance finie si, et seulement si, G_X est dérivable en 1 et l'on a alors :

$$\mathbb{E}(X) = G'_X(1).$$

Démonstration page 949

Remarque Il résulte de la démonstration du lemme que G_X est dérivable en 1 si, et seulement si, G_X est dérivable à gauche en 1. Il suffit donc, pour montrer que X est d'espérance finie, de considérer la restriction de G_X à $[0,1]$.

Théorème 59

Une variable aléatoire X à valeurs dans $\mathbb{N}$ possède un moment d'ordre 2 si, et seulement si, G_X est deux fois dérivable en 1 et l'on a alors : $G''_X(1) = \mathbb{E}(X(X-1))$.

Démonstration page 949

Corollaire 60

Si la fonction génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est deux fois dérivable en 1, alors X possède un moment d'ordre 2 et on a :

$$\mathbb{E}(X^2) = G''_X(1) + G'_X(1).$$

Remarques

- L'existence et le calcul de l'espérance et de la variance d'une variable aléatoire ne posent aucun problème quand la série définissant la fonction génératrice a un rayon de convergence strictement supérieur à 1.
- Par récurrence sur $r \geqslant 1$, on peut démonter que si X une variable aléatoire à valeurs dans $\mathbb{N}$, alors G_X est r fois dérivable en 1 si, et seulement si, X possède un moment d'ordre r et qu'alors on a :

$$G_X^{(r)}(1) = \mathbb{E}\left(X(X-1)\dots(X-r+1)\right).$$

Fonctions génératrices des lois usuelles

p.949 **Exercice 27** Déterminer la fonction génératrice d'une variable aléatoire qui suit :

1. la loi de Bernoulli de paramètre p ;
2. la loi binomiale de paramètre (n,p) ;
3. la loi géométrique de paramètre $p \in]0,1[$;
4. la loi de Poisson de paramètre λ.

Retrouver ainsi dans chaque cas son espérance et sa variance.

Théorème 61

Soit X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. Alors, pour tout réel t tel que $G_X(t)$ et $G_Y(t)$ sont définis, $G_{X+Y}(t)$ est défini et :

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

Démonstration page 950

Corollaire 62

Soit X_1, ..., X_n des variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. On pose $X = \sum\limits_{k=1}^{n} X_k$. Alors, pour tout réel t, tel que $G_{X_k}(t)$ est défini pour tout $k \in [\![1,n]\!]$, $G_X(t)$ est défini et :

$$G_X(t) = \prod_{k=1}^{n} G_{X_k}(t).$$

L'utilisation de la fonction de répartition peut permettre de déterminer la loi d'une variable aléatoire difficile à déterminer directement.

p.950 **Exercice 28** Soit ℓ, m et n des entiers naturels non nuls tels que $n = \ell m$, ainsi que X et Y des variables aléatoires indépendantes à valeurs dans $\mathbb{N}$; on pose $Z = X + Y$. On suppose que X suit la loi uniforme sur $[\![0,\ell-1]\!]$ et que Z suit la loi uniforme sur $[\![0,n-1]\!]$.

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

p.951 **Exercice 29** À l'aide du théorème 61, retrouver les théorèmes de stabilité de la loi binomiale et de la loi de Poisson.

Démonstrations et solutions des exercices du cours

Proposition 1 On a :

$$X^{-1}(A) = X^{-1}\big(A \cap X(\Omega)\big) = \bigcup_{x \in A \cap X(\Omega)} X^{-1}\big(\{x\}\big).$$

Chaque $X^{-1}(\{x\})$ appartient à $\mathcal{A}$ et $A \cap X(\Omega)$ est au plus dénombrable, car inclus dans $X(\Omega)$ qui est au plus dénombrable. Ainsi $X^{-1}(A)$ est un événement, car c'est une réunion au plus dénombrable d'événements.

Théorème 2

- L'application $\mathbb{P}_X$ est définie sur $\mathcal{P}\big(X(\Omega)\big)$ et à valeurs dans $[0,1]$.
- De $\{X \in X(\Omega)\} = \Omega$, on déduit $\mathbb{P}_X\left(X(\Omega)\right) = 1$.
- Si (A_n) est une suite de parties deux à deux disjointes de $X(\Omega)$, alors les ensembles $\{X \in A_n\}$ sont également deux à deux disjoints et :

$$\left\{X \in \bigcup_{n \in \mathbb{N}} A_n\right\} = \bigcup_{n \in \mathbb{N}} \{X \in A_n\}.$$

  On en déduit

$$\begin{aligned}\mathbb{P}_X\left(\bigcup_{n \in \mathbb{N}} A_n\right) &= \mathbb{P}\left(\left\{X \in \bigcup_{n \in \mathbb{N}} A_n\right\}\right) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} \{X \in A_n\}\right)\\ &= \sum_{n \in \mathbb{N}} \mathbb{P}\big(\{X \in A_n\}\big) = \sum_{n \in \mathbb{N}} \mathbb{P}_X(A_n).\end{aligned}$$

Proposition 3

- L'ensemble $X(\Omega)$ est au plus dénombrable.
- Par définition, $\{X = x\}$ est un événement pour tout $x \in X(\Omega)$.
- Si x et x' sont deux éléments de $X(\Omega)$, distincts, les événements $\{X = x\}$ et $\{X = x'\}$ sont incompatibles, car pour tout $\omega \in \Omega$, $\omega \in \{X = x\}$ équivaut à $X(\omega) = x$.
- On a enfin :

$$\bigcup_{x \in X(\Omega)} \{X = x\} = \Omega$$

  car tout élément ω de Ω appartient à $\{X = x\}$ pour $x = X(\omega)$.

Proposition 5 On prend $\Omega = E$ et $\mathcal{A} = \mathcal{P}(E)$. D'après le théorème 9 de la page 854, il existe une unique probabilité $\mathbb{P}$ sur cet espace probabilisé telle que $\mathbb{P}\left(\{x\}\right) = p_x$ pour tout $x \in E$. On considère l'application $X = \mathrm{Id}_E$. On a alors $X(\Omega) = E$, donc $X(\Omega)$ est au plus dénombrable et, pour tout $x \in E$, $X^{-1}\left(\{x\}\right) = \{x\} \in \mathcal{A}$. Donc X est une variable aléatoire discrète et pour tout $x \in E$:

$$\mathbb{P}(X = x) = \mathbb{P}\left(X^{-1}\left(\{x\}\right)\right) = \mathbb{P}\left(\{x\}\right) = p_x.$$

Exercice 1

1. La suite $(\{X \leqslant n\})_{n\in\mathbb{N}}$ est une suite croissante d'événements. On a donc :
$$\lim_{n\to+\infty} \mathbb{P}(X \leqslant n) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{X \leqslant n\}\right) = \mathbb{P}(\Omega) = 1.$$
La fonction $F : x \mapsto \mathbb{P}(X \leqslant x)$ est croissante. En effet si $x \leqslant y$, alors $\{X \leqslant x\} \subset \{X \leqslant y\}$ et donc $F(x) \leqslant F(y)$ par croissance de $\mathbb{P}$.
Elle possède donc une limite en $+\infty$ et :
$$\lim_{x\to+\infty} \mathbb{P}(X \leqslant x) = \lim_{x\to+\infty} F(x) = \lim_{n\to+\infty} F(n) = 1.$$
2. De même, la suite $(\{X \leqslant -n\})_{n\in\mathbb{N}}$ est une suite décroissante d'événements. On a donc :
$$\lim_{n\to+\infty} \mathbb{P}(X \leqslant -n) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} \{X \leqslant -n\}\right) = \mathbb{P}(\varnothing) = 0.$$
On en déduit :
$$\lim_{x\to-\infty} \mathbb{P}(X \leqslant x) = \lim_{x\to-\infty} F(x) = \lim_{n\to+\infty} F(-n) = 0.$$

Proposition 6 L'ensemble $f(X)(\Omega) = f(X(\Omega))$ est au plus dénombrable, car c'est l'image par f d'un ensemble au plus dénombrable. Pour tout $y \in f(X)(\Omega)$, on a :
$$\begin{aligned}\{f(X) = y\} &= \big\{\omega \in \Omega \mid X(\omega) \in f^{-1}(\{y\})\big\}\\ &= \big\{X \in f^{-1}(\{y\})\big\} = \bigcup_{x\in f^{-1}(\{y\})} \{X = x\}.\end{aligned}$$
L'ensemble $f^{-1}\big(\{y\}\big)$ est au plus dénombrable, car c'est une partie de $X(\Omega)$. On en déduit que $\{f(X) = y\} \in \mathcal{A}$ car c'est une union au plus dénombrable d'éléments de $\mathcal{A}$. Ainsi, $f(X)$ est une variable aléatoire discrète. De plus, puisqu'on a une union au plus dénombrable d'événements incompatibles :
$$\forall y \in f(X)(\Omega) \quad \mathbb{P}\big(f(X) = y\big) = \sum_{x\in f^{-1}(\{y\})} \mathbb{P}(X = x).$$

Proposition 7 Pour tout $k \in \mathbb{N}$,
$$\mathbb{P}(X > k) = \sum_{n=k+1}^{+\infty} \mathbb{P}(X = n) = \sum_{n=k+1}^{+\infty} pq^{n-1} = \frac{pq^k}{1-q} = q^k.$$

Corollaire 8 Comme $\{X > k + \ell\} \subset \{X > k\}$, on a :
$$\mathbb{P}\big(X > k + \ell \mid X > k\big) = \frac{\mathbb{P}\big(X > k + \ell\big)}{\mathbb{P}\big(X > k\big)} = \frac{q^{k+\ell}}{q^\ell} = q^k = \mathbb{P}\big(X > k\big).$$

Proposition 9 Posons, pour tout $n \in \mathbb{N}$, $u_n = \mathbb{P}(X > n)$. Pour tout $n \in \mathbb{N}$, on a $u_n \neq 0$ car, par hypothèse, la probabilité conditionnelle à $\{X > n\}$ est définie. On obtient, pour tout $n \in \mathbb{N}$:
$$u_1 = \mathbb{P}\big(X > n + 1 \mid X > n\big) = \frac{\mathbb{P}(X > n + 1)}{\mathbb{P}(X > n)} = \frac{u_{n+1}}{u_n}.$$

La suite (u_n) est géométrique de raison u_1, donc $u_n = u_1 \times u_1^{n-1} = u_1^n$, pour tout $n \geqslant 1$. Cela reste vrai pour $n = 0$, car $u_0 = \mathbb{P}(X > 0) = 1$. On en déduit, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(X = n) = \mathbb{P}(X > n-1) - \mathbb{P}(X > n) = u_{n-1} - u_n = u_1^{n-1} - u_1^n = u_1^{n-1}(1 - u_1).$$

On sait que $u_1 \neq 0$. D'autre part, $u_1 \neq 1$, car sinon $\mathbb{P}(X = n) = 0$ pour tout $n \geqslant 1$, ce qui est impossible car X est à valeurs dans $\mathbb{N}^*$. Ainsi $1 - u_1 \in]0,1[$ et X suit la loi géométrique de paramètre $1 - u_1$.

Exercice 2 Pour $n \in \mathbb{N}^*$, on note A_n l'événement « la n-ième épreuve est un succès ».
Soit $k \in \mathbb{N}^*$. Si $k < r$, alors on a $\{Y_r = k\} = \varnothing$ et si $k \geqslant r$, alors on a :

$$\{Y_r = k\} = \bigcup_{\substack{I \subset [\![1,k-1]\!] \\ \operatorname{card} I = r-1}} \left(\bigcap_{i \in I} A_i \right) \cap \left(\bigcap_{i \in [\![1,k-1]\!] \setminus I} \overline{A_i} \right) \cap A_r.$$

En effet, la k-ième épreuve correspond au r-ième succès et dans les $k-1$ premières épreuves, il doit y avoir $r-1$ succès (et donc $k-r$ échecs). Ainsi $\{Y_r = k\} \in \mathcal{A}$ et

$$\mathbb{P}(Y_r = k) = \binom{k-1}{r-1} p^{r-1} q^{k-r} p = \binom{k-1}{r-1} p^r q^{k-r},$$

car c'est la réunion disjointe de $\binom{k-1}{r-1}$ événements de probabilité $p^r q^{k-r}$.

On a $\{Y_r = \infty\} = \bigcup_{k \in \mathbb{N}^*} \overline{\{Y_r = k\}}$, donc $\{Y_r = \infty\}$ est un événement. Il est inclus dans l'événement H : « il existe un rang à partir duquel toutes les épreuves donnent des échecs » qui peut s'écrire $H = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geqslant n} \overline{A_k}$. On a, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}\left(\bigcap_{k \geqslant n} \overline{A_k} \right) = \lim_{N \to +\infty} \mathbb{P}\left(\bigcap_{k=n}^{N} \overline{A_k} \right) = \lim_{N \to +\infty} q^{N-n+1} = 0.$$

L'événement H qui est une réunion dénombrable d'événements négligeables est donc négligeable. *A fortiori* l'événement $\{Y_r = \infty\}$ est négligeable.

Proposition 10 Pour tout entier k et pour tout entier $n \geqslant k$, on a :

$$\mathbb{P}\big(X_n = k\big) = \binom{n}{k} p_n^k \, (1 - p_n)^{n-k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \, p_n^k \exp\big((n-k)\ln(1-p_n)\big).$$

On a, quand n tend vers $+\infty$, $n(n-1)\cdots(n-k+1) \sim n^k$ et $p_n \sim \dfrac{\lambda}{n}\cdot$

Ainsi (p_n) tend vers 0 et donc :

$$(n-k)\ln(1-p_n) \sim n(-p_n) \to -\lambda.$$

On en déduit :

$$\mathbb{P}\big(X_n = k\big) \sim \frac{n^k}{k!} \left(\frac{\lambda}{n}\right)^k e^{-\lambda} \sim e^{-\lambda} \frac{\lambda^k}{k!}\cdot$$

On obtient $\lim\limits_{n \to +\infty} \mathbb{P}\big(X_n = k\big) = e^{-\lambda} \dfrac{\lambda^k}{k!}\cdot$

Proposition 11

- Si X et Y sont deux variables aléatoires, l'image $(X,Y)(\Omega)$ est la partie du produit cartésien $X(\Omega)\times Y(\Omega)$ formée des couples $\big(X(\omega),Y(\omega)\big)$ où ω décrit Ω. Le produit cartésien $X(\Omega)\times Y(\Omega)$ est au plus dénombrable car $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables. *A fortiori*, $(X,Y)(\Omega)$ est au plus dénombrable.

  Pour tout $(x,y)\in(X,Y)(\Omega)$, l'ensemble $\{(X,Y)=(x,y)\}=\{X=x\}\cap\{Y=y\}$ appartient à $\mathcal{A}$, car c'est l'intersection de deux éléments de $\mathcal{A}$. Donc (X,Y) est une variable aléatoire à valeurs dans $E\times E'$.
- Si Z est une variable aléatoire à valeurs dans $E\times E'$, on pose $X=\pi_1(Z)$ et $Y=\pi_2(Z)$, où $\pi_1:(x,y)\mapsto x$ et $\pi_2:(x,y)\mapsto y$ sont les projections canoniques de $E\times E'$. Alors X et Y sont des variables aléatoires discrètes à valeurs respectivement dans E et E' et $Z=(X,Y)$.

Proposition 12

- Pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$, $\{X=x\}\cap\{Y=y\}$ est un événement.
- Si (x,y) et (x',y') sont des éléments distincts de $X(\Omega)\times Y(\Omega)$, on a soit $x\neq x'$ et $\{X=x\}\cap\{X=x'\}=\varnothing$, soit $y\neq y'$ et $\{Y=y\}\cap\{Y=y'\}=\varnothing$. On en déduit que $\big(\{X=x\}\cap\{Y=y\}\big)\cap\big(\{X=x'\}\cap\{Y=y'\}\big)=\varnothing$.
- On a enfin $\displaystyle\bigcup_{(x,y)\in X(\Omega)\times Y(\Omega)}\{X=x\}\cap\{Y=y\}=\Omega$ car tout élément ω de Ω appartient à $\{X=x\}\cap\{Y=y\}$ pour $x=X(\omega)$ et $y=Y(\omega)$.

Proposition 13 La fonction nulle est une variable aléatoire discrètes. Si X et Y sont des variables aléatoires réelles sur $(\Omega,\mathcal{A})$ et λ un réel, posons $Z=(X,Y)$. Alors Z est une variable aléatoire discrète à valeurs dans $\mathbb{R}^2$ et $\lambda X+Y=f(Z)$, où f est l'application $(x,y)\mapsto\lambda x+y$, donc $\lambda X+Y$ est une variable aléatoire réelle discrète, d'après la proposition 6 de la page 895. Ainsi, l'ensemble des variables aléatoires réelles discrètes sur $(\Omega,\mathcal{A})$ est un sous-espace vectoriel de l'ensemble des applications de Ω dans $\mathbb{R}$.

En gardant les mêmes notations, on a $XY=g(Z)$, où g est l'application $(x,y)\mapsto xy$, donc XY est une variable aléatoire réelle.

Exercice 3 On sait que $\displaystyle\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty}\mathbb{P}\big(\{X=i\}\cap\{Y=j\}\big)=1$. Pour tout $n\in\mathbb{N}$, on a :

$$\sum_{i+j=n}\frac{(i+j)\lambda^{i+j}}{i!\,j!}=n\lambda^n\sum_{i+j=n}\frac{1}{i!j!}=\frac{n\lambda^n}{n!}\sum_{i=0}^{n}\binom{n}{i}=\frac{n(2\lambda)^n}{n!}.$$

D'autre part, on a :

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

D'après le théorème de sommation par paquets, on a :

$$\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty}\frac{(i+j)\lambda^{i+j}}{i!\,j!}=\sum_{n=0}^{+\infty}\sum_{i+j=n}\frac{(i+j)\lambda^{i+j}}{i!\,j!}=(2\lambda)e^{2\lambda}\qquad\text{et}\qquad a=\frac{e^{-2\lambda}}{2\lambda}\cdot$$

Exercice 4 Pour tout $i \in \mathbb{N}$, on a :

$$\begin{aligned}\mathbb{P}(X=i)=\sum_{j=0}^{+\infty}\mathbb{P}\big(\{X=i\}\cap\{Y=j\}\big) &= a\sum_{j=0}^{+\infty}\frac{(i+j)\lambda^{i+j}}{i!\,j!}\\ &= a\frac{i\lambda^i}{i!}\sum_{j=0}^{+\infty}\frac{\lambda^j}{j!}+a\frac{\lambda^i}{i!}\sum_{j=1}^{+\infty}\frac{\lambda^j}{(j-1)!}\\ &= a\frac{i\lambda^i e^\lambda}{i!}+a\frac{\lambda^{i+1}e^\lambda}{i!}=\frac{ae^\lambda\lambda^i(i+\lambda)}{i!}\cdot\end{aligned}$$

Comme $a=\dfrac{e^{-2\lambda}}{2\lambda}$, on obtient $\mathbb{P}(X=i)=\dfrac{e^{-\lambda}\lambda^{i-1}(i+\lambda)}{2\,i!}$ et, par symétrie de la loi conjointe, $\mathbb{P}(Y=j)=\dfrac{e^{-\lambda}\lambda^{j-1}(ij+\lambda)}{2\,j!}\cdot$

Exercice 5

1. Soit $(i,j)\in\mathbb{N}^2$. On a :

$$\mathbb{P}\big(\{X=i\}\cap\{Y=j\}\big)=\begin{cases}0 & \text{si} \quad i\geqslant j\\ p^2(1-p)^{j-2} & \text{si} \quad i<j.\end{cases}$$

   En effet, il faut avoir des succès à la i-ème et à la j-ème épreuve et des échecs jusqu'à la $(i-1)$-ième épreuve et entre la i-ème et la j-ème.

2. Pour tout $i\in\mathbb{N}^*$, on a :

$$\begin{aligned}\mathbb{P}(X=i)=\sum_{j=i+1}^{+\infty}\mathbb{P}\big(\{X=i\}\cap\{Y=j\}\big) &= \sum_{j=i+1}^{+\infty}p^2(1-p)^{j-2}\\ &= \frac{p^2(1-p)^{i-1}}{1-(1-p)}=p(1-p)^{i-1}.\end{aligned}$$

   Sans surprise, on trouve que X suit une loi géométrique.
   La variable aléatoire Y est à valeurs dans $[\![2,+\infty[\![$. Pour $j\geqslant 2$, on a :

$$\mathbb{P}(Y=j)=\sum_{i=1}^{j-1}\mathbb{P}(X=i,Y=j)=\sum_{i=1}^{j-1}p^2(1-p)^{j-2}=(j-1)p^2(1-p)^{j-2}.$$

Exercice 6

- Pour $j\geqslant 2$ et $i\geqslant 1$, on a :

$$\begin{aligned}\mathbb{P}(X=i\mid Y=j) &= \frac{\mathbb{P}\big(\{X=i\}\cap\{Y=j\}\big)}{\mathbb{P}(Y=j)}\\ &= \begin{cases}0 & \text{si} \quad i\geqslant j\\ \dfrac{p^2(1-p)^{j-2}}{(j-1)p^2(1-p)^{j-2}}=\dfrac{1}{j-1} & \text{si} \quad i<j.\end{cases}\end{aligned}$$

  La loi conditionnelle de X sachant $\{Y=j\}$ est la loi uniforme sur $[\![1,j-1]\!]$. Le rang du deuxième tirage étant connu, égal à j, le premier se répartit uniformément entre 1 et $j-1$. Cela reflète l'indépendance des différentes épreuves.

- Pour $i \geqslant 1$, on a, pour tout $j \geqslant 1$:

$$\begin{aligned}\mathbb{P}(Y-i=j \mid X=i) &= \frac{\mathbb{P}\big(\{Y-i=j\}\cap\{X=i\}\big)}{\mathbb{P}(X=i)} \\ &= \frac{\mathbb{P}\big(\{Y=i+j\}\cap\{X=i\}\big)}{\mathbb{P}(X=i)} \\ &= \frac{p^2(1-p)^{i+j-2}}{p(1-p)^{i-1}} = p(1-p)^{j-1}.\end{aligned}$$

La loi conditionnelle de $Y-i$ conditionnelle sachant $\{X=i\}$ est une loi géométrique de paramètre p. Une fois obtenu le premier succès au rang i, le temps d'attente $Y-i$ du succès suivant suit encore une loi géométrique. Cela reflète la propriété de la loi géométrique d'être sans mémoire.

Exercice 7

- Soit $(m,k) \in \mathbb{N}^2$. Par définition de la loi binomiale, on a :

$$\mathbb{P}\big(Y=k \mid X=m\big) = \begin{cases} \binom{m}{k} p^k (1-p)^{m-k} & \text{si} \quad k \in [\![0,m]\!] \\ 0 & \text{si} \quad k > m, \end{cases}$$

et donc, par définition de la loi conditionnelle :

$$\begin{aligned}\mathbb{P}\left(\{Y=k\}\cap\{X=m\}\right) &= \mathbb{P}\big(X=k \mid Y=m\big)\,\mathbb{P}(Y=m) \\ &= \begin{cases} \binom{m}{k} p^k (1-p)^{m-k} \dfrac{\lambda^m}{m!}\, e^{-\lambda} & \text{si} \quad k \in [\![0,m]\!] \\ 0 & \text{si} \quad k > m. \end{cases}\end{aligned}$$

- On en déduit que, pour tout entier naturel k :

$$\begin{aligned}\mathbb{P}(Y=k) = \sum_{m=0}^{+\infty} \mathbb{P}\big(\{Y=k\}\cap\{X=m\}\big) &= \sum_{m=k}^{+\infty} \binom{m}{k} p^k (1-p)^{m-k} \frac{\lambda^m}{m!}\, e^{-\lambda} \\ &= (p\lambda)^k \frac{1}{k!}\, e^{-\lambda} \sum_{m=k}^{+\infty} \frac{((1-p)\lambda)^{m-k}}{(m-k)!} \\ &= (p\lambda)^k \frac{1}{k!}\, e^{-\lambda} e^{(1-p)\lambda} = \frac{(p\lambda)^k}{k!}\, e^{-p\lambda}.\end{aligned}$$

Donc Y suit la loi de Poisson de paramètre $p\lambda$.

Théorème 17 Le $(n-1)$-uplet $Z = (X_1,\ldots,X_{k-1},X_{k+1},\ldots,X_n)$, où figurent toutes les variables aléatoires sauf X_k, est une variable aléatoire dont le système complet d'événements associé est :

$$(Z=z)_{z\in X_1(\Omega)\times\ldots\times X_{k-1}(\Omega)\times X_{k+1}(\Omega)\times\ldots\times X_n(\Omega)}.$$

On a donc, d'après la formule des probabilités totales, pour tout $x_k \in X_k(\Omega)$:

$$\begin{aligned}\mathbb{P}(X_k = x_k) &= \sum_{\substack{x_1\in X_1(\Omega),\dots,x_{k-1}\in X_{k-1}(\Omega),\\ x_{k+1}\in X_{k+1}(\Omega),\dots,x_n\in X_n(\Omega)}} \mathbb{P}\left(Z = (x_1,\dots,x_{k-1},x_{k+1},\dots,x_n), X_k = x_k\right)\\ &= \sum_{\substack{x_1\in X_1(\Omega),\dots,x_{k-1}\in X_{k-1}(\Omega),\\ x_{k+1}\in X_{k+1}(\Omega),\dots,x_n\in X_n(\Omega)}} \mathbb{P}(X_1 = x_1,\dots,X_n = x_n).\end{aligned}$$

Exercice 8 Si, pour $1 \leqslant i \leqslant n$, la variable aléatoire X_i est à valeurs dans E_i et si f est une fonction quelconque définie sur $E_1 \times \cdots \times E_n$, alors $f(X_1,\dots,X_n)$ est une variable aléatoire, d'après la proposition 6 de la page 895 appliquée à la fonction f et à la variable aléatoire $(X_1,\dots,X_n)$.

Exercice 9 En sommant sur $x \in X(\Omega)$, on obtient, pour tout $y \in Y(\Omega)$:

$$\mathbb{P}(Y = y) = \sum_{x\in X(\Omega)} \mathbb{P}\big(\{X = x\}\cap\{Y = y\}\big) = \sum_{x\in X(\Omega)} \varphi(x)\,\psi(y) = C'\psi(y),$$

où $C' = \sum\limits_{x\in X(\Omega)} \varphi(x)$.

De même, pour tout $x \in X(\Omega)$, on a $\mathbb{P}(X = x) = C\varphi(x)$, où $C = \sum\limits_{y\in Y(\Omega)} \psi(y)$.

De $\sum\limits_{y\in Y(\Omega)} \mathbb{P}(Y = y) = 1$, on tire $CC' = 1$ et donc, pour tout $(x,y) \in X(\Omega)\times Y(\Omega)$:

$$\mathbb{P}\left(\{X = x\}\cap\{Y = y\}\right) = \big(C\varphi(x)\big)\big(C'\psi(y)\big) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y),$$

donc X et Y sont indépendantes.

Proposition 18 Soit $A \subset E$ et $B \subset F$. Les ensembles $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables, donc il en est de même de leurs sous-ensembles $A' = A\cap X(\Omega)$ et $B' = B\cap Y(\Omega)$, puis de $A' \times B'$. On a :

$$\{X \in A\} = \{X \in A'\} = \bigcup_{x\in A'} \{X = x\} \quad \text{et} \quad \{Y \in B\} = \{Y \in B'\} = \bigcup_{y\in B'} \{Y = y\}.$$

On en déduit $\{X \in A\}\cap\{Y \in B\} = \bigcup\limits_{(x,y)\in A'\times B'} \{X = x\}\cap\{Y = y\}$.

On a une réunion dénombrable d'événements incompatibles. On en déduit :

$$\begin{aligned}\mathbb{P}\big(\{X \in A\}\cap\{Y \in B\}\big) &= \sum_{(x,y)\in A'\times B'} \mathbb{P}\big(\{X = x\}\cap\{Y = y\}\big)\\ &= \sum_{(x,y)\in A'\times B'} \mathbb{P}\big(X = x\big)\,\mathbb{P}\big(Y = y\big)\\ &= \left(\sum_{x\in A'} \mathbb{P}\big(X = x\big)\right)\left(\sum_{y\in B'} \mathbb{P}\big(Y = y\big)\right).\end{aligned}$$

Comme $\mathbb{P}(X \in A) = \sum\limits_{x\in A'} \mathbb{P}\big(X = x\big)$ et $\mathbb{P}(Y \in B) = \sum\limits_{y\in B'} \mathbb{P}\big(Y = y\big)$, on a donc :

$$\mathbb{P}\big(\{X \in A\}\cap\{Y \in B\}\big) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B).$$

Proposition 19 Montrons l'équivalence entre (i) et (ii).

- Supposons que (i) soit vérifié.
  Soit $y \in Y(\Omega)$ tel que $\mathbb{P}\big(Y = y\big) \neq 0$. On a, pour tout $x \in X(\Omega)$:
  $$\mathbb{P}_{\{Y=y\}}(X = x) = \frac{\mathbb{P}(\{X = x\} \cap \{Y = y\})}{\mathbb{P}(Y = y)} = \frac{\mathbb{P}(X = x)\mathbb{P}(Y = y)}{\mathbb{P}(Y = y)} = \mathbb{P}(X = x).$$
  La loi de X sachant $\{Y = y\}$ est égale à la loi de X : donc (ii) est vérifié.
- Supposons que (ii) soit vérifié. Soit $y \in Y(\Omega)$.
  * Si $\mathbb{P}(Y = y) = 0$, alors $\{Y = y\}$ est indépendant de $\{X = x\}$ pour tout $x \in X(\Omega)$, car un événement quasi-impossible est indépendant de tout événement.
  * Sinon on a, pour tout $x \in X(\Omega)$:
    $$\mathbb{P}\left(\{X = x\} \cap \{Y = y\}\right) = \mathbb{P}_{\{Y=y\}}\left(X = x\right) \mathbb{P}\left(Y = y\right) = \mathbb{P}\left(X = x\right) \mathbb{P}\left(Y = y\right).$$

  Les variables aléatoires X et Y sont donc indépendantes.

L'équivalence de (i) et (iii) s'en déduit par symétrie.

Proposition 20 Pour tout $(x, y) \in f(X)(\Omega) \times g(Y)(\Omega)$, on a, d'après la proposition 18 de la page 904 :
$$\begin{aligned}\mathbb{P}\big(\{f(X) = x\} \cap \{g(Y) = y\}\big) &= \mathbb{P}\left(\big\{X \in f^{-1}\big(\{x\}\big)\big\} \cap \big\{Y \in g^{-1}\big(\{y\}\big)\big\}\right)\\ &= \mathbb{P}\big(X \in f^{-1}(\{x\})\big)\, \mathbb{P}\big(Y \in g^{-1}(\{y\})\big)\\ &= \mathbb{P}\big(f(X) = x\big)\, \mathbb{P}\big(g(Y) = y\big).\end{aligned}$$
Ainsi $f(X)$ et $g(Y)$ sont indépendantes.

Proposition 21 Dans le premier cas, on écrit $\{X + Y = k\} = \bigcup\limits_{i=0}^{k} \{X = i\} \cap \{Y = k - i\}$ et dans le second $\{X + Y = k\} = \bigcup\limits_{i \in \mathbf{Z}} \{X = i\} \cap \{Y = k - i\}$ et l'on utilise l'additivité, puis l'indépendance des variables.

Exercice 10 Posons $S = X+Y$. Alors S est à valeurs dans $\mathbb{N}$ et, pour tout k dans $\mathbb{N}$:
$$\begin{aligned}\mathbb{P}\big(S = k\big) &= \sum_{i+j=k} \mathbb{P}\big(X = i\big)\, \mathbb{P}\big(Y = j\big)\\ &= \sum_{i=0}^{k} \mathbb{P}\big(X = i\big)\mathbb{P}\big(Y = k - i\big)\\ &= \sum_{i=0}^{k} \frac{e^{-\lambda}\lambda^i}{i!} \frac{e^{-\mu}\mu^{k-i}}{(k-i)!}\\ &= e^{-(\lambda+\mu)} \frac{1}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{k-i}\\ &= \frac{e^{-\lambda+\mu}}{k!} (\lambda + \mu)^k,\end{aligned}$$
d'après la formule du binôme. Donc S suit la loi de Poisson de paramètre $\lambda + \mu$.

Exercice 11

1. L'hypothèse implique que, pour tout $n \in \mathbb{N}$, on a $\mathbb{P}(N_{a,c}=n) \neq 0$. Par ailleurs, il est clair que $N_{a,c} = N_{a,b} + N_{b,c}$. On a, pour tout $(k,\ell) \in \mathbb{N}^2$, par indépendance de $N_{a,b}$ et N_{bc} :
$$\begin{aligned}\mathbb{P}(N_{a,b}=k)\,\mathbb{P}(N_{b,c}=\ell) &= \mathbb{P}\big(\{N_{a,b}=k\}\cap\{N_{b,c}=\ell\}\big)\\ &= \mathbb{P}\big(\{N_{a,b}=k\}\cap\{N_{a,c}-N_{a,b}=\ell\}\big)\\ &= \mathbb{P}\big(\{N_{a,b}=k\}\cap\{N_{a,c}=k+\ell\}\big)\\ &= \mathbb{P}(N_{ac}=k+\ell)\,\mathbb{P}(N_{a,b}=k \mid N_a=k+\ell)\\ &= \mathbb{P}(N_{a,c}=k+\ell)\binom{k+\ell}{k}p^k\,(1-p)^\ell.\end{aligned}$$
En divisant par $\mathbb{P}(N_{a,c}=k+\ell)$ qui n'est pas nul, on obtient le résultat voulu.

2. De la première question, on déduit $\mathbb{P}(N_{a,b}=k) \neq 0$ et $\mathbb{P}(N_{b,c}=\ell) \neq 0$, pour tout $(k,\ell) \in \mathbb{N}^2$, car le second membre de l'égalité ne s'annule pas. Pour $k \in \mathbb{N}^*$ et $\ell \in \mathbb{N}$, on peut écrire de même :
$$\frac{\mathbb{P}(N_{a,b}=k-1)\,\mathbb{P}(N_{b,c}=\ell+1)}{\mathbb{P}(N_{a,c}=k+\ell)} = \binom{k+\ell}{k-1}p^{k-1}\,(1-p)^{\ell+1}.$$
En divisant l'égalité obtenue dans la première question par celle-ci, on obtient :
$$\frac{\mathbb{P}(N_{a,b}=k)\,\mathbb{P}(N_{b,c}=\ell)}{\mathbb{P}(N_{a,b}=k-1)\,\mathbb{P}(N_{b,c}=\ell+1)} = \frac{p(\ell+1)}{k(1-p)},$$
c'est-à-dire :
$$\frac{\mathbb{P}(N_{a,b}=k)}{\mathbb{P}(N_{a,b}=k-1)} = \frac{p(\ell+1)\,\mathbb{P}(N_{b,c}=\ell+1)}{(1-p)\mathbb{P}(N_{b,c}=\ell)}\times\frac{1}{k}\cdot$$
Si l'on fixe $\ell \in \mathbb{N}$ et que l'on pose $\lambda = \dfrac{p(\ell+1)\,\mathbb{P}(N_{b,c}=\ell+1)}{(1-p)\mathbb{P}(N_{b,c}=\ell)}$, on obtient $\lambda > 0$ et :
$$\forall k \in \mathbb{N}^* \quad \frac{\mathbb{P}(N_{a,b}=k)}{\mathbb{P}(N_{a,b}=k-1)} = \frac{\lambda}{k}\cdot$$

3. En réitérant l'égalité démontrée dans la question précédente, on obtient, pour tout $k \in \mathbb{N}$:
$$\begin{aligned}\mathbb{P}(N_{a,b}=k) = \frac{\lambda}{k}\,\mathbb{P}(N_{a,b}=k-1) &= \frac{\lambda^2}{k(k-1)}\,\mathbb{P}(N_{a,b}=k-2) = \cdots\\ &= \frac{\lambda^k}{k!}\,\mathbb{P}(N_{a,b}=0).\end{aligned}$$
Sachant que $\sum\limits_{k=0}^{+\infty}\mathbb{P}(N_{a,b}=k) = 1$, on obtient $e^\lambda\mathbb{P}(N_{a,b}=0) = 1$ et donc $\mathbb{P}(N_{a,b}=0) = e^{-\lambda}$. Ainsi $N_{a,b}$ suit la loi $\mathcal{P}(\lambda)$.

Proposition 22 On suppose $X_1, \ldots, X_n$ mutuellement indépendantes.

Pour tout $i \in [\![1,n]\!]$, on considère $A_i \subset E_i$ et $A'_i = A_i \cap X_i(\Omega)$; on écrit :

$$\{X_i \in A_i\} = \{X_i \in A'_i\} = \bigcup_{x_i \in A'_i} \{X_i = x_i\}.$$

On obtient :

$$\bigcap_{1 \leqslant i \leqslant n} \{X_i \in A_i\} = \bigcap_{1 \leqslant i \leqslant n} \{X_i \in A'_i\} = \bigcup_{(x_i)_{1 \leqslant i \leqslant n} \in \prod\limits_{1 \leqslant i \leqslant n} A'_i} \left(\bigcap_{1 \leqslant i \leqslant n} \{X_i = x_i\} \right).$$

Chaque A'_i est au plus dénombrable car inclus dans $X_i(\Omega)$ qui est au plus dénombrable. Il en est de même de leur produit. On a donc une réunion dénombrable d'événements incompatibles. On en déduit :

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{1 \leqslant i \leqslant n} \{X_i \in A_i\} \right) &= \sum_{(x_i)_{1 \leqslant i \leqslant n} \in \prod\limits_{1 \leqslant i \leqslant n} A'_i} \mathbb{P}\left(\bigcap_{1 \leqslant i \leqslant n} \{X_i = x_i\} \right) \\
&= \sum_{(x_i)_{1 \leqslant i \leqslant n} \in \prod\limits_{1 \leqslant i \leqslant n} A'_i} \left(\prod_{1 \leqslant i \leqslant n} \mathbb{P}(X_i = x_i) \right) \\
&= \prod_{1 \leqslant i \leqslant n} \left(\sum_{x_i \in A'_i} \mathbb{P}(X_i = x_i) \right) = \prod_{1 \leqslant i \leqslant n} \mathbb{P}(X_i \in A_i).
\end{aligned}$$

Exercice 12 Les variables Y et Z sont à valeurs dans $\mathbb{N}^*$.

1. Pour tous réels $x_1, \ldots, x_n$, y, on a :

$$\min(x_1, \ldots, x_n) \geqslant y \iff \forall i \in [\![1,n]\!] \quad x_i \geqslant y.$$

Pour tout $k \in \mathbb{N}^*$, on a donc $\{Y \geqslant k\} = \{X_1 \geqslant k\} \cap \ldots \cap \{X_n \geqslant k\}$, d'où l'on déduit, par indépendance des variables X_i :

$$\begin{aligned}
\mathbb{P}(Y \geqslant k) &= \mathbb{P}(\{X_1 \geqslant k\} \cap \ldots \cap \{X_n \geqslant k\}) \\
&= \prod_{i=1}^n \mathbb{P}(X_i \geqslant k) = \prod_{i=1}^n (1-p_i)^{k-1} = \left(\prod_{i=1}^n (1-p_i) \right)^{k-1}.
\end{aligned}$$

On obtient :

$$\begin{aligned}
\mathbb{P}(Y = k) &= \mathbb{P}(Y \geqslant k) - \mathbb{P}(Y \geqslant k+1) \\
&= \left(\prod_{i=1}^n (1-p_i) \right)^{k-1} \left(1 - \prod_{i=1}^n (1-p_i) \right).
\end{aligned}$$

Ainsi, Y suit la loi géométrique de paramètre $1 - \prod\limits_{i=1}^n (1-p_i)$.

2. Pour tous réels x_1,..., x_n, y, on a :

$$\max(x_1,\ldots,x_n) \leqslant y \Longleftrightarrow \forall i \in [\![1,n]\!] \quad x_i \leqslant y.$$

Pour tout $k \in \mathbb{N}^*$, on a donc $\{Z \leqslant k\} = \{X_1 \leqslant k\} \cap \ldots \cap \{X_n \leqslant k\}$, d'où l'on déduit, par indépendance des variables X_i :

$$\begin{aligned}\mathbb{P}(Z \leqslant k) &= \mathbb{P}(\{X_1 \leqslant k\} \cap \ldots \cap \{X_n \leqslant k\}) = \prod_{i=1}^{n} \mathbb{P}(X_i \leqslant k)\\ &= \prod_{i=1}^{n} \big(1 - \mathbb{P}(X_i > k)\big) = \prod_{i=1}^{n} \big(1 - (1-p_i)^k\big).\end{aligned}$$

On obtient :

$$\begin{aligned}\mathbb{P}(Z = k) &= \mathbb{P}(Z \leqslant k) - \mathbb{P}(Z \leqslant k-1)\\ &= \prod_{i=1}^{n} \big(1 - (1-p_i)^k\big) - \prod_{i=1}^{n} \big(1 - (1-p_i)^{k-1}\big).\end{aligned}$$

Proposition 23 Soit $I \subset [\![1,n]\!]$ et, pour tout $i \in I$ un élément x_i de $X_i(\Omega)$. On pose $A_i = \{x_i\}$ si $i \in I$ et $A_i = E_i$ si $i \notin I$. On a alors $\{X_i \in A_i\} = \Omega$ pour $i \notin I$. On obtient, en utilisant la proposition 22 :

$$\mathbb{P}\left(\bigcap_{i\in I}\{X_i = x_i\}\right) = \mathbb{P}\left(\bigcap_{i\in[\![1,n]\!]}\{X_i \in A_i\}\right) = \prod_{i\in[\![1,n]\!]} \mathbb{P}(X_i \in A_i) = \prod_{i\in I} \mathbb{P}(X_i = x_i).$$

Exercice 13

On remarque que, pour tout événement A, on a $\{\mathbb{1}_A = 1\} = A$ et $\{\mathbb{1}_A = 0\} = \overline{A}$.

- Supposons les événements A_1,..., A_n mutuellement indépendants.
  On a, pour tout $(x_1,\ldots,x_n) \in \{0,1\}^n$:

  $$\mathbb{P}\left(\{\mathbb{1}_{A_1} = x_1\} \cap \ldots \cap \{\mathbb{1}_{A_n} = x_n\}\right) = \mathbb{P}\left(B_1 \cap \ldots \cap B_n\right),$$

  où $B_i = A_i$ si $x_i = 1$ et $B_i = \overline{A_i}$ si $x_i = 0$. Les événements A_1, A_2, ..., A_n étant indépendants, il en est de même des événements B_1, B_2, ..., B_n. On a donc :

  $$\mathbb{P}\left(\{\mathbb{1}_{A_1} = x_1\} \cap \ldots \cap \{\mathbb{1}_{A_n} = x_n\}\right) = \prod_{i=1}^{n} \mathbb{P}(B_i) = \prod_{i=1}^{n} \mathbb{P}\left(\mathbb{1}_{A_i} = x_i\right).$$

  Les variables $\mathbb{1}_{A_1}$,..., $\mathbb{1}_{A_n}$ sont donc mutuellement indépendantes.
- Supposons les variables aléatoires mutuellement indépendantes. Soit I une partie non vide de $[\![1,n]\!]$. D'après la proposition 23, $(\mathbb{1}_{A_i})_{i\in I}$ est encore une famille de variables aléatoires mutuellement indépendants, donc on a :

  $$\mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \mathbb{P}\left(\bigcap_{i\in I}\{\mathbb{1}_{A_i} = 1\}\right) = \prod_{i\in I} \mathbb{P}(\mathbb{1}_{A_i} = 1) = \prod_{i\in I} \mathbb{P}(A_i).$$

  Les événements A_1,..., A_n sont donc mutuellement indépendants.

Proposition 25 Soit $(y_1,\dots,y_n) \in f_1(X_1)(\Omega)\times\cdots\times f_n(X_n)(\Omega)$. On a :

$$\begin{aligned}\mathbb{P}\big(\{f_1(X_1)=y_1\}\cap\ldots\cap\{f_n(X_n)=y_n\}\big)\\ &= \mathbb{P}\Big(\big\{X_1\in f_1^{-1}(\{y_1\})\big\}\cap\cdots\cap\big\{X_n\in f_n^{-1}(\{y_n\})\big\}\Big)\\ &= \mathbb{P}\big(X_1\in f_1^{-1}(\{y_1\})\big)\ldots\mathbb{P}\big(X_n\in f_n^{-1}(\{y_n\})\big)\\ &= \mathbb{P}\big(f_1(X_1)=y_1\big)\ldots\mathbb{P}\big(f_n(X_n)=y_n\big).\end{aligned}$$

Proposition 26

- Les variables aléatoires $(X_1,\dots,X_p)$ et $(X_{p+1},\dots,X_n)$ sont indépendantes. En effet, on a, pour tout $(x_1,\dots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)$:

$$\begin{aligned}\mathbb{P}\Big(\big\{(X_1,\dots,X_p)=(x_1,\dots,x_p)\big\}\cap\big\{(X_{p+1}\dots,X_n)=(x_{p+1},\dots,x_n)\big\}\Big)\\ &= \mathbb{P}\big(X_1=x_1,\dots,X_n=x_n\big)\\ &= \prod_{k=1}^{n}\mathbb{P}\big(X_k=x_k\big)\\ &= \prod_{k=1}^{p}\mathbb{P}(X_k=x_k)\prod_{k=p+1}^{n}\mathbb{P}(X_k=x_k)\\ &= \mathbb{P}\big((X_1,\dots,X_p)=(x_1,\dots,x_p)\big)\mathbb{P}\big((X_{p+1},\dots,X_n)=(x_{p+1},\dots,x_n)\big),\end{aligned}$$

  car $X_1,\dots,X_p$ d'une part, $X_{p+1},\dots,X_n$ d'autre part, sont indépendantes.

- On en déduit que $f(X_1,\dots,X_p)$ et $g(X_{p+1},\dots,X_n)$ sont indépendantes, d'après la proposition 20 de la page 904.

Exercice 14 On procède par récurrence sur n, la propriété étant évidente pour $n=1$ et ayant été démontré pour $n=2$ dans l'exercice 10 de la page 905.
Supposons que la propriété soit vraie au rang n et considérons $n+1$ variables indépendantes $X_1,\dots,X_{n+1}$ suivant des lois de Poisson. Les variables aléatoires $X_1,\dots,X_n$ sont indépendantes et donc, par hypothèse de récurrence, la variable $X_1+\cdots+X_n$ suit la loi de Poisson de paramètre $\sum\limits_{i=1}^{n}\lambda_i$. Les variables $X_1,\dots,X_n$, X_{n+1} étant mutuellement indépendantes, les variables $X_1+\cdots+X_n$ et X_{n+1} sont indépendantes. Elles suivent des lois de Poisson de paramètres respectifs $\sum\limits_{i=1}^{n}\lambda_i$ et λ_{n+1}. D'après le cas de deux variables, leur somme $X_1+\cdots+X_n+X_{n+1}$ suit la loi de Poisson de paramètre $\sum\limits_{i=1}^{n}\lambda_i+\lambda_{n+1}=\sum\limits_{i=1}^{n+1}\lambda_i$. Cela termine la démonstration par récurrence.

Exercice 15 On pose $q=1-p$. Les variables aléatoires Z_n sont à valeurs dans $\mathbb{N}^*$. Pour $n\in\mathbb{N}^*$ et $(k_1,\dots,k_n)\in(\mathbb{N}^*)^n$, on a :

$$\mathbb{P}(Z_1=k_1,Z_2=k_2,\dots,Z_n=k_n)=\mathbb{P}(Y_1=k_1,Y_2=k_1+k_2,\dots,Y_n=k_1+\ldots+k_n).$$

On posons $\ell_1 = k_1$, $\ell_2 = k_1 + k_2, \ldots,$ $\ell_n = k_1 + \cdots + k_n$, on obtient :

$$\{Z_1 = k_1, Z_2 = k_2, \ldots, Z_n = k_n\} = \bigcap_{1 \leqslant i \leqslant n} \{X_{\ell_i} = 1\} \cap \bigcap_{j \in [\![1, k_1 + \cdots + k_n]\!] \setminus \{\ell_1, \ldots, \ell_n\}} \{X_j = 0\}$$

et donc :

$$\mathbb{P}(Z_1 = k_1, Z_2 = k_2, \ldots, Z_n = k_n) = p^n q^{k_1 + k_2 + \ldots + k_n - n}.$$

car parmi les variables X_i pour $1 \leqslant i \leqslant k_1 + k_2 + \ldots + k_n$, il y en a n qui sont égales à 1 et donc $k_1 + k_2 + \ldots + k_n - n$ qui sont égales à 0. On peut réécrire

$$\mathbb{P}(Z_1 = k_1, Z_2 = k_2, \ldots, Z_n = k_n) = \prod_{i=1}^{n} p q^{k_i - 1}.$$

En sommant par rapport à k_1, on obtient :

$$\begin{aligned}
\mathbb{P}(Z_2 = k_2, \ldots, Z_n = k_n) &= \sum_{k_1=1}^{+\infty} \mathbb{P}(Z_1 = k_1, Z_2 = k_2, \ldots, Z_n = k_n) \\
&= \sum_{k_1=1}^{+\infty} \prod_{i=1}^{n} p q^{k_i - 1} \\
&= \prod_{i=2}^{n} p q^{k_i - 1} \underbrace{\sum_{k_1=1}^{+\infty} p q^{k_1 - 1}}_{=1} = \prod_{i=2}^{n} p q^{k_i - 1}.
\end{aligned}$$

En réitérant le procédé, en sommant par rapport à $k_2, \ldots,$ k_{n-1} successivement, on obtient :

$$\forall k_n \in \mathbb{N}^* \quad \mathbb{P}(Z_n = k_n) = p q^{k_n - 1}.$$

Pour tout $n \in \mathbb{N}^*$, la variable Z_n suit la loi géométrique de paramètre p. On en déduit que, pour $n \in \mathbb{N}^*$ et $(k_1, \ldots, k_n) \in (\mathbb{N}^*)^n$, on a :

$$\mathbb{P}(Z_1 = k_1, Z_2 = k_2, \ldots, Z_n = k_n) = \prod_{i=1}^{n} p q^{k_i - 1} = \prod_{i=1}^{n} \mathbb{P}(Z_i = k_i).$$

Ainsi, pour tout $n \in \mathbb{N}^*$, les variables aléatoires $Z_1, \ldots, Z_n$ sont indépendantes. Donc $(Z_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires indépendantes.

Exercice 16

1. La série de terme général $\dfrac{1}{n(n+1)}$ est convergence, de somme 1. Donc il existe une variable aléatoire X à valeurs dans $\mathbb{N}^*$ dont la loi est donnée par :

$$\forall n \in \mathbb{N}^* \quad \mathbb{P}(X = n) = \frac{1}{n(n+1)}.$$

La série de terme général $\mathbb{P}(X = n)\, n$ diverge donc l'espérance de X est infinie.

2. On a $\displaystyle\sum_{n \in \mathbb{Z}^*} \frac{1}{2|n|(|n|+1)} = 1$. Donc il existe une variable aléatoire Y à valeurs dans $\mathbb{Z}^*$ dont la loi est donnée par :

$$\forall n \in \mathbb{Z}^* \quad \mathbb{P}(Y = n) = \frac{1}{2|n|(|n|+1)}.$$

La série de terme général $\mathbb{P}(X = n)\, n$ diverge donc *a fortiori* la famille $(\mathbb{P}(X=n)\, n)_{n \in \mathbf{Z}^*}$ n'est pas sommable. Ainsi, Y n'a pas d'espérance.

Exercice 17 Soit X une variable aléatoire discrète bornée sur $(\Omega, \mathcal{A}, \mathbb{P})$. Il existe $M \in \mathbb{R}_+$ tel que, pour tout $x \in X(\Omega)$, $|x| \leqslant M$. On a donc :

$$\forall x \in X(\Omega) \quad |\mathbb{P}(X=x)\, x| \leqslant \mathbb{P}(X=x)\, M.$$

La famille $(\mathbb{P}(X=x)\, M)_{x \in X(\Omega)}$ est sommable de somme M, donc la famille $(\mathbb{P}(X=x)\, x)_{x \in X(\Omega)}$ est également sommable et X est d'espérance finie.

Proposition 29

1. Pour $n \in \mathbb{N}^*$, on a $\mathbb{P}(X=n)\, n = pn(1-p)^{n-1}$. La série entière $\sum nx^{n-1}$ a pour rayon de convergence 1 et $\displaystyle\sum_{n=1}^{+\infty} nx^{n-1} = \frac{1}{(1-x)^2}$ pour tout $x \in]-1,1[$. On en déduit que l'espérance de X est finie et :

$$\mathbb{E}(X) = p \sum_{n=1}^{+\infty} n(1-p)^{n-1} = p\frac{1}{p^2} = \frac{1}{p}\cdot$$

2. Pour $n \in \mathbb{N}^*$, on a $\mathbb{P}(X=n)\, n = e^{-\lambda}\dfrac{n\lambda^n}{n!} = \lambda e^{-\lambda}\dfrac{\lambda^{n-1}}{(n-1)!}\cdot$ De $\displaystyle\sum_{n=1}^{+\infty}\frac{\lambda^{n-1}}{(n-1)!} = e^{\lambda}$, on déduit que l'espérance de X est finie et $\mathbb{E}(X) = \lambda e^{-\lambda} e^{\lambda} = \lambda$.

Exercice 18 Pour tout $N \in \mathbb{N}$, on a :

$$\begin{aligned}
\sum_{n=0}^{N} n\mathbb{P}(X=n) &= \sum_{n=0}^{N} n\,(\mathbb{P}(X>n-1) - \mathbb{P}(X>n)) \\
&= \sum_{n=-1}^{N-1} (n+1)\mathbb{P}(X>n) - \sum_{n=0}^{N} n\mathbb{P}(X>n) \\
&= \sum_{n=0}^{N} \mathbb{P}(X>n) - (N+1)\mathbb{P}(X>N). \quad (*)
\end{aligned}$$

- Supposons que X ne soit pas d'espérance finie. Alors la série de terme général $n\,\mathbb{P}(X=n)$ diverge. Il en est de même *a fortiori* de la série de terme général $\mathbb{P}(X>n)$, car pour tout $N \in \mathbb{N}$, on a, d'après $(*)$:

$$\sum_{n=0}^{N} \mathbb{P}(X>n) \geqslant \sum_{n=0}^{N} n\mathbb{P}(X=n).$$

  On a donc :

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X>n) = +\infty.$$

- Supposons que X soit d'espérance finie. La série de terme général $n\mathbb{P}(X = n)$ converge. On a :

$$(N+1)\mathbb{P}(X > N) = (N+1)\sum_{k=N+1}^{+\infty} \mathbb{P}(X = k) \leqslant \sum_{k=N+1}^{+\infty} k\,\mathbb{P}(X = k) = R_N,$$

où R_N est le reste d'indice N de la série convergente $\sum n\mathbb{P}(X = n)$. Donc la suite $((N+1)\mathbb{P}(X > N))$ converge vers 0. On en déduit, par passage à la limite dans la relation $(*)$:

$$\sum_{n=0}^{+\infty} \mathbb{P}(X > n) = \sum_{n=0}^{+\infty} n\mathbb{P}(X = n) = \mathbb{E}(X).$$

La deuxième formule s'obtient par un changement d'indice.
Si X suit la loi géométrique, on a $\mathbb{P}(X > n) = (1-p)^n$ pour $n \in \mathbb{N}$, d'après la proposition 7 de la page 896. On en déduit :

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n) = \sum_{n=0}^{+\infty} (1-p)^n = \frac{1}{1-(1-p)} = \frac{1}{p}\cdot$$

Théorème 31

1. La famille $\big(\mathbb{P}(X = x)\,x\big)_{x\in X(\Omega)}$ étant sommable, il en est de même de la famille $\big(\mathbb{P}(X = x)\,\lambda x\big)_{x\in X(\Omega)}$, donc d'après la formule de transfert appliquée à X et à la fonction $x \mapsto \lambda x$, la variable aléatoire λX est d'espérance finie et :

$$\mathbb{E}(\lambda X) = \sum_{x\in X(\Omega)} \mathbb{P}(X = x)\,\lambda x = \lambda \sum_{x\in X(\Omega)} \mathbb{P}(X = x)\,x = \lambda\mathbb{E}(X).$$

2. On écrit $X + Y = f(X, Y)$ où f est l'application $(x, y) \mapsto x + y$.
Comme $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$, $X + Y$ est d'espérance finie si, et seulement si, la famille $\Big(\mathbb{P}\big((X, Y) = (x, y)\big)\,(x + y)\Big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable. On a :

$$\mathbb{P}\big((X, Y) = (x, y)\big)\,|x + y| \leqslant \mathbb{P}\big((X, Y) = (x, y)\big)|x| + \mathbb{P}\big((X, Y) = (x, y)\big)|y|,$$

donc il suffit de montrer que les deux familles

$$\big(\mathbb{P}\left((X, Y) = (x, y)\right)\,|x|\big)_{(x,y)\in X(\Omega)\times Y(\Omega)} \quad \text{et}$$
$$\big(\mathbb{P}\left((X, Y) = (x, y)\right)\,|y|\big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$$

sont sommables.
Montrons-le pour la première famille. Pour tout $y \in Y(\Omega)$, on a :

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

La famille $(|x|\mathbb{P}(X = x))_{x\in X(\Omega)}$ est sommable car X est d'espérance finie. Le théorème de sommation par paquets permet de conclure que la famille $(\mathbb{P}\left((X, Y) = (x, y)\right)\,|x|)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable.

On montre de même que la famille $\Big(\mathbb{P}\big((X,Y)=(x,y)\big)\,|y|\Big)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable. Ainsi $X+Y$ possède une espérance donnée par :

$$\begin{aligned}\mathbb{E}(X+Y) &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbb{P}\big((X,Y)=(x,y)\big)\,(x+y)\\ &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbb{P}\big((X,Y)=(x,y)\big)\,x + \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbb{P}\big(X,Y)=(x,y)\big)\,y.\end{aligned}$$

Mais le théorème de sommation par paquets donne maintenant :

$$\begin{aligned}\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbb{P}\big((X,Y)=(x,y)\big)\,x &= \sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)} \mathbb{P}\big((X,Y)=(x,y)\big)\,x\\ &= \sum_{x\in X(\Omega)} \mathbb{P}(X=x)\,x = \mathbb{E}(X).\end{aligned}$$

De même, on a $\sum_{(x,y)\in X(\Omega)\times Y(\Omega)} \mathbb{P}\big((X,Y)=(x,y)\big)\,y = \mathbb{E}(Y)$, ce qui permet de conclure :

$$\mathbb{E}(X+Y)=\mathbb{E}(X)+\mathbb{E}(Y).$$

Exercice 19

1. Soit $k\in[\![2,n]\!]$ et $(i,j)\in\mathbb{N}^*$. Si $\mathbb{P}(T_{k-1}=i)\neq 0$, on a :

$$\begin{aligned}\mathbb{P}(Z_k=j\mid T_{k-1}=i) &= \mathbb{P}(T_k-T_{k-1}=j\mid T_{k-1}=i)\\ &= \mathbb{P}(T_k=i+j\mid T_{k-1}=i).\end{aligned}$$

Si $\{T_{k-1}=i\}$ est réalisé, les variables aléatoires $X_1,\dots,X_{i-1}$ ont donné $k-2$ valeurs différentes et X_i une valeur différentes des $k-2$ précédentes. Alors $\{T_k=i+j\}$ est réalisé si les variables aléatoires $X_{i+1},\dots,X_{i+j-1}$ redonnent une des $k-1$ valeurs déjà obtenues, ce qui pour chaque variable aléatoire se fait avec la probabilité $\dfrac{k-1}{n}$ et si X_{i+j} donne une valeur différente des précédentes, ce qui se fait avec une probabilité $1-\dfrac{k-1}{n}\cdot$

Par indépendance des variables aléatoires $X_{i+1},\dots,X_{i+j}$, on obtient :

$$\mathbb{P}(Z_k=j\mid T_{k-1}=i)=\left(\frac{k-1}{n}\right)^{j-1}\left(1-\frac{k-1}{n}\right).$$

Ce résultat ne dépend pas de i, donc pour tout $j\in\mathbb{N}^*$, on a :

$$\begin{aligned}\mathbb{P}(Z_k=j) &= \sum_{i=1}^{+\infty}\mathbb{P}(Z_k=j\mid T_{k-1}=i)\mathbb{P}(T_{k-1}=i)\\ &= \left(\frac{k-1}{n}\right)^{j-1}\left(1-\frac{k-1}{n}\right)\sum_{i=1}^{+\infty}\mathbb{P}(T_{k-1}=i)\\ &= \left(\frac{k-1}{n}\right)^{j-1}\left(1-\frac{k-1}{n}\right).\end{aligned}$$

La variable Z_k suit la loi $\mathcal{G}\left(\dfrac{n-k+1}{n}\right)$.

2. On a $T_n = \sum\limits_{k=1}^{n} Z_k$. La variable aléatoire $Z_1 = T_1$ est la variable certaine égale à 1. Pour $2 \leqslant k \leqslant i$, on a $\mathbb{E}(Z_k) = \dfrac{n}{n-k+1}\cdot$ Comme la formule est valable pour $k=1$, on obtient :

$$\mathbb{E}(T_n) = \sum_{k=1}^{n} \mathbb{E}(Z_k) = \sum_{k=1}^{n} \frac{n}{n-k+1} = n\sum_{j=1}^{n} \frac{1}{j}\cdot$$

On en déduit $\mathbb{E}(T_n) \underset{n\to+\infty}{\sim} n \ln n$.

Proposition 33

- La famille $\big(\mathbb{P}(X=x)\,x\big)_{x\in X(\Omega)}$ est à termes positifs, donc sa somme $\mathbb{E}(X)$ est positive.
- Si $X=0$ presque sûrement alors, pour tout $x \in X(\Omega)\setminus\{0\}$, on a $\mathbb{P}(X=x)=0$. On en déduit $\mathbb{P}(X=x)\,x=0$ pour tout $x\in X(\Omega)$ et donc $\mathbb{E}(X)=0$.
- Réciproquement, si $\mathbb{E}(X)=0$, on a alors, pour tout $x \in X(\Omega)$, $0 \leqslant x\mathbb{P}(X=x) \leqslant \mathbb{E}(X)=0$ et donc $\mathbb{P}(X=x)=0$ si $x\neq 0$. On en déduit :

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

Corollaire 35 D'après le théorème de transfert, $|X|$ est d'espérance finie si, et seulement si, la famille $(\mathbb{P}(X=x)\,|x|)_{x\in X(\Omega)}$ est sommable, ce qui équivaut par définition au fait que la famille $(\mathbb{P}(X=x)\,x)_{x\in X(\Omega)}$ soit sommable, c'est-à-dire que X soit d'espérance finie.
De $-|X| \leqslant X \leqslant |X|$, on tire, par linéarité et croissance de l'espérance, dans le cas où X est d'espérance finie :

$$-\mathbb{E}\big(|X|\big) \leqslant \mathbb{E}\big(X\big) \leqslant \mathbb{E}\big(|X|\big), \quad \text{c'est-à-dire } \big|\mathbb{E}(X)\big| \leqslant \mathbb{E}\big(|X|\big).$$

Proposition 36 Considérons le couple de variables aléatoires $Z=(X,Y)$ et les deux projections canoniques de IR^2, $\pi_1 : (x,y)\mapsto x$ et $\pi_2 : (x,y)\mapsto y$. On a donc $X=\pi_1(Z)$ et $Y=\pi_2(Z)$.
Comme Y est d'espérance finie, d'après le théorème de transfert, appliqué à la variable aléatoire Z et à la fonction π_2, la famille $\big(\pi_2(x,y)\mathbb{P}(Z=(x,y))\big)_{(x,y)\in Z(\Omega)}$ est sommable et :

$$\mathbb{E}(Y) = \sum_{(x,y)\in Z(\Omega)} \pi_2(x,y)\,\mathbb{P}\big(Z=(x,y)\big) = \sum_{(x,y)\in Z(\Omega)} y\,\mathbb{P}\big(Z=(x,y)\big).$$

Si $(x,y)\in Z(\Omega)$, alors il existe $\omega\in\Omega$ tel que $(x,y)=(X(\omega),Y(\omega))$.
Comme $|X|\leqslant Y$, on alors $|x|\leqslant y$. On a donc, pour tout $(x,y)\in Z(\Omega)$:

$$|x|\mathbb{P}\big(Z=(x,y)\big) \leqslant y\mathbb{P}\big(Z=(x,y)\big).$$

De la sommabilité de la famille $\big(y\mathbb{P}(Z=(x,y))\big)_{(x,y)\in Z(\Omega)}$, on déduit la sommabilité de la famille $\big(|x|\,\mathbb{P}(Z=(x,y))\big)_{(x,y)\in Z(\Omega)}$, c'est-à-dire celle de $\big(x\,\mathbb{P}(Z=(x,y))\big)_{(x,y)\in Z(\Omega)}$ ou encore de la famille $(\pi_1(x,y)\mathbb{P}(Z=(x,y))_{(x,y)\in Z(\Omega)}$. Comme $\pi_1(Z)=X$, on en déduit, par le théorème de transfert, que X est d'espérance finie.

Proposition 37 On pose $Z=(X,Y)$ et on écrit $XY=u(Z)$, où u est l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $u(x,y)=xy$.

D'après le théorème de transfert, Z est d'espérance finie si, et seulement si, la famille :

$$\Big(u(x,y)\mathbb{P}\big(Z=(x,y)\big)\Big)_{(x,y)\in Z(\Omega)} \quad \text{c'est-à-dire} \quad \Big(xy\mathbb{P}\big((X,Y)=(x,y)\big)\Big)_{(x,y)\in Z(\Omega)}$$

est sommable. Si $(x,y)\in X(\Omega)\times Y(\Omega)\setminus(X,Y)(\Omega)$, alors $\mathbb{P}(((X,Y)=(x,y))=0$. Il revient donc au même de dire que la famille $(xy\mathbb{P}\left((X,Y)=(x,y)\right))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable.

Dans le cas où XY est d'espérance finie, on a, toujours par la formule de transfert :

$$\mathbb{E}(XY)=\sum_{(x,y)\in(X,Y)(\Omega)} xy\mathbb{P}\left((X,Y)=(x,y)\right)=\sum_{x,y)\in X(\Omega)\times Y(\Omega)} xy\mathbb{P}\left((X,Y)=(x,y)\right).$$

Théorème 38

Supposons X et Y indépendantes. On a $\mathbb{P}\big((X,Y)=(x,y)\big)=\mathbb{P}(X=x)\mathbb{P}(Y=y)$, pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$. D'après la proposition 37, XY est d'espérance finie si, et seulement si, la famille $(xy\mathbb{P}(X=x)\,\mathbb{P}(Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable.

Par hypothèse, les familles $\big(x\,\mathbb{P}(X=x)\big)$ et $\big(y\,\mathbb{P}(Y=y)\big)$ sont sommables. D'après l'exercice III.3 de la page 462, on en déduit que la famille $(xy\mathbb{P}(X=x)\,\mathbb{P}(Y=y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable et que sa somme est :

$$\left(\sum_{x\in X(\Omega)} x\,\mathbb{P}(X=x)\right)\left(\sum_{y\in Y(\Omega)} y\,\mathbb{P}(Y=y)\right),$$

ce qui est le résultat voulu.

Exercice 20 Les variables aléatoires $f(X)$, $g(Y)$ et $f(X)g(Y)$ sont bornées, donc possèdent une espérance.

- Si X et Y sont indépendantes, il en est de même de $f(X)$ et $g(Y)$ et l'égalité découle directement du théorème 38.
- Réciproquement supposons que $\mathbb{E}\big(f(X)g(Y)\big)=\mathbb{E}\big(f(X)\big)\mathbb{E}\big(g(Y)\big)$, pour toutes fonctions f et g bornées.
  Soit $(x,y)\in X(\Omega)\times Y(\Omega)$. Prenons $f=\mathbb{1}_{\{x\}}$ et $g=\mathbb{1}_{\{y\}}$. On a alors
  $$f(X)=\mathbb{1}_{\{X=x\}},\ \ g(Y)=\mathbb{1}_{\{Y=y\}} \qquad \text{et} \qquad f(X)g(Y)=\mathbb{1}_{\{X=x\}\cap\{Y=y\}}.$$
  L'hypothèse se traduit par $\mathbb{P}(\{X=x\}\cap\{Y=y\})=\mathbb{P}(X=x)\mathbb{P}(Y=y)$. Cela est vrai pour tout $(x,y)\in X(\Omega)\times Y(\Omega)$, ce qui établit l'indépendance de X et Y.

Théorème 39 Soit $\omega \in \Omega$.

- Si $\omega \in \{X \geqslant a\}$, alors on a $\mathbb{1}_{\{X \geqslant a\}}(\omega) = 1 \leqslant \dfrac{X(\omega)}{a}$, car $X(\omega) \geqslant a$.
- Dans le cas contraire, on a $\mathbb{1}_{\{X \geqslant a\}}(\omega) = 0 \leqslant \dfrac{X(\omega)}{a}$, car X est une variable positive.

Ainsi, $\mathbb{1}_{\{X \geqslant a\}} \leqslant \dfrac{X}{a}$, et comme les deux variables aléatoires sont d'espérance finie, on obtient par croissance de l'espérance :

$$\mathbb{P}(X \geqslant a) = \mathbb{E}(\mathbb{1}_{\{X \geqslant a\}}) \leqslant \mathbb{E}\left(\frac{X}{a}\right) = \frac{\mathbb{E}(X)}{a}\cdot$$

Exercice 21 Comme X est positive, on a, pour tout $a > 0$:

$$\{X > a\} = \{X^r > a^r\}.$$

En appliquant l'inégalité de Markov à la variable aléatoire X^r, on obtient :

$$\mathbb{P}(X \geqslant a) = \mathbb{P}(X^r \geqslant a^r) \leqslant \frac{\mathbb{E}(X^r)}{a^r}\cdot$$

Proposition 40 On a pour tout $x \in \mathrm{IR}$, $|x| \leqslant x^2 + 1$, car $|x| \leqslant x^2$ si $|x| \geqslant 1$. On a donc

$$|X| \leqslant X^2 + 1.$$

La variable X^2+1 est d'espérance finie, car somme de variables d'espérances finies, donc X est d'espérance finie d'après la proposition 36 de la page 914.

Exercice 22 Même preuve que celle de la proposition 40, en utilisant l'inégalité :

$$\forall x \in \mathrm{IR} \quad |x^k| \leqslant |x^r| + 1.$$

.

Lemme 41 On a $|XY| \leqslant \dfrac{1}{2}(X^2+Y^2)$. Les variables aléatoires X^2 et Y^2 étant d'espérance finie, il en est de même de $\dfrac{1}{2}(X^2+Y^2)$. On en déduit que XY est d'espérance finie, d'après la proposition 36 de la page 914.

Proposition 42 La variable aléatoire nulle appartient à $L_2(\Omega, \mathcal{A}, \mathbb{P})$.

Si X et Y sont deux variables aléatoires réelles discrètes admettant un moment d'ordre 2 et λ est un réel, on a $(\lambda X+Y)^2 = \lambda^2 X^2 + 2\lambda XY + Y^2$. Par hypothèse les variables X^2 et Y^2 sont d'espérance finie. Il en est de même de XY, d'après le lemme 41. La variable aléatoire $(\lambda X + Y)^2$ est donc d'espérance finie, comme combinaison linéaire de variables aléatoires d'espérance finie, donc $\lambda X + Y \in L_2(\Omega, \mathcal{A}, \mathbb{P})$.

Théorème 43 Il a été déjà démontré que XY est d'espérance finie, dans le lemme 41. Pour tout $\lambda \in \mathrm{IR}$, la variable $(\lambda X+Y)^2 = \lambda^2 X^2 + 2\lambda XY + Y^2$ est d'espérance finie, comme combinaison linéaire de variables aléatoires ayant une espérance finie. Son espérance est positive car c'est une variable aléatoire positive. Par linéarité de l'espérance, on obtient :

$$\forall \lambda \in \mathrm{IR} \quad \lambda^2\mathbb{E}(X^2) + 2\lambda\mathbb{E}(XY) + \mathbb{E}(Y^2) = \mathbb{E}\left((\lambda X + Y)^2\right) \geqslant 0.$$

- Si $\mathbb{E}(X^2) = 0$, alors X^2 est presque sûrement nulle, donc X est presque sûrement nulle. On en déduit que XY est presque sûrement nulle. On a donc $\mathbb{E}(XY) = 0$ et $\mathbb{E}(XY)^2 = \mathbb{E}(X^2)\mathbb{E}(Y^2)$.
- Si $\mathbb{E}(X) \neq 0$, la fonction trinôme $\lambda \longmapsto \mathbb{E}\left((\lambda X + Y)^2\right)$ qui garde un signe constant. Son discriminant est donc négatif, ce qui donne l'inégalité voulue.

Exercice 23 On garde les notations du théorème 43.

- Si X est presque sûrement nulle, on a $\mathbb{E}(XY)^2 = \mathbb{E}(X^2)\mathbb{E}(Y^2) = 0$.
- Sinon, l'égalité $\mathbb{E}(XY)^2 = \mathbb{E}(X^2)\mathbb{E}(Y^2)$ équivaut à la nullité du discriminant de la fonction trinôme $\lambda \longmapsto \mathbb{E}\left((\lambda X + Y)^2\right)$, c'est-à-dire à l'existence de $\lambda \in \mathbb{R}$ tel que $\mathbb{E}\left((\lambda X + Y)^2\right) = 0$. Cette dernière condition est réalisée si, et seulement si, $Y = -\lambda X$ presque sûrement.

On conclut qu'il y a égalité dans l'inégalité de Cauchy-Schwarz si, et seulement si :

$$\mathbb{P}(X = 0) = 1 \quad \text{ou} \quad \exists \alpha \in \mathbb{R} \quad \mathbb{P}(Y = \alpha X) = 1.$$

Exercice 24 Si X une variable aléatoire réelle discrète sur $(\Omega, \mathcal{A}, \mathbb{P})$, possédant un moment d'ordre 2, on a :

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

donc, d'après la proposition 33 de la page 913, $\mathbb{V}(X) = 0$ équivaut à :

$$\mathbb{P}\big(X - \mathbb{E}(X) = 0\big) = 1, \ \text{ c'est-à-dire à } \ \mathbb{P}\big(X = \mathbb{E}(X)\big) = 1.$$

Si cette condition est réalisée, alors X est presque sûrement constante.

Réciproquement, s'il existe $a \in \mathbb{R}$ tel que $\mathbb{P}(X = a) = 1$, alors :

$$\mathbb{E}(X) = a \qquad \text{et} \qquad \mathbb{P}\big(X = \mathbb{E}(X)\big) = 1.$$

La variance de X est donc nulle si, et seulement si, X est presque sûrement constante.

Théorème 44

- Si X possède un moment d'ordre 2, alors, d'après la proposition 40 de la page 915, elle est d'espérance finie. De l'égalité :

$$\big(X - \mathbb{E}(X)\big)^2 = X^2 - 2\mathbb{E}(X)\,X + (\mathbb{E}(X))^2,$$

on déduit par linéarité que $\big(X - \mathbb{E}(X)\big)^2$ est d'espérance finie. Ainsi $\mathbb{V}(X)$ existe et :

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}\Big(\big(\mathbb{E}(X)\big)^2\Big) \\
&= \mathbb{E}(X^2) - 2\big(\mathbb{E}(X)\big)^2 + \big(\mathbb{E}(X)\big)^2 \\
&= \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2,
\end{aligned}$$

puisque $\mathbb{E}(X)$ est une variable aléatoire constante dont l'espérance est $\mathbb{E}(X)$.

- Réciproquement si $\mathbb{V}(X)$ existe, X et $\big(X - \mathbb{E}(X)\big)^2$ ont des espérances. Il en est de même de X^2, car $X^2 = \big(X - \mathbb{E}(X)\big)^2 + 2\mathbb{E}(X)\,X - \big(\mathbb{E}(X)\big)^2$. La variable X possède donc un moment d'ordre 2.

Proposition 45 Dans les deux cas, il est plus simple de commencer par calculer $\mathbb{E}\left(X(X-1)\right)$, plutôt que de calculer directement $\mathbb{E}(X^2)$.

1. Pour $n \in \mathbb{N}^*$, on a $\mathbb{P}(X=n)\,n(n-1) = p(1-p)^{n-1}n(n-1)$.
   De $\displaystyle\sum_{n=2}^{+\infty} n(n-1)x^{n-2} = \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{1}{1-x}\right) = \frac{2}{(1-x)^3}$ pour tout $x \in \left]-1,1\right[$, on déduit que la variable aléatoire $X(X-1)$ possède une espérance finie donnée par la formule de transfert :
   $$\mathbb{E}(X(X-1)) = p(1-p)\sum_{n=2}^{+\infty} n(n-1)(1-p)^{n-2} = p(1-p)\frac{2}{p^3} = \frac{2(1-p)}{p^2}\cdot$$
   Comme $\mathbb{E}(X) = \dfrac{1}{p}$, on en déduit par linéarité :
   $$\mathbb{E}(X^2) = \mathbb{E}\big(X(X-1)\big) + \mathbb{E}(X) = \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2-p}{p^2},$$
   puis :
   $$\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \frac{1-p}{p^2}\cdot$$
2. Pour $n \geqslant 2$, on a
   $$\mathbb{P}(X=n)\,n(n-1) = e^{-\lambda}\frac{n(n-1)\lambda^n}{n!} = \lambda^2 e^{-\lambda}\frac{\lambda^{n-2}}{(n-2)!}\cdot$$
   De $\displaystyle\sum_{n=2}^{+\infty} \frac{\lambda^{n-2}}{(n-2)!} = e^{\lambda}$, on déduit que $X(X-1)$ possède une espérance finie et :
   $$\mathbb{E}(X(X-1)) = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2.$$
   On en déduit, comme dans le premier point :
   $$\mathbb{V}(X) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - (\mathbb{E}(X))^2 = \lambda.$$

Proposition 46 La variable X admet un moment d'ordre 2 donc est d'espérance finie. On en déduit que $aX+b$ est d'espérance finie et $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b$. On obtient :
$$\big(aX+b-\mathbb{E}(aX+b)\big)^2 = a^2\big(X-\mathbb{E}(X)\big)^2.$$
Comme X admet une variance $\big(X-E(X)\big)^2$ est d'espérance finie. Par linéarité, il en est de même de $a^2\big(X-\mathbb{E}(X)\big)^2$. Ainsi $aX+b$ admet une variance et :
$$\begin{aligned}\mathbb{V}(aX+b) = \mathbb{E}\big(aX+b-\mathbb{E}(aX+b)\big)^2 &= \mathbb{E}\big(a^2(X-\mathbb{E}(X))^2\big)\\ &= a^2\mathbb{E}\big((X-\mathbb{E}(X))^2\big) = a^2\mathbb{V}(X).\end{aligned}$$

Théorème 47 La variable $\big(X-\mathbb{E}(X)\big)^2$ est positive et possède une espérance égale à $\mathbb{V}(X)$. D'après l'inégalité de Markov, on a pour tout $\varepsilon > 0$:
$$\mathbb{P}\Big(\big(X-\mathbb{E}(X)\big)^2 \geqslant \varepsilon^2\Big) \leqslant \frac{\mathbb{V}(X)}{\varepsilon^2}\cdot$$

Comme les événements $\{|X-\mathbb{E}(X)| \geqslant \varepsilon\}$ et $\{(X-\mathbb{E}(X))^2 \geqslant \varepsilon^2\}$ sont égaux, on a le résultat voulu.

Proposition 48 Par linéarité de l'espérance, on a $\mathbb{E}(X^*) = \dfrac{\mathbb{E}(X)-\mathbb{E}(X)}{\sigma(X)} = 0$. En utilisant la proposition 46, on obtient :

$$\mathbb{V}(X^*) = \frac{\mathbb{V}(X)}{\sigma(X)^2} = 1.$$

Théorème 49 Par linéarité de l'espérance, la variable aléatoire :

$$\big(X-\mathbb{E}(X)\big)\big(Y-\mathbb{E}(Y)\big) = XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\,\mathbb{E}(Y)$$

admet une espérance. Donc le couple (X,Y) possède une covariance et :

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}\Big(\big(X-\mathbb{E}(X)\big)\big(Y-\mathbb{E}(Y)\big)\Big)\\ &= \mathbb{E}(XY) - \mathbb{E}(Y)\,\mathbb{E}(X) - \mathbb{E}(X)\,\mathbb{E}(Y) + \mathbb{E}(X)\,\mathbb{E}(Y)\\ &= \mathbb{E}(XY) - \mathbb{E}(Y)\,\mathbb{E}(X).\end{aligned}$$

Si X et Y ont un moment d'ordre 2, alors elles ont une espérance et XY a une espérance d'après le théorème 43 de la page 915.

Exercice 25

- Pour $(m,n) \in (\mathbb{N}^*)^2$, on a :

$$\begin{aligned}&\mathbb{P}(L_1 = m, L_2 = n)\\ &= \mathbb{P}\left(\{X_1 = \cdots = X_m = 1\} \cap \{X_{m+1} = \cdots = X_{m+n} = 0\} \cap \{X_{m+n+1} = 1\}\right)\\ &+ \mathbb{P}\left(\{X_1 = \cdots = X_m = 0\} \cap \{X_{m+1} = \cdots = X_{m+n} = 1\} \cap \{X_{m+n+1} = 0\}\right)\\ &= p^{m+1}q^n + q^{m+1}p^n.\end{aligned}$$

  On en déduit, pour tout $m \in \mathbb{N}^*$:

$$\mathbb{P}(L_1 = m) = \sum_{n=1}^{+\infty}\left(p^{m+1}q^n + q^{m+1}p^n\right) = p^{m+1}\frac{q}{1-q} + q^{m+1}\frac{p}{1-p} = qp^m + pq^m.$$

  On trouve de même que, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(L_2 = n) = p^2q^{n-1} + q^2p^{n-1}.$$

  On vérifie au passage que $\displaystyle\sum_{m=1}^{+\infty}\mathbb{P}(L_1 = m) = \sum_{n=1}^{+\infty}\mathbb{P}(L_2 = n) = 1$, ce qui montre que L_1 et L_2 sont bien des variables aléatoires à valeurs dans $\mathbb{N}^*$, définies presque sûrement.

- On calcule l'espérance de L_1 et de L_2. On obtient :

$$\begin{aligned}\mathbb{E}(L_1) &= \sum_{m=1}^{+\infty} m(qp^m + pq^m) = pq\sum_{m=1}^{+\infty}(mp^{m-1} + mq^{m-1})\\ &= pq\left(\frac{1}{(1-p)^2} + \frac{1}{(1-q)^2}\right) = \frac{p}{q} + \frac{q}{p}.\end{aligned}$$

On trouve de même :

$$\mathbb{E}(L_2) = \frac{p^2}{(1-q)^2} + \frac{q^2}{(1-p)^2} = 2.$$

- On a

$$\begin{aligned}
\mathbb{E}(L_1 \, L_2) &= \sum_{(m,n)\in(\mathbb{N}^*)^2} mn\mathbb{P}(L_1 = m, L_2 = n) \\
&= \sum_{(m,n)\in(\mathbb{N}^*)^2} mn(p^{m+1}q^n + q^{m+1}p^n) \\
&= \left(\sum_{m=1}^{+\infty} mp^{m+1}\right)\left(\sum_{n=1}^{+\infty} nq^n\right) + \left(\sum_{m=1}^{+\infty} mq^{m+1}\right)\left(\sum_{n=1}^{+\infty} np^n\right) \\
&= \frac{p^2}{(1-p)^2}\,\frac{q}{(1-q)^2} + \frac{q^2}{(1-q)^2}\,\frac{p}{(1-p)^2} = \frac{1}{q} + \frac{1}{p}\cdot
\end{aligned}$$

Finalement, on obtient :

$$\begin{aligned}
\mathrm{Cov}(L_1, L_2) = \mathbb{E}(L_1\,L_2) - \mathbb{E}(L_1)\,\mathbb{E}(L_2) &= \frac{1}{q} + \frac{1}{p} - \frac{2p}{q} - \frac{2q}{p} \\
&= -\frac{2p^2 + 2q^2 - p - q}{pq} = -\frac{2p^2 + 2q^2 - (p+q)^2}{pq} \\
&= -\frac{(p-q)^2}{pq}\cdot
\end{aligned}$$

Proposition 51

Soit X, Y et X' trois variables aléatoires appartenant à $L_2(\Omega, \mathcal{A}, \mathbb{P})$ et $\lambda \in \mathbb{R}$.

- On a $\mathrm{Cov}(X,X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{V}(X) \geqslant 0$.
- Par définition, on a :

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}\Big(\big(X - \mathbb{E}(X)\big)\big(Y - \mathbb{E}(Y)\big)\Big) \\
&= \mathbb{E}\Big(\big(Y - \mathbb{E}(Y)\big)\big(X - \mathbb{E}(X)\big)\Big) = \mathrm{Cov}(Y,X).
\end{aligned}$$

- D'après le théorème 43, $X + X'$ possède un moment d'ordre 2. Ainsi le couple $(X + X', Y)$ possède une covariance. Par linéarité de l'espérance, on obtient :

$$\begin{aligned}
\mathrm{Cov}(X + X', Y) &= \mathbb{E}\big((X + X')Y\big) - \mathbb{E}(X + X')E(Y) \\
&= \mathbb{E}(XY) + \mathbb{E}(X'Y) - \big(\mathbb{E}(X) + \mathbb{E}(X')\big)\mathbb{E}(Y) \\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X'Y) - \mathbb{E}(X')\mathbb{E}(Y) \\
&= \mathrm{Cov}(X,Y) + \mathrm{Cov}(X',Y),
\end{aligned}$$

  et de même :

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

- La linéarité par rapport à la deuxième variable résulte de la symétrie et de la linéarité par rapport à la première variable.

Exercice 26 La matrice M est symétrique réelle donc diagonalisable. Soit λ une valeur propre et $Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$ un vecteur propre associé. On a par bilinéarité :

$$\mathbb{V}\left(\sum_{i=1}^{n} z_i X_i\right) = \sum_{1\leqslant i,j\leqslant n} z_i z_j \operatorname{Cov}(X_i, X_j) = {}^t Z M Z = {}^t Z \lambda Z = \lambda \|Z\|^2.$$

On a $\mathbb{V}\left(\sum\limits_{i=1}^{n} z_i X_i\right) \geqslant 0$ et $\|Z\|^2 > 0$, donc $\lambda \geqslant 0$.

Si $\lambda = 0$, on obtient $\mathbb{V}\left(\sum\limits_{i=1}^{n} z_i X_i\right) = 0$ et donc $\sum\limits_{i=1}^{n} z_i X_i = 0$ presque sûrement.

Réciproquement, s'il existe $(z_1, \dots, z_n) \in \mathbb{R}^n$ non nul tel que $\sum\limits_{i=1}^{n} z_i X_i = 0$ presque sûrement, on a, pour tout $i \in [\![1, n]\!]$:

$$0 = \operatorname{Cov}\left(X_i, \sum_{j=1}^{n} z_j X_j\right) = \sum_{j=1}^{n} \operatorname{Cov}(X_i, X_j) z_j,$$

donc 0 est valeur propre de M, un vecteur propre associé étant $(z_1, z_2, \dots, z_n)$.

Théorème 52 Les variables X_i, pour $1 \leqslant i \leqslant n$, appartiennent à l'espace vectoriel $L_2(\Omega, \mathcal{A}, \mathbb{P})$ donc il en est de même de $X_1 + \cdots + X_n$ et on obtient, par bilinéarité de la covariance :

$$\mathbb{V}(X_1 + \cdots + X_n) = \operatorname{Cov}(X_1 + \cdots + X_n, X_1 + \cdots + X_n) = \sum_{1\leqslant i,j\leqslant n} \operatorname{Cov}(X_i, X_j).$$

Sachant que $\operatorname{Cov}(X_i, X_i) = \mathbb{V}(X_i)$ et $\operatorname{Cov}(X_i, X_j) = \operatorname{Cov}(X_j, X_i)$, on obtient :

$$\mathbb{V}(X + \cdots + X_n) = \sum_{k=1}^{n} \mathbb{V}(X_k) + 2 \sum_{1\leqslant i<j\leqslant n} \operatorname{Cov}(X_i, X_j).$$

Théorème 54 Par linéarité de l'espérance, on a $\mathbb{E}\left(\dfrac{S_n}{n}\right) = \dfrac{1}{n} \sum\limits_{k=1}^{n} \mathbb{E}(X_k) = m$. D'autre part, comme les variables X_k sont deux à deux indépendantes et possèdent une variance, d'après le corollaire 53 de la page 919, S_n possède une variance et, en notant $\sigma^2 = \mathbb{V}(X_1)$:

$$\mathbb{V}\left(\frac{S_n}{n}\right) = \frac{\mathbb{V}(S_n)}{n^2} = \frac{1}{n^2} \sum_{k=1}^{n} \mathbb{V}(X_k) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}\cdot$$

On a donc, pour $\varepsilon > 0$, d'après l'inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)\right| \geqslant \varepsilon\right) \leqslant \frac{\mathbb{V}\left(\dfrac{S_n}{n}\right)}{\varepsilon^2} \leqslant \frac{\sigma^2}{n\varepsilon^2}\cdot$$

Le résultat découle de cette inégalité.

Proposition 55 On a, pour tous $t \in [-1,1]$ et $n \in \mathbb{N}$, $|\mathbb{P}(X=n)t^n| \leqslant \mathbb{P}(X=n)$. Comme $\sum \mathbb{P}(X=n)$ converge, la série $\sum \mathbb{P}(X=n)t^n$ converge normalement (et donc absolument) sur $[-1,1]$. Donc G_X est définie sur $[-1,1]$ et est continue sur $[-1,1]$, car somme d'une série de fonctions continues convergeant normalement.

La convergence sur $[-1,1]$ montre que le rayon de convergence de la série entière est supérieur ou égale à 1.

Proposition 56 Cela résulte immédiatement de la formule donnant les coefficients du développement d'une fonction en série entière.

Lemme 57 On note R le rayon de convergence de la série $\sum a_n t^n$. Comme la série converge pour $t=1$, on a $R \geqslant 1$. La fonction f est dérivable sur $]-R,R[$ et :

$$\forall t \in]-R,R[\quad f'(t) = \sum_{n=1}^{+\infty} n a_n t^{n-1}.$$

Si $R>1$, la fonction f est dérivable en 1 et $f'(1) = \sum\limits_{n=1}^{+\infty} n a_n$, donc $\sum n a_n$ converge.

On suppose désormais $R=1$. On a, pour tout $t \in [0,1[$,

$$\frac{f(t)-f(1)}{t-1} = \sum_{n=1}^{+\infty} a_n \frac{t^n-1}{t-1} = \sum_{n=1}^{+\infty} a_n (t^{n-1} + \cdots + t + 1).$$

La fonction $t \longmapsto \dfrac{f(t)-f(1)}{t-1}$ est donc croissante sur $[0,1[$.

- Si la fonction f est dérivable en 1, on a pour tout $t \in [0,1[$, $\dfrac{f(t)-f(1)}{t-1} \leqslant f'(1)$ et *a fortiori* pour tout entier $N \in \mathbb{N}^*$:

  $$\sum_{n=1}^{N} a_n (t^{n-1} + \cdots + t + 1) \leqslant f'(1).$$

  En faisant tendre t vers 1, on obtient $\sum\limits_{n=1}^{N} n a_n \leqslant f'(1)$. Comme cela est vrai pour tout $N \in \mathbb{N}^*$, la série à termes positifs $\sum n a_n$ converge et $\sum\limits_{n=1}^{+\infty} n a_n \leqslant f'(1)$.

- Réciproquement, si la série de terme général $n a_n$ converge, on a, pour tout $t \in [0,1[$,

  $$\frac{f(t)-f(1)}{t-1} \leqslant \sum_{n=1}^{+\infty} n a_n.$$

  La fonction $t \longmapsto \dfrac{f(t)-f(1)}{t-1}$ est croissante et majorée. Elle possède donc une limite finie en 1, inférieure à $\displaystyle\sum_{n=1}^{+\infty} n a_n$. La fonction f est donc dérivable en 1 et $f'(1) \leqslant \sum\limits_{n=1}^{+\infty} n a_n$.

Dans le cas où f est dérivable, on a donc $f'(1) = \sum\limits_{n=1}^{+\infty} n a_n$.

Théorème 58 D'après le lemme 57, la fonction G_X est dérivable en 1 si, et seulement si, la série de terme général $n\mathbb{P}(X=n)$ converge, c'est-à-dire si, et seulement si, X est d'espérance finie et l'on a alors $G'_X(1)=\mathbb{E}(X)$.

Théorème 59 La variable X possède un moment d'ordre 2 si, et seulement si, la série de terme général $n^2\mathbb{P}(X=n)$ converge, c'est-à-dire si, et seulement si, la série de terme général $n(n-1)\mathbb{P}(X=n)$ converge. En effet, ces séries à termes positifs ont des termes généraux équivalents. En cas de convergence $\sum\limits_{n=2}^{+\infty} n(n-1)\mathbb{P}(X=n)=\mathbb{E}(X(X-1))$, d'après la formule de transfert.

- Si la fonction G_X est deux fois dérivable en 1, elle est en particulier dérivable en 1 et d'après le lemme 57 on a :

$$\forall t\in[0,1]\quad G'_X(t)=\sum_{n=1}^{+\infty}\mathbb{P}(X=n)nt^{n-1}=\sum_{n=0}^{+\infty}\mathbb{P}(X=n+1)(n+1)t^n.$$

  Toujours d'après le lemme 57, appliqué à la série définissant G'_X, la dérivabilité de G'_X en 1 implique la convergence de la série de terme général $n\mathbb{P}(X=n+1)(n+1)$. De plus, on a :

$$G''_X(1)=\sum_{n=1}^{\infty}n\mathbb{P}(X=n+1)(n+1)=\sum_{n=2}^{\infty}n(n-1)\mathbb{P}(X=n).$$

- Réciproquement, si la série de terme général $n(n-1)\mathbb{P}(X=n)$ converge, alors la série de terme général $nP(X=n)$ converge *a fortiori* donc, d'après le théorème 58, G_X est dérivable en 1 et :

$$\forall t\in[0,1]\quad G'_X(t)=\sum_{n=1}^{+\infty}\mathbb{P}(X=n)nt^{n-1}=\sum_{n=0}^{+\infty}\mathbb{P}(X=n+1)(n+1)t^n.$$

  La série de terme général $n(n+1)\mathbb{P}(X=n+1)$ converge donc, d'après le lemme 57, G'_X est dérivable en 1 et, d'après le premier point, $G''_X(1)=\sum\limits_{n=2}^{\infty}n(n-1)\mathbb{P}(X=n)$.

Exercice 27

1. La variable X est finie donc G_X est définie sur $\mathbb{R}$ par :
$$G_X(t)=\mathbb{P}(X=0)t^0+\mathbb{P}(X=1)t=(1-p)+pt.$$
   On en déduit, pour tout réel t, $G'_X(t)=p$ et $G''_X(t)=0$. On a donc $\mathbb{E}(X)=p$ et $\mathbb{E}(X^2)=p+0=p$, puis $\mathbb{V}(X)=p-p^2=p(1-p)$.
2. La variable X est finie donc G_X est définie sur $\mathbb{R}$ par :
$$\begin{aligned}G_X(t)=\sum_{k=0}^{n}\mathbb{P}(X=k)t^k&=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}t^k\\&=\sum_{k=0}^{n}\binom{n}{k}(pt)^k(1-p)^{n-k}=\left((1-p)+pt\right)^n.\end{aligned}$$

On en déduit, pour tout réel t :

$$G'_X(t) = np\,((1-p)+pt)^{n-1} \quad ;$$
$$G''_X(t) = n(n-1)p^2\,((1-p)+pt)^{n-2}.$$

On a donc $\mathbb{E}(X) = np$ et $\mathbb{E}(X^2) = np + n(n-1)p^2$, puis :

$$\mathbb{V}(X) = np + n(n-1)p^2 - (np)^2 = np(1-p).$$

3. La série de terme général $\mathbb{P}(X=n)t^n = pt(t(1-p))^{n-1}$ est une série géométrique qui converge si $\left|p(1-t)\right| < 1$. On a :

$$\forall t \in \left]-\frac{1}{1-p}, \frac{1}{1-p}\right[\quad G_X(t) = \sum_{n=1}^{\infty} pt(t(1-p))^{n-1} = \frac{pt}{1-(1-p)t}.$$

On en déduit, pour tout $t \in \left]-\dfrac{1}{1-p}, \dfrac{1}{1-p}\right[$:

$$G'_X(t) = \frac{p}{(1-(1-p)t)^2} \quad \text{et} \quad G''_X(t) = \frac{2p(1-p)}{(1-(1-p)t)^3}.$$

On a donc :

$$\mathbb{E}(X) = G'_X(1) = \frac{p}{p^2} = \frac{1}{p}$$
$$\mathbb{E}(X^2) = G'_X(1) + G''_X(1) = \frac{1}{p} + \frac{2(1-p)}{p^2},$$

puis $\mathbb{V}(X) = \dfrac{1}{p} + \dfrac{2(1-p)}{p^2} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$.

4. La série de terme général $\mathbb{P}(X=n)t^n = e^{-\lambda}\dfrac{(\lambda t)^n}{n!}$ converge pour tout réel t donc :

$$\forall t \in \mathbb{R} \quad G_X(t) = e^{-\lambda}\sum_{n=0}^{+\infty} \frac{(\lambda t)^n}{n!} = e^{-\lambda}e^{\lambda t} = e^{\lambda(t-1)}.$$

On a, pour tout réel t, $G'_X(t) = \lambda e^{\lambda(t-1)}$ et $G''_X(t) = \lambda^2 e^{\lambda(t-1)}$. On en déduit $\mathbb{E}(X) = G'_X(1) = \lambda$ et $\mathbb{E}(X^2) = G'_X(1) + G''_X(1) = \lambda + \lambda^2$, puis $\mathbb{V}(X) = \lambda$.

Théorème 61 Supposons que $G_X(t)$ et $G_Y(t)$ soient définis, c'est-à-dire que t^X et t^Y aient une espérance. Les variables aléatoires t^X et t^Y étant indépendantes, car fonctions de variables aléatoires indépendantes, $t^X t^Y = t^{X+Y}$ possède une espérance et :

$$G_{X+Y}(t) = \mathbb{E}\left(t^X t^Y\right) = \mathbb{E}\left(t^X\right)\mathbb{E}\left(t^Y\right) = G_X(t)G_Y(t).$$

Exercice 28 Les variables X, Z et Y (car $Y \leqslant Z$) sont finies, donc leur fonction génératrice est définie sur $\mathbb{R}$. Pour tout $t \in \mathbb{R} \setminus \{1\}$, on a :

$$G_X(t) = \frac{1}{\ell}\sum_{i=0}^{\ell-1} t^i = \frac{t^\ell - 1}{\ell(t-1)} \text{ et, de même, } G_Z(t) = \frac{t^n - 1}{n(t-1)}.$$

Les variables X et Y étant indépendantes, on a $G_Z = G_X\,G_Y$ et donc :

$$\forall t \in \mathbb{R} \setminus \{1\} \quad G_Y(t) = \frac{G_Z(t)}{G_X(t)} = \frac{\ell(t^n - 1)}{n(t^\ell - 1)} = \frac{t^{\ell m} - 1}{m(t^\ell - 1)} = \frac{1}{m}\sum_{i=0}^{m-1} t^{\ell i}.$$

Cette égalité est encore valable pour $t = 1$ par continuité.

On en déduit que $\mathbb{P}(Y = \ell i) = \dfrac{1}{m}$, pour tout $i \in [\![0, m-1]\!]$. Ainsi Y suit la loi uniforme sur $\{0, \ell, 2\ell, \dots, (m-1)\ell\}$.

Exercice 29

1. Soit X_1, …, X_k des variables aléatoires indépendantes, telles que, pour tout $i \in [\![1, k]\!]$, la variable X_i suive la loi binomiale de paramètre (n_i, p). On pose $X = \sum\limits_{i=1}^{n} X_i$. On a alors, pour tout réel t,

$$G_X(t) = \prod_{i=1}^{k} G_{X_i}(t) = \prod_{i=1}^{k}(1 - p + pt)^{n_i} = (1 - p + pt)^n,$$

   où $n = \sum\limits_{i=1}^{k} n_i$. On reconnaît la fonction génératrice d'une loi binomiale de paramètre (n, p).

   D'après la proposition 56, X suit la loi binomiale de paramètre (n, p).

2. Soit X_1, …, X_k des variables aléatoires indépendantes, telles que, pour tout $i \in [\![1, k]\!]$, la variable X_i suive la loi de Poisson de paramètre λ_i. On pose $X = \sum\limits_{i=1}^{n} X_i$. On a alors, pour tout réel t :

$$G_X(t) = \prod_{i=1}^{k} G_{X_i}(t) = \prod_{i=1}^{k} e^{\lambda_i(t-1)} = e^{\lambda(t-1)},$$

   où $\lambda = \sum\limits_{i=1}^{k} \lambda_i$. On reconnaît la fonction génératrice d'une loi de Poisson de paramètre λ. D'après la proposition 56, X suit la loi de Poisson de paramètre λ.

S'entraîner et approfondir

16.1 Soit X une variable aléatoire discrète, à valeur dans $\mathbb{N}$ et $f : \mathbb{N} \longrightarrow \mathbb{Z}$ définie par :

$$f(n) = \frac{n}{2} \quad \text{si} \quad n \text{ est pair} \qquad \text{et} \qquad f(n) = \frac{1-n}{2} \quad \text{si} \quad n \text{ est impair.}$$

On pose $Y = f(X)$. Déterminer la loi de Y et son espérance dans les cas suivants :

$$X \sim \mathcal{G}(p) \qquad\qquad X \sim \mathcal{P}(\lambda).$$

16.2 Soit X et Y deux variables indépendantes suivant toutes deux la loi géométrique de paramètre $\frac{1}{2}\cdot$ Déterminer la probabilité des événements suivants :

1. $\mathbb{P}(X = Y)$;
2. $\mathbb{P}(X > Y)$;
3. $\mathbb{P}(X \geqslant kY)$ pour un entier $k \geqslant 1$ fixé.

16.3 *(Mines-Ponts 2015)*
Soit X et Y deux variables aléatoires indépendantes suivant une loi géométrique de même paramètre $p \in \,]0,1[$. Donner la loi de $|X - Y|$.

16.4 Dans une salle de cinéma, il arrive X personnes souhaitant voir le film. On suppose que X suit la loi géométrique de paramètre p. La capacité de la salle est de n places. On note Y le nombre de personnes ne pouvant entrer dans la salle.

1. Déterminer la loi de Y.
2. Déterminer la fonction génératrice de Y.
3. Calculer l'espérance de Y.

16.5 On considère X et Y deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $\mathbb{N}$, indépendantes, de même loi. On pose $D = X - Y$ et $I = \min(X, Y)$.

1. On suppose que pour tout k dans $\mathbb{N}$, $\mathbb{P}(X = k) = pq^k$, où $p \in \,]0,1[$ et $q = 1 - p$.
   (a) Déterminer la loi conjointe de (D, I).
   (b) Déterminer les lois marginales de D et I. Vérifier que D et I sont indépendantes.
2. On suppose que les variables D et I sont indépendantes et que $\mathbb{P}(X = n) \neq 0$ pour tout $n \in \mathbb{N}$.
   Montrer qu'il existe $p \in \,]0,1[$, tel que, pour tout $k \in \mathbb{N}$,
   $$\mathbb{P}(X = k) = pq^k.$$

16.6 (Une caractérisation de la loi de Poisson)

On considère une variable aléatoire discrète N sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ telle que $N(\Omega) = \mathbb{N}$ et $\mathbb{P}(N = n) \neq 0$ pour tout $n \in \mathbb{N}$. Si la variable aléatoire N prend la valeur n, on procède à une succession de n épreuves de Bernoulli indépendantes de paramètre $p \in]0,1[$. On note S et E les variables aléatoires représentant respectivement le nombre de succès et d'échecs dans ces n épreuves.

1. Montrer que si N suit une loi de Poisson de paramètre $\lambda > 0$, les variables S et E suivent aussi des lois de Poisson dont on déterminera les paramètres. Montrer que les variables E et S sont indépendantes.
2. Montrer réciproquement que si S et E sont indépendantes, alors N suit une loi de Poisson. Pour cela, on montrera :
   (a) qu'il existe deux suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ telles que :
   $$\forall (m,n) \in \mathbb{N}^2 \quad (m+n)!\,\mathbb{P}(N = m+n) = u_m\, v_n \ ;$$
   (b) que les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont géométriques.

16.7 On lance une pièce de monnaie (la probabilité d'obtenir pile étant $p \in]0,1[$) jusqu'à l'obtention du premier pile. Soit N la variable aléatoire représentant le nombre de lancers nécessaire. Si $N = n$, on relance ensuite n fois la pièce et on appelle X la variable aléatoire représentant le nombre de piles obtenu.

1. Déterminer la loi de N, celle du couple (N, X), puis la loi de X.
2. Montrer que X a même loi que le produit de deux variables indépendantes Y et Z telles que Y suive une loi de Bernoulli et Z une loi géométrique de même paramètre.
3. En déduire l'espérance et la variance de X.

16.8 Soit $\mathcal{P}$ l'ensemble des nombres premiers. Pour $s > 1$, on note $\zeta(s) = \sum_{n\geqslant 1} n^{-s}$ et X une variable aléatoire à valeurs dans $\mathbb{N}^*$ dont la loi est définie par :

$$\forall n \in \mathbb{N}^* \quad \mathbb{P}(X = n) = \frac{n^{-s}}{\zeta(s)}\cdot$$

1. Justifier qu'on définit bien ainsi la loi d'une variable aléatoire.
2. Pour tout $n \in \mathbb{N}^*$, on considère A_n : « n divise X ». Montrer que $(A_p)_{p\in\mathcal{P}}$ est une famille d'événements indépendants. En déduire une preuve probabiliste de :
   $$\prod_{p\in\mathcal{P}} \left(1 - \frac{1}{p^s}\right) = \frac{1}{\zeta(s)}\cdot$$
3. Montrer que la probabilité qu'aucun carré différent de 1 ne divise X vaut $\dfrac{1}{\zeta(2s)}\cdot$

16.9 (Taux de panne)

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, X une variable aléatoire discrète, définie sur cet espace probabilisé, à valeurs dans $\mathbb{N}^*$, et vérifiant :

$$\forall n \in \mathbb{N}^* \quad \mathbb{P}(X \geqslant n) > 0. \tag{1}$$

On appelle taux de panne associé à X la suite réelle $(x_n)_{n\in\mathbb{N}^*}$ définie par :

$$\forall n \in \mathbb{N}^* \quad x_n = \mathbb{P}(X = n \mid X \geqslant n).$$

1. Exprimer $p_n = \mathbb{P}(X = n)$ en fonction des x_k.
2. (a) Montrer que l'on a $0 \leqslant x_n < 1$, pour tout $n \in \mathbb{N}^*$, et que la série de terme général x_n diverge.
   (b) Réciproquement, soit $(x_n)_{n\in\mathbb{N}^*}$ une suite à valeur dans $[0,1[$ telle que la série de terme général x_n diverge. Montrer qu'il existe une variable aléatoire dont le taux de panne est la suite (x_n).
3. Montrer que la variable X suit une loi géométrique si, et seulement si, son taux de panne est constant.

⋆ **16.10** Soit $n \in \mathbb{N}^*$. On désigne par Ω l'ensemble des permutations de $[\![1,n]\!]$. On munit Ω de la probabilité uniforme. Pour $\sigma \in \Omega$ et $i \in [\![1,n]\!]$, on dit que $\sigma(i)$ est un maximum (resp.minimum) provisoire de σ si :

$$\sigma(i) = \max\big(\sigma(1), \sigma(2), \ldots, \sigma(i)\big) \quad \big(\text{resp. } \sigma(i) = \min\big(\sigma(1), \sigma(2), \ldots, \sigma(i)\big)\big).$$

On désigne par X_n (resp. Y_n) les variables aléatoires représentant le nombre de maximums (resp. minimums) provisoires des permutations de $[\![1,n]\!]$.

1. Montrer que les variables X_n et Y_n ont même loi.
2. (a) Déterminer la loi de X_3, son espérance et sa variance.
   (b) Déterminer la loi du couple (X_3, Y_3) et sa covariance.
3. Pour $n \in \mathbb{N}^*$, on note g_n la fonction génératrice de X_n.
   (a) Pour $1 \leqslant k \leqslant n$, on note Z_k la variable indicatrice de l'événement « $\sigma(k)$ est un maximum provisoire ».
   Montrer que les variables Z_1, Z_2,..., Z_n sont indépendantes.
   (b) Exprimer X_n en fonction de Z_1, Z_2,..., Z_n. En déduire g_n.
   (c) En déduire $\mathbb{P}(X_n = 1)$, $\mathbb{P}(X_n = 2)$, $\mathbb{P}(X_n = n)$.
   (d) Déterminer $\mathbb{E}(X_n)$ et $\mathbb{V}(X_n)$ (sous forme de sommes) et un équivalent de $\mathbb{E}(X_n)$ et de $\mathbb{V}(X_n)$ quand n tend vers $+\infty$.

16.11 (Formule du crible)

1. Soit $A_1, A_2, \ldots, A_n$ des événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Montrer que $\mathbb{1}_{\bigcup_{i=1}^n A_i} = 1 - \prod_{i=1}^n (1 - \mathbb{1}_{A_i})$. En déduire la *formule du crible* :

$$\mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big) = \sum_{k=1}^n \bigg((-1)^{k-1} \sum_{\substack{I\subset[\![1,n]\!]\\ \text{card } I=k}} \mathbb{P}\Big(\bigcap_{i\in I} A_i\Big)\bigg).$$

2. Soit $n \in \mathbb{N}^*$ et $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables indépendantes d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, suivant toutes la loi uniforme sur $[\![1,n]\!]$.
   On note X la variable aléatoire égale au nombre de tirages nécessaires pour obtenir tous les numéros entre 1 et n au moins une fois (et à $+\infty$ si on n'obtient jamais les n numéros). Pour $j \in [\![1,n]\!]$ et $m \in \mathbb{N}$, on note $B_{j,m}$ l'événement : « au bout de m tirages, le numéro j n'est pas encore apparu ».
   (a) Calculer $\mathbb{P}(B_{j_1,m} \cap B_{j_2,m} \cap \cdots \cap B_{j_k,m})$, où j_1, j_2, ..., j_k sont des indices distincts compris entre 1 et n.
   (b) En déduire que $\mathbb{P}(X > m) = \displaystyle\sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} \left(\frac{n-k}{n}\right)^m$.
   Calculer $\displaystyle\lim_{m\to+\infty} \mathbb{P}(X > m)$. Interpréter.
   (c) Montrer que $\mathbb{E}(X) = n \displaystyle\sum_{k=1}^{n} (-1)^{k-1} \frac{\binom{n}{k}}{k}$, en utilisant l'exercice 18 de la page 911
   .
   (d) Montrer que $\mathbb{E}(X) = n\left(1 + \dfrac{1}{2} + \cdots + \dfrac{1}{n}\right)$. En déduire un équivalent de $\mathbb{E}(X)$ quand n tend vers $+\infty$.

16.12 Soit $n \geqslant 1$ un entier, et $(U_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de même loi uniforme sur $[\![1,n]\!]$. Pour tout $i \in [\![1,n]\!]$, on définit :

$$X_i^{(m)} = \operatorname{card}\{k \in [\![1,m]\!] \mid U_k = i\} \quad \text{si} \quad m \geqslant 1 \qquad \text{et} \qquad X_i^{(0)} = 0.$$

1. Quelle est la loi de $X_i^{(m)}$ pour $i \in [\![1,n]\!]$ et $m \geqslant 1$?
2. Soit $m \geqslant 1$ et $(i,j) \in [\![1,n]\!]^2$, avec $i \neq j$. Calculer la covariance des variables aléatoires $X_i^{(m)}$ et $X_j^{(m)}$. Sont-elles indépendantes ?
3. Soit $\lambda > 0$ et N une variable aléatoire suivant une loi de Poisson de paramètre λ, indépendante des variables U_k. On pose

$$\forall i \in [\![1,n]\!] \quad Y_i = X_i^{(N)}.$$

   (a) Déterminer, en fonction de λ et n, la loi de Y_i pour tout $i \in [\![1,n]\!]$
   (b) Déterminer la loi conjointe de $(Y_1, \ldots, Y_n)$.

$\star\star$ **16.13** *(Centrale 2015)*
Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(E_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ une suite d'événements. On suppose que $\sum\limits_{n=0}^{+\infty} \mathbb{P}(E_n) < +\infty$, c'est-à-dire que la série converge.

1. On note $\mathbb{1}_X$ la fonction indicatrice d'un ensemble X. Soit $Z = \sum\limits_{n=0}^{+\infty} \mathbb{1}_{E_n}$ (on convient que $Z = +\infty$ si la série diverge). Prouver que Z est une variable aléatoire discrète.
2. Soit $F = \{\omega \in \Omega \mid \omega \text{ appartient à un nombre fini de } E_n \text{ (pour } n \in \mathbb{N})\}$.
   Prouver que F est un événement et que $\mathbb{P}(F) = 1$.
3. Prouver que Z est d'espérance finie.

16.14 Une action vaut initialement 1 euro. À chaque instant $n \geqslant 1$, sa valeur est multipliée par une quantité aléatoire Z_n. On suppose que les variables Z_n sont indépendantes et de même loi, telles que :

$$\mathbb{P}(Z_n = 1 + a) = \mathbb{P}(Z_n = 1 - a) = 1/2, \quad \text{avec} \quad a \in]0,1[.$$

On note X_n la valeur de l'action à l'instant n. On pose $Y_k = \ln(Z_k)$, et l'on définit, pour tout entier naturel non nul n, la variable :

$$\widehat{Y}_n = \frac{Y_1 + \cdots + Y_n}{n}\cdot$$

1. Montrer que, pour tout $n \geqslant 0$, on a $\mathbb{E}(X_n) = 1$.
2. Calculer la limite de $\mathbb{V}(X_n)$ quand n tend vers l'infini ?
3. Montrer qu'il existe $\delta > 0$ tel que

$$\lim_{n \to +\infty} \mathbb{P}\left(\widehat{Y}_n > -\delta\right) = 0.$$

4. En déduire que pour tout $\varepsilon > 0$,

$$\lim_{n \to +\infty} \mathbb{P}(X_n > \varepsilon) = 0.$$

⋆ **16.15** *(Centrale 2015)*

On dispose de n urnes et de $N = na$ boules, où a et n sont des entiers naturels non nuls. Ces boules sont réparties de façon indépendante et équiprobable entre les urnes. On nomme Y_n la variable aléatoire donnant le nombre d'urnes vides et $S_n = \dfrac{Y_n}{n}\cdot$

1. Énoncer et démontrer l'inégalité de Bienaymé-Tchebychev.
2. Calculer $\mathbb{E}(S_n)$ et $\mathbb{V}(S_n)$.
3. Montrer :

$$\forall \varepsilon > 0 \qquad \lim_{n \to +\infty} \mathbb{P}(|S_n - e^{-a}| \geqslant \varepsilon) = 0.$$

16.16 (**Modèle de Galton-Watson)**

On observe des virus qui se reproduisent tous selon la même loi avant de mourir : un virus donne naissance en une journée à X virus, où X est une variable aléatoire à valeurs dans $\mathbb{N}$. Pour tout $k \in \mathbb{N}$, on note $\mathbb{P}(X = k) = p_k$. On suppose $p_1 > 0$ et $p_0 + p_1 < 1$. On note f la fonction génératrice de X.

On part au jour zéro de $X_0 = 1$ virus. Au premier jour, on a donc X_1 virus, où X_1 suit la loi de X ; chacun de ces X_1 virus évolue alors indépendamment des autres virus et se reproduit selon la même loi avant de mourir : cela conduit à avoir X_2 virus au deuxième jour ; et le processus continue de la sorte. On note $u_n = \mathbb{P}(X_n = 0)$.

1. Calculer u_0, u_1.
2. Montrer que la suite $(u_n)_{n \in \mathbb{N}}$ est convergente.
3. Montrer que pour tout entier $n \geqslant 0$, on a $u_{n+1} = f(u_n)$.
4. Que peut-on dire de la limite de $(u_n)_{n \in \mathbb{N}}$. Discuter selon la valeur de $\mathbb{E}(X)$. Interpréter le résultat.

⋆ 16.17 (Somme aléatoire de variables aléatoires)
Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles discrètes, toutes de même loi, et N une variable aléatoires à valeurs dans $\mathbb{N}$. On suppose que N et les variables X_n, pour $n \in \mathbb{N}^*$, forment une suite de variables aléatoires indépendantes. On pose :

$$\forall n \in \mathbb{N}^* \quad S_n = \sum_{k=1}^{n} X_k \qquad \text{et} \qquad S_0 = 0.$$

1. Montrer que S_N est une variable aléatoire.
2. (a) Déterminer la loi de S_N, lorsque les X_k suivent la loi de Bernoulli de paramètre p et N la loi de Poisson de paramètre λ.
   (b) Déterminer la loi de S_N lorsque les X_k suivent la loi géométrique de paramètre p et N la loi géométrique de paramètre p'.
3. On suppose que les variables aléatoires X_n sont à valeurs dans $\mathbb{N}$.
   (a) Montrer que $G_{S_N} = G_N \circ G_{X_1}$ sur $[0,1]$.
   (b) Montrer que, si X_1 et N sont d'espérance finie, alors S_N est d'espérance finie et vérifie la première *formule de Wald* :
   $$\mathbb{E}(S_N) = \mathbb{E}(X_1)\,\mathbb{E}(N).$$
   (c) Montrer que, si X_1 et N possèdent un moment d'ordre 2, alors S_N possède aussi un moment d'ordre 2 et vérifie la seconde *formule de Wald* :
   $$\mathbb{V}(S_N) = \mathbb{V}(X_1)\,\mathbb{E}(N) + \left(\mathbb{E}(X_1)\right)^2\,\mathbb{V}(N).$$
4. On revient au cas général. On suppose que X_1 et N sont d'espérance finie.
   (a) Démontrer que la famille $\left(x\,\mathbb{P}(S_n = x)\,\mathbb{P}(N = n)\right)_{(x,n)\in S_N(\Omega)\times N(\Omega)}$ est sommable.
   (b) En déduire que S_N est d'espérance finie et :
   $$\mathbb{E}(S_N) = \mathbb{E}(X_1)\mathbb{E}(N).$$

16.18 On considère une suite d'épreuves de Bernoulli indépendantes. À chaque épreuve, la probabilité de succès est $p \in \,]0,1[$. On se donne un entier r strictement positif. Pour $n \in \mathbb{N}^*$, on note Π_n la probabilité qu'au cours des n premières épreuves, on ait obtenu r succès consécutifs (au moins une fois).

1. (a) Calculer Π_0, Π_1, …, Π_r.
   (b) Montrer que, pour $n \geqslant r$, on a $\Pi_{n+1} = \Pi_n + (1-\Pi_{n-r})p^r(1-p)$.
   (c) Montrer que la suite $(\Pi_n)_{n\in\mathbb{N}}$ est convergente. Calculer sa limite.
2. (a) Déduire de la question 1 que l'on peut définir une variable aléatoire T égale au temps d'attente de r succès consécutifs. On définira $\{T = k\}$ comme l'événement « on a obtenu des succès aux épreuves de rang $k-r+1$, $k-r+2$, …, k sans jamais avoir obtenu r succès consécutifs auparavant ».
   (b) Montrer en utilisant le résultat de l'exercice 18 de la page 911 que :
   $$\mathbb{E}(T) = \frac{1-p^r}{(1-p)p^r}\cdot$$

16.19 (Marche aléatoire dans $\mathbb{Z}$: premier retour à l'origine)
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires, sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes et de même loi définie par

$$\mathbb{P}(X_n = 1) = p \qquad \text{et} \qquad \mathbb{P}(X_n = -1) = 1 - p,$$

où $p \in [0,1]$. On pose $S_0 = 0$ et, pour tout $n \in \mathbb{N}^*$, $S_n = \displaystyle\sum_{k=1}^{n} X_k$.

La suite (S_n) est appelé *marche aléatoire dans* $\mathbb{Z}$. On peut imaginer un mobile partant de l'origine et se déplaçant à chaque instant (entier) de ± 1, les déplacements successifs étant indépendants. Alors S_n représente la position du mobile au bout de n déplacements.

1. (a) Déterminer $u_n = \mathbb{P}(S_n = 0)$, pour tout $n \in \mathbb{N}$.
   (b) On note $f(x)$ la somme de la série entière $\sum u_n x^n$. Montrer que :
   $$\forall x \in \left]-1,1\right[\quad f(x) = \frac{1}{\sqrt{1 - 4p(1-p)x^2}}.$$

2. Pour tout entier naturel non nul k, on note A_k l'événement « le mobile retourne pour la première fois à l'origine au bout n déplacements », c'est-à-dire
   $$A_k = \{S_k = 0\} \cap \left(\bigcap_{i=1}^{k-1} \{S_i \neq 0\} \right).$$
   On pose $v_k = \mathbb{P}(A_k)$, pour tout $k \geqslant 1$ et $v_0 = 0$
   (a) Montrer que, pour tout entier naturel n non nul, on a :
   $$\{S_n = 0\} = \sum_{k=1}^{n} \mathbb{P}\left(\{S_n = 0\} \cap A_k\right).$$
   (b) En déduire que, pour tout entier naturel non nul n, on a :
   $$u_n = \sum_{k=0}^{n} u_{n-k} v_k.$$

3. On note $g(x)$ la somme de la série entière $\sum v_n x^n$.
   (a) Montrer que le rayon de la série entière définissant $g(x)$ est supérieur ou égal à 1. Montrer que :
   $$\forall x \in \left]-1,1\right[\quad g(x) = \frac{f(x) - 1}{f(x)} = 1 - \sqrt{1 - 4p(1-p)x^2}.$$
   (b) Déterminer la probabilité de l'événement A : « il existe $n \in \mathbb{N}^*$ tel que $S_n = 0$ ».

On suppose dans le reste de l'exercice que $p = \frac{1}{2}\cdot$

4. (a) Montrer que l'on peut définir une variable aléatoire T égale au premier indice n non nul pour lequel l'événement $\{S_n = 0\}$ est réalisé.

   (b) Montrer que l'on a, pour tout $n \in \mathbb{N}^*$,
$$v_{2n} = \frac{2\binom{2n-2}{n-1}}{n4^n}\cdot$$

   (c) La variable T est-t-elle d'espérance finie ?

5. (a) Montrer que, pour tout $n \in \mathbb{N}^*$, on a
$$v_{2n} = u_{2n-2} - u_{2n}.$$

   (b) Démontrer que la probabilité que le mobile soit à l'origine à l'issue des $2n$ premiers déplacements est égal à la probabilité qu'il ne soit jamais à l'origine à l'issue d'aucun des $2n$ premiers déplacements. En déduire $\mathbb{P}(S_1 > 0, S_2 > 0, \ldots, S_{2n} > 0)$.

$\star$ **16.20 (Loi faible des grands nombres dans L_1)**

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles discrètes, deux à deux indépendantes, de même loi, possédant un espérance finie m.

On pose, pour tout $n \in \mathbb{N}^*$, $Y_n = \dfrac{X_1 + \cdots + X_n}{n}\cdot$

Dans les deux premières questions, on suppose $m = 0$.

1. Soit $\varepsilon > 0$.

   (a) Pour $c > 0$, on définit $g : \mathbb{R} \longmapsto \mathbb{R}$ par $g(x) = \begin{cases} x & \text{si } |x| \leqslant c \\ 0 & \text{sinon.} \end{cases}$

   Montrer que la variable aléatoire $g(X_1)$ est d'espérance finie et que l'on peut choisir c tel que $\mathbb{E}\left(|g(X_1) - X_1|\right) \leqslant \dfrac{\varepsilon}{2}\cdot$

   Dans la suite de la question, c est ainsi fixé.

   (b) On pose $a = \mathbb{E}(g(X_1))$. Montrer que :
$$\mathbb{E}\left(|g(X_1) - X_1 - a|\right) \leqslant \varepsilon.$$

   (c) On pose, pour tout $n \in \mathbb{N}^*$, $U_n = g(X_n) - a$ et $Y'_n = \dfrac{U_1 + \cdots + U_n}{n}\cdot$

   Justifier que les variables U_n admettent un moment d'ordre 2.

   Montrer que $\lim\limits_{n\to+\infty} \mathbb{V}(Y'_n) = 0$.

   En déduire que $\lim\limits_{n\to+\infty} \mathbb{E}\left(|Y'_n|\right) = 0.$

   (d) Montrer que $\lim\limits_{n\to+\infty} \mathbb{E}\left(|Y_n|\right) = 0$.

2. Montrer que, pour tout $\varepsilon > 0$, on a $\lim\limits_{n\to+\infty} \mathbb{P}\big(|Y_n| \geqslant \varepsilon\big) = 0.$

3. On ne suppose plus $m = 0$. Montrer que l'on a :
$$\forall \varepsilon > 0 \quad \lim_{n\to+\infty} \mathbb{P}\big(|Y_n - m| \geqslant \varepsilon\big) = 0.$$

⋆⋆ **16.21 (Chaînes de Markov)**
Soit $N \in \mathbb{N}^*$ et (X_n) une suite de variables aléatoires sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, à valeurs dans $[\![1, N]\!]$. On dit que (X_n) est une *chaîne de Markov homogène* s'il existe une matrice $P = \big(p_{i,j}\big)_{1 \leqslant i,j \leqslant N} \in \mathcal{M}_N(\mathbb{R})$ telle que, pour tout entier n et tous éléments x_0, x_1, …, x_{n+1} de $[\![1, N]\!]$, on ait :

$$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1} \,|\, X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n)&\\ = \mathbb{P}(X_{n+1} = x_{n+1} \,|\, X_n = x_n) = p_{x_n, x_{n+1}},&\end{aligned}$$

quand les probabilités conditionnelles sont définies. La matrice P est appelée matrice de transition de la chaîne. Dans la suite, on considère une telle chaîne de Markov.

1. Montrer que P est une matrice à coefficients positifs dont la somme des coefficients de chaque ligne est égale à 1 (on dit que P est une *matrice stochastique*).
   Montrer que 1 est valeur propre de P.
2. (a) Soit $x_0 \in [\![1, N]\!]$ tel que $\mathbb{P}(X = x_0) \neq 0$. Montrer que, pour $n \in \mathbb{N}^*$ et x_1, …, x_n dans $[\![1, N]\!]$, on a :

   $$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n \mid X_0 = x_0) = p_{x_0,x_1} p_{x_1,x_2} \cdots p_{x_{n-1},x_n}.$$

   En déduire que :

   $$\mathbb{P}(X_n = x_n \mid X_0 = x_0) = p^{(n)}_{x_0,x_n},$$

   où $p^{(n)}_{x_0,x_n}$ est le coefficient d'indice (x_0, x_n) de la matrice P^n.
   (b) Montrer que, pour $n \in \mathbb{N}$, $k \in \mathbb{N}^*$ et x_0, x_1, …, x_{n+k} dans $[\![1, N]\!]$, tels que $\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n) \neq 0$, on a :

   $$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+k} = x_{n+k} \mid X_0 = x_0, \ldots, X_n = x_n)&\\ = \mathbb{P}(X_1 = x_{n+1}, \ldots, X_k = x_{n+k} \mid X_0 = x_n).&\end{aligned}$$

3. On suppose désormais que tous les coefficients de P sont strictement positifs. On pose $\varepsilon = \min\limits_{(i,j) \in [\![1,N]\!]^2} p_{i,j} > 0$. On fixe un élément j, et l'on considère les suites

   $$u_n = \max_{i \in [\![1,N]\!]} p^{(n)}_{i,j} \quad \text{et} \quad v_n = \min_{i \in [\![1,N]\!]} p^{(n)}_{i,j}.$$

   (a) Montrer que pour tout entier naturel n :

   $$u_{n+1} \leqslant (1 - \varepsilon) u_n + \varepsilon v_n \quad \text{et} \quad v_{n+1} \geqslant (1 - \varepsilon) v_n + u_n.$$

   (b) Montrer que les suite (u_n) et (v_n) convergent vers la même limite.
   (c) En déduire qu'il existe une probabilité Q sur $[\![1, N]\!]$, c'est-à-dire une famille $Q = (q_j)_{j \in [\![1,N]\!]}$ d'entiers positifs, de somme 1, telle que

   $$\forall (i,j) \in [\![1, N]\!]^2 \qquad \lim_{n \to +\infty} \mathbb{P}(X_n = j \mid X_0 = i) = q_j.$$

   Montrer que Q est la seule probabilité invariante par P, c'est-à-dire vérifiant $QP = Q$.

16.22 Toutes les variables considérées dans cet exercice sont à valeurs dans $\mathbb{Z}$. Une variable aléatoire à valeurs dans $\mathbb{Z}$ est dite symétrique si :

$$\forall n \in \mathbb{Z} \quad \mathbb{P}(X = n) = \mathbb{P}(X = -n).$$

1. (a) Montrer que si X est symétrique, alors 0 est une médiane de X, c'est-à-dire :

$$\mathbb{P}(X > 0) \leqslant \frac{1}{2} \leqslant \mathbb{P}(X \geqslant 0).$$

   (b) Montrer que si X et Y sont deux variables aléatoires indépendantes de même loi, alors $X - Y$ est symétrique.

   (c) Montrer que si X et Y sont deux variables aléatoires symétriques indépendantes, alors $X + Y$ est symétrique.

2. On considère des variables aléatoires symétriques X_1, X_2, …, X_n, indépendantes. On se donne $x \geqslant 0$. Pour $k \in [\![1, n]\!]$, on pose $S_k = \sum\limits_{j=1}^{k} X_j$ et, de plus, on note Ω_k l'événement $\left\{\max\limits_{1 \leqslant j \leqslant k-1} S_j \leqslant x\right\} \cap \{S_k > x\}$ si $k \geqslant 2$ et $\Omega_1 = \{S_1 > x\}$.

   (a) Montrer que $X_1 + \cdots + X_n$ est symétrique.

   (b) Montrer que, pour $k \in [\![1, n]\!]$, on a

$$\{S_n - S_k \geqslant 0\} \cap \Omega_k \subset \{S_n > x\} \cap \Omega_k \quad \text{et}$$

$$\mathbb{P}\left(\{S_n - S_k \geqslant 0\} \cap \Omega_k\right) \geqslant \frac{1}{2}\mathbb{P}(\Omega_k).$$

   (c) Prouver que $\bigcup\limits_{k=1}^{n} \Omega_k = \left\{\max\limits_{1 \leqslant j \leqslant n} S_j > x\right\}$.

   (d) En déduire l'inégalité de Paul Lévy :

$$\mathbb{P}\left(\max_{1 \leqslant j \leqslant n} S_j > x\right) \leqslant 2\mathbb{P}(S_n > x).$$

16.23 (Inégalité de Kolmogorov)

Soit X_1, …, X_n des variables aléatoires réelles discrètes de l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, indépendantes, ayant un moment d'ordre 2, centrées, ainsi que $a \in \mathbb{R}_+^*$. On pose, pour tout $i \in [\![1, n]\!]$:

$$S_i = X_1 + \cdots + X_i, \quad B_i = \big\{|S_1| < a\big\} \cap \ldots \cap \big\{|S_{i-1}| < a\big\} \cap \big\{|S_i| \geqslant a\big\}.$$

1. Montrer que, pour $i \in [\![1, n]\!]$, les variables $S_i \mathbb{1}_{B_i}$ et $S_n - S_i$ sont indépendantes. En déduire que :

$$\mathbb{E}\left(S_n^2 \, \mathbb{1}_{B_i}\right) = \mathbb{E}\left(S_i^2 \, \mathbb{1}_{B_i}\right) + \mathbb{E}\left((S_n - S_i)^2 \, \mathbb{1}_{B_i}\right) \geqslant a^2 \mathbb{P}(B_i).$$

2. (a) On pose $C = \big\{ \sup(|S_1|, |S_2|, \ldots, |S_n|) \geqslant a \big\}$.

   Montrer que $\mathbb{P}(C) = \sum\limits_{i=1}^{n} \mathbb{P}(B_i)$.

   (b) En déduire l'inégalité de Kolmogorov :

$$\mathbb{P}\big(\sup(|S_1|, |S_2|, \ldots, |S_n|) \geqslant a\big) \leqslant \frac{\mathbb{V}(S_n)}{a^2}.$$

16.24 (Inégalité de Le Cam)

L'objet de l'exercice est d'étudier l'approximation de la loi binomiale par la loi de Poisson. Toutes les variables aléatoires considérées sont définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et sont à valeurs dans $\mathbb{N}$.

1. Soit X et Y deux telles variables aléatoires. Pour tout $k \in \mathbb{N}$, on pose $p_k = \mathbb{P}(X = k)$ et $q_k = \mathbb{P}(Y = k)$. On définit la distance entre X et Y par :

$$\mathrm{d}(X, Y) = \frac{1}{2} \sum_{k=0}^{+\infty} |p_k - q_k|.$$

(a) Montrer que pour toute partie A de $\mathbb{N}$, on a :

$$\big|\mathbb{P}\,(X \in A) - \mathbb{P}\,(Y \in A)\big| \leqslant \mathrm{d}(X, Y).$$

(b) Démontrer la formule :

$$\mathrm{d}\,(X, Y) = 1 - \sum_{k=0}^{+\infty} \min(p_k, q_k).$$

(c) En déduire :

$$\mathrm{d}(X, Y) \leqslant \mathbb{P}\,(X \neq Y)\,.$$

2. (a) Vérifier que pour tout $x \in [0,1]$, le réel $f\,(x) = 1 - (1 - x)\exp\,(x)$ appartient à $[0,1]$.
Soit $U_1, \ldots, U_n$, $Y_1, \ldots, Y_n$ des variables aléatoires mutuellement indépendantes. On suppose que, pour $1 \leqslant i \leqslant n$, U_i suit la loi de Bernoulli de paramètre $f\left(\frac{\lambda}{n}\right)$ et Y_i suit la loi de Poisson de paramètre $\frac{\lambda}{n}\cdot$ On pose $Y = \sum\limits_{i=1}^{n} Y_i$.
Enfin, pour $i \in [\![1, n]\!]$, on considère la variable de Bernoulli X_i telle que $X_i = 0$ si $U_i = Y_i = 0$ et $X_i = 1$ sinon.

(b) Déterminer pour tout $i \in [\![1, n]\!]$, la loi de X_i. En déduire la loi de $X = \sum\limits_{i=1}^{n} X_i$.
Quelle est la loi de Y ?

(c) Montrer que pour tout $i \in [\![1, n]\!]$:

$$\mathbb{P}\big(X_i \neq Y_i\big) \leqslant \frac{\lambda^2}{n^2}\cdot$$

(d) En déduire l'*inégalité de Le Cam* :

$$\mathrm{d}(X, Y) \leqslant \frac{\lambda^2}{n}\cdot$$

16.25 (Convergence presque sûre) Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle définies sur $(\Omega,\mathcal{A},\mathbb{P})$. On pose :

$$B=\{\omega\in\Omega \mid \lim_{n\to+\infty} X_n(\omega)=X(\omega)\}$$

$$\forall k\in\mathbb{N}^* \quad C_k=\bigcup_{n\in\mathbb{N}}\bigcap_{p\geqslant n}\left\{|X_p-X|\leqslant\frac{1}{k}\right\}.$$

On dit que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X si $\mathbb{P}(B)=1$.

1. Montrer que l'on a $\mathbb{P}(B)=\lim_{k\to+\infty}\mathbb{P}(C_k)$.
2. On suppose que :
$$\forall\varepsilon>0\quad \mathbb{P}\Big(\bigcap_{n\in\mathbb{N}}\bigcup_{p\geqslant n}\big\{|X_p-X|>\varepsilon\big\}\Big)=0.$$
Montrer que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X.
3. Montrer que si la série de terme général $\mathbb{P}\left(|X_n-X|>\varepsilon\right)$ converge pour tout $\varepsilon>0$, alors la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X.

$\star\star$ **16.26 (Fonction génératrice des moments)**

Soit X une variable aléatoire discrète, pas presque sûrement constante, sur l'espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$. On pose, pour $t\in\mathbb{R}$, $L_X(t)=\mathbb{E}\left(e^{tX}\right)$ (la fonction L_X est appelée *fonction génératrice des moments* de la variable aléatoire X). On suppose qu'il existe un intervalle $]\alpha,\beta[$ contenant 0 tel que $L_X(t)<+\infty$ pour tout $t\in\,]\alpha,\beta[$.

1. (a) Soit $a<b$ deux réels tels que $[a,b]\subset\,]\alpha,\beta[$. On considère $\delta>0$ tel que $[a-\delta,b+\delta]\subset\,]\alpha,\beta[$. Soit $k\in\mathbb{N}$. Montrer qu'il existe $C>0$ tel que :
$$\forall t\in[a,b]\quad\forall u\in\mathbb{R}\quad |u|^k e^{tu}\leqslant C\big(e^{(a-\delta)u}+e^{(b+\delta)u}\big).$$
En déduire que $X^k e^{tX}$ est d'espérance finie pour tout $t\in\,]\alpha,\beta[$.

   (b) Montrer que L_X est de classe $\mathcal{C}^\infty$ sur $]\alpha,\beta[$ et vérifie :
$$\forall t\in\,]\alpha,\beta[\quad\forall k\in\mathbb{N}\quad L_X^{(k)}(t)=\mathbb{E}(X^k e^{tX}).$$
En déduire, pour tout $k\in\mathbb{N}$, une expression du moment d'ordre k de X. On note m l'espérance de X.

2. (a) On pose, pour tout $t\in\,]\alpha,\beta[$, $\Psi_X(t)=\ln L_X(t)$.
Montrer que Ψ_X est strictement convexe.

   (b) On note $I=\Psi_X'\big(]\alpha,\beta[\big)$ et on pose $g(c)=\max\limits_{t\in]\alpha,\beta[}(ct-\Psi_X(t)$, pour tout $c\in I$.
Montrer que $m\in I$. Calculer $g(m)$; montrer que $g(c)>0$ si $c\neq m$.

   (c) Montrer que $g(c)=\begin{cases}\max\limits_{t\in]\alpha,0[}\big(ct-\Psi_X(t)\big) & \text{si } c<m\\ \max\limits_{t\in]0,\beta[}\big(ct-\Psi_X(t)\big) & \text{si } c>m.\end{cases}$

   (d) En déduire les *inégalités de Chernov* :
$$\mathbb{P}(X\leqslant c)\leqslant e^{-g(c)}\quad\text{si}\quad c<m$$
$$\mathbb{P}(X\geqslant c)\leqslant e^{-g(c)}\quad\text{si}\quad c>m.$$

3. On considère une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires indépendantes suivant la même loi que X. On pose, pour tout $n\in\mathbb{N}^*$, $S_n=\sum\limits_{k=1}^{n} X_k$.

   (a) Soit $c\in I$. Montrer que :
   $$\mathbb{P}\left(\frac{S_n}{n}\leqslant c\right)\leqslant e^{-ng(c)} \qquad \text{si} \qquad c<m$$
   $$\mathbb{P}\left(\frac{S_n}{n}\geqslant c\right)\leqslant e^{-ng(c)} \qquad \text{si} \qquad c>m.$$

   (b) Soit $\varepsilon>0$. Montrer que, pour ε assez petit, on a :
   $$\mathbb{P}\left(\left|\frac{S_n}{n}-m\right|\geqslant\varepsilon\right)\leqslant 2e^{-n\min(g(m+\varepsilon),g(m-\varepsilon))}.$$

   (c) En utilisant le résultat de l'exercice 16.25, montrer que la suite $\left(\dfrac{S_n}{n}\right)$ converge presque sûrement vers m.

16.27 (Théorème de Weierstrass)
Soit f une fonction continue de $[0,1]$ dans $\mathbb{R}$. Soit $x\in[0,1]$. On considère une suite $(X_n)_{n\geqslant 1}$ de variables de Bernoulli de paramètre x, indépendantes, sur le même espace probabilisé. Pour $n\geqslant 1$, on pose $Y_n=\dfrac{X_1+\cdots+X_n}{n}\cdot$

1. Soit $\varepsilon>0$. Par uniforme continuité de f sur $[0,1]$, il existe $\eta>0$ tel que :
   $$\forall (t,u)\in[0,1]^2 \quad |t-u|\leqslant\eta\Longrightarrow|f(t)-f(u)|\leqslant\varepsilon.$$

   (a) Montrer que :
   $$\forall (t,u)\in[0,1]^2 \quad |f(t)-f(u)|\leqslant\frac{2\|f\|_\infty(t-u)^2}{\eta^2}+\varepsilon.$$

   (b) En déduire que
   $$|\mathbb{E}\left(f(Y_n)\right)-f(x)|\leqslant\frac{2\|f\|_\infty\mathbb{V}(Y_n)}{\eta^2}+\varepsilon\leqslant\frac{2\|f\|_\infty}{n\eta^2}+\varepsilon.$$

2. On considère les polynômes de Bernstein définis par :
   $$\forall x\in\mathbb{R} \quad B_n(f)=\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^k.$$
   Montrer que la suite $(B_n(f))$ converge uniformément vers f sur $[0,1]$.

16.28 Soit n un entier supérieur ou égal à 2. On considère une file d'attente avec un guichet et n clients qui attendent. Chaque minute, un guichet se libère. Le guichetier choisit alors le client qu'il appelle selon le processus aléatoire suivant :

- avec probabilité $\dfrac{1}{2}$, il appelle le client en première position dans la file,
- sinon, il choisit de manière équiprobable parmi les $n-1$ autres clients.

Enfin, un nouveau client arrive dans la file et se place en dernière position (de telle sorte qu'il y a toujours exactement n clients qui attendent). Pour tout $k \in [\![1, n]\!]$, on note T_k le temps d'attente d'un client qui se trouve en position k dans la file.

1. Quelle est la loi de T_1 ? Donner son espérance, sa variance.
2. Montrer que, pour tout $k \in [\![1, n]\!]$, la variable T_k est d'espérance finie.
3. Écrire une relation entre $\mathbb{E}(T_k)$ et $\mathbb{E}(T_{k-1})$ pour tout $k \geqslant 2$. En déduire une expression de $\mathbb{E}(T_k)$ en fonction de k et n.
   On pourra considérer la suite $((n+k-2)\,\mathbb{E}(T_k))_{1 \leqslant k \leqslant n}$.
4. Comparer les caractéristiques de cette file d'attente et d'une file d'attente « classique » (premier arrivé, premier servi).

⋆ **16.29** *(Cachan-Rennes 2015)*

Soit X et Y deux variables aléatoires réelles d'espérance finie sur le même espace probabilisé. On considère les trois propriétés suivantes :

(i) X et Y sont presque sûrement constante ;

(ii) X et Y sont indépendantes ;

(iii) XY est d'espérance finie et $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.

1. Montrer que $(i) \Rightarrow (ii)$ et que $(ii) \Rightarrow (iii)$, mais qu'aucune réciproque n'est vraie.
2. Montrer que $(iii) \Rightarrow (i)$ est vrai s'il existe $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ strictement croissantes ainsi qu'une variable aléatoire réelle Z telle que $X = f(Z)$ et $Y = g(Z)$.
   On admettra qu'aucune généralité n'est perdue à supposer qu'il existe une seconde variable aléatoire Z' indépendante de Z et suivant la même loi que Z. On introduira la fonction :
   $$\theta : (a,b) \in \mathbb{R}^2 \mapsto (f(a) - f(b))(g(a) - g(b))$$
   et on s'intéressera à la variable $\theta(Z, Z')$.

⋆ **16.30** *(Polytechnique 2015)*

Soit X_1 et X_2 deux variables aléatoires sur le même espace probabilisé, à valeurs dans $\mathbb{Z}$, tel que $\mathbb{E}(|X_2|) < +\infty$.

1. Montrer qu'il existe une variable aléatoire Y_1 d'espérance finie, qui s'écrit $Y_1 = h(X_1)$, où h est une fonction de $\mathbb{Z}$ dans $\mathbb{R}$, et telle que, pour tout application bornée f de $\mathbb{Z}$ dans $\mathbb{R}$, on ait :
   $$\mathbb{E}(X_2\, f(X_1)) = \mathbb{E}(Y_1 f(X_1)).$$
2. Montrer que si une autre variable aléatoire a les mêmes propriétés que Y_1, elle est égale à Y_1 presque sûrement.
3. La variable aléatoire Y_1 étant définie et unique, on la note $\mathbb{E}(X_2 \mid X_1)$: c'est l'espérance conditionnelle de X_2 par rapport à X_1. Soit g une application bornée de $\mathbb{Z}$ vers $\mathbb{R}$. Montrer que :
   $$\mathbb{E}(X_2\, g(X_1) \mid X_1) = g(X_1)\,\mathbb{E}(X_2 \mid X_1).$$

Solution des exercices

16.1 Dans les deux cas, Y est une variable aléatoire à valeurs dans $\mathbb{Z}$ et, pour tout $k \in \mathbb{Z}$, on a :

$$\mathbb{P}(Y = k) = \begin{cases} \mathbb{P}(X = 1 - 2k) & \text{si} \quad k < 0 \\ \mathbb{P}(X = 0) + \mathbb{P}(X = 1) & \text{si} \quad k = 0 \\ \mathbb{P}(X = 2k) & \text{si} \quad k > 0. \end{cases}$$

- Supposons que $X \sim \mathcal{G}(p)$. On obtient

$$\mathbb{P}(Y = k) = \begin{cases} p(1-p)^{-2k} & \text{si} \quad k < 0 \\ p & \text{si} \quad k = 0 \\ p(1-p)^{2k-1} & \text{si} \quad k > 0. \end{cases}$$

  On en déduit :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k \in \mathbb{Z}_-^*} kp(1-p)^{-2k} + \sum_{k \in \mathbb{N}^*} kp(1-p)^{2k-1} \\ &= -p(1-p)^2 \sum_{k \in \mathbb{N}^*} k(1-p)^{2(k-1)} + p(1-p) \sum_{k \in \mathbb{N}^*} k(1-p)^{2(k-1)} \\ &= \frac{-p(1-p)^2 + p(1-p)}{(1-p^2)^2} = \frac{p^2}{(1-p)(1+p)^2}. \end{aligned}$$

- Supposons que $X \sim \mathcal{P}(\lambda)$. On obtient :

$$\mathbb{P}(Y = k) = \begin{cases} \dfrac{\lambda^{1-2k} e^{-\lambda}}{(1-2k)!} & \text{si} \quad k < 0 \\ e^{-\lambda}(1+\lambda) & \text{si} \quad k = 0 \\ \dfrac{\lambda^{2k} e^{-\lambda}}{(2k)!} & \text{si} \quad k > 0. \end{cases}$$

  On en déduit :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k \in \mathbb{Z}_-^*} k \frac{\lambda^{1-2k} e^{-\lambda}}{(1-2k)!} + \sum_{k \in \mathbb{N}^*} \frac{\lambda^{2k} e^{-\lambda}}{(2k)!} \\ &= e^{-\lambda} \left(- \sum_{k \in \mathbb{N}} k \frac{\lambda^{1+2k}}{(1+2k)!} + \sum_{k \in \mathbb{N}^*} k \frac{\lambda^{2k}}{(2k)!} \right) \\ &= e^{-\lambda} \left(-\frac{1}{2} \sum_{k \in \mathbb{N}} \lambda^{1+2k} \left(\frac{1}{(2k)!} - \frac{1}{(1+2k)!} \right) + \frac{1}{2} \sum_{k \in \mathbb{N}^*} \frac{\lambda^{2k}}{(2k-1)!} \right) \\ &= \frac{e^{-\lambda}}{2} \left(-\lambda \operatorname{ch}(\lambda) + \operatorname{sh}(\lambda) + \lambda \operatorname{sh}(\lambda) \right) \\ &= \frac{1}{4} \left(1 - (2\lambda + 1) e^{-2\lambda} \right). \end{aligned}$$

16.2 On a, pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = \left(\dfrac{1}{2}\right)^k$.

1. On a $\mathbb{P}(X = Y) = \sum\limits_{k=1}^{+\infty} \mathbb{P}\big(\{X = k\} \cap \{Y = k\}\big)$ et par indépendance de X et Y, on obtient :

$$\mathbb{P}(X = Y) = \sum_{k=1}^{+\infty} (\mathbb{P}(X = k))^2 = \sum_{k=1}^{+\infty} \left(\frac{1}{4}\right)^k = \frac{1}{3}\cdot$$

2. On obtient, de même :

$$P(X > Y) = \sum_{k=1}^{+\infty} \mathbb{P}(\{X > k\} \cap \{Y = k\}) = \sum_{k=1}^{+\infty} \mathbb{P}(X > k)\,\mathbb{P}(Y = k)$$
$$= \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^k = \frac{1}{3}.$$

3. Pour $k \geqslant 1$ fixé, on a :

$$\mathbb{P}(X \geqslant kY) = \sum_{j=1}^{+\infty} \mathbb{P}(\{Y = j\} \cap \{X \geqslant kj\}) = \sum_{j=1}^{+\infty} \mathbb{P}(Y = j)\mathbb{P}(X \geqslant kj)$$
$$= \sum_{j=1}^{+\infty} \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^{kj-1} = 2\sum_{j=1}^{+\infty} \left(\frac{1}{2}\right)^{j(k+1)}$$
$$= 2\frac{\left(\frac{1}{2}\right)^{k+1}}{1 - \left(\frac{1}{2}\right)^{k+1}} = \frac{2}{2^{k+1} - 1}\cdot$$

16.3 On pose $q = 1 - p$. La variable $|X - Y|$ est à valeurs dans $\mathbb{N}$. Pour tout $n \in \mathbb{N}$,

$$\{|X - Y| = n\} = \{X - Y = n\} \cup \{Y - X = n\}$$
$$= \bigcup_{k \in \mathbb{N}^*} \big(\{Y = k\} \cap \{X = n + k\} \cup \{X = k\} \cap \{Y = n + k\}\big).$$

Pour $n \in \mathbb{N}^*$, tous ces événements sont incompatibles. On en déduit :

$$\mathbb{P}(|X - Y| = n) = 2\sum_{k=1}^{+\infty} \big(\mathbb{P}\,\{Y = k\} \cap \{X = k + n\}) + \mathbb{P}\,(\{X = k\} \cap \{Y = n + k\})\,\big)$$
$$= 2\sum_{k=1}^{+\infty} \big(\mathbb{P}(Y = k)\mathbb{P}(X = k + n)$$
$$= 2\sum_{k=1}^{+\infty} p^2 q^{2k+n-2} = \frac{2p^2 q^n}{1 - q^2} = \frac{2pq^n}{1 + q}\cdot$$

De plus, on a $\mathbb{P}(|X - Y| = 0) = \mathbb{P}(X = Y) = \sum\limits_{k=1}^{+\infty} \mathbb{P}(\{X = k\} \cap \{Y = k\})$

$$= \sum_{k=1}^{+\infty} p^2 q^{2k-2} = \frac{p^2}{1 - q^2} = \frac{p}{1 + q}\cdot$$

16.4 On pose $q = 1 - p$.

1. On a $\mathbb{P}(Y = 0) = \mathbb{P}(X \leqslant n) = 1 - \mathbb{P}(X > n) = 1 - q^n$.
   Pour $k \geqslant 1$, on a $\mathbb{P}(Y = k) = \mathbb{P}(X = n + k) = pq^{n+k-1}$.
2. On en déduit, sous réserve d'absolue convergence :
$$G_Y(t) = (1 - q^n) + \sum_{k=1}^{+\infty} pq^{n+k-1}t^k.$$
   Ainsi, $G_Y(t)$ est défini pour $|tq| < 1$ et $G_Y(t) = 1 - q^n + \dfrac{q^n pt}{1 - qt}\cdot$
3. Comme $1 \in \left]-1/q, 1/q\right[$, la fonction G_X est dérivable en 1, donc Y est d'espérance finie :
$$\mathbb{E}(Y) = G_Y'(1) = \frac{q^n p}{1-q} + \frac{q^n pq}{(1-q)^2} = \frac{q^n}{p}\cdot$$

16.5 1. (a) On a $D(\Omega) = \mathbb{Z}$ et $I(\Omega) = \mathbb{N}$.

- Si $k \geqslant 0$, alors on a :
$$\{D = k\} \cap \{I = \ell\} = \{X - Y = k\} \cap \{Y = \ell\} = \{X = k + \ell\} \cap \{Y = \ell\}$$
  et donc, par indépendance de X et Y :
$$\begin{aligned}\mathbb{P}\left(\{D = k\} \cap \{I = \ell\}\right) &= \mathbb{P}\left(\{X = k + \ell\} \cap \{Y = \ell\}\right)\\ &= pq^{k+\ell}pq^{\ell} = p^2 q^{k+2\ell}.\end{aligned}$$
- Si $k < 0$, alors on a :
$$\{D = k\} \cap \{I = \ell\} = \{X - Y = k\} \cap \{X = \ell\} = \{X = \ell\} \cap \{Y = -k + \ell\}$$
  et donc, par indépendance de X et Y :
$$\begin{aligned}\mathbb{P}\left(\{D = k\} \cap \{I = \ell\}\right) &= \mathbb{P}\left(\{X = \ell\} \cap \{Y = -k + \ell\}\right)\\ &= pq^{\ell}pq^{-k+\ell} = p^2 q^{-k+2\ell}.\end{aligned}$$

Dans tous les cas, on trouve $\mathbb{P}\left(\{D = k\} \cap \{I = \ell\}\right) = p^2 q^{|k|+2\ell}$.

(b)
- Pour $k \in \mathbb{Z}$, on a :
$$\begin{aligned}\mathbb{P}(D = k) &= \sum_{\ell=0}^{+\infty} \mathbb{P}\left(\{D = k\} \cap \{I = \ell\}\right) = \sum_{\ell=0}^{+\infty} p^2 q^{|k|+2\ell}\\ &= p^2 q^{|k|} \sum_{\ell=0}^{+\infty} \left(q^2\right)^{\ell} = \frac{p^2 q^{|k|}}{1 - q^2} = \frac{pq^{|k|}}{1+q}\cdot\end{aligned}$$
- Pour $\ell \in \mathbb{N}$, on a :
$$\begin{aligned}\mathbb{P}(I = \ell) &= \sum_{k=-\infty}^{+\infty} \mathbb{P}\left(\{D = k\} \cap \{I = \ell\}\right) = \sum_{k=-\infty}^{+\infty} p^2 q^{|k|+2\ell}\\ &= p^2 q^{2\ell} \sum_{k=-\infty}^{+\infty} q^{|k|} = p^2 q^{2\ell}\left(2\sum_{k=0}^{+\infty} q^k - 1\right)\\ &= p^2 q^{2\ell}\left(\frac{2}{1-q} - 1\right) = pq^{2\ell}(1+q).\end{aligned}$$

On a bien :

$$\forall (k,\ell) \in \mathbb{Z} \times \mathbb{N} \quad \mathbb{P}\left(\{D = k\} \cap \{I = \ell\}\right) = \mathbb{P}(D = k)\, \mathbb{P}(I = \ell),$$

ce qui prouve que D et I sont indépendantes.

2. Comme précédemment, on a :

$$\forall k \in \mathbb{N} \quad \{D = k\} \cap \{I = \ell\} = \{X = k + \ell\} \cap \{Y = \ell\}.$$

Par indépendance de X et Y d'une part, de D et I d'autre part, on en déduit :

$$\mathbb{P}(D = k)\mathbb{P}(I = \ell) = \mathbb{P}(X = k + \ell)\mathbb{P}(Y = \ell) = \mathbb{P}(X = k + \ell)\mathbb{P}(X = \ell) \neq 0,$$

par hypothèse. On a en particulier, pour tout $k \in \mathbb{N}$:

$$\mathbb{P}(D = k)\mathbb{P}(I = 0) = \mathbb{P}(X = k)\mathbb{P}(X = 0)$$

$$\mathbb{P}(D = k)\mathbb{P}(I = 1) = \mathbb{P}(X = k + 1)\mathbb{P}(X = 1).$$

En divisant les égalités, on obtient :

$$\frac{\mathbb{P}(X = k + 1)}{\mathbb{P}(X = k)} = \frac{\mathbb{P}(I = 1)\mathbb{P}(X = 0)}{\mathbb{P}(I = 0)\mathbb{P}(X = 1)}.$$

Ce rapport est indépendant de k et strictement positif. On le note q.

La suite $(\mathbb{P}(X = k))$ est géométrique de raison q.

Pour tout $k \in \mathbb{N}$, on a $\mathbb{P}(X = k) = \mathbb{P}(X = 0)q^k$. La série $\sum \mathbb{P}(X = k)$ converge et a pour somme 1 donc $q < 1$ et $\mathbb{P}(X = 0) = 1 - q$. En posant $p = 1 - q$, on a le résultat voulu.

16.6 1. Pour tout $n \in \mathbb{N}$, la loi de S sachant $\{N = n\}$ est la loi binomiale de paramètre (n,p). D'autre part, N suit une loi de Poisson de paramètre λ. On montre, comme dans l'exercice 7 de la page 902, que S suit la loi de Poisson de paramètre λp. De la même façon, la loi de E sachant $N = n$ est la loi binomiale de paramètre $(n, 1 - p)$ donc E suit la loi de Poisson de paramètre $\lambda(1 - p)$.

On remarque que $S + E = N$. On a donc, pour tout $(m,n) \in \mathbb{N}^2$:

$$\begin{aligned}
\mathbb{P}(\{S = m\} \cap \{E = n\}) &= \mathbb{P}(\{S = m\} \cap \{N = m + n\}) \\
&= \mathbb{P}(N = m + n)\, \mathbb{P}(X = m \mid N = m + n) \\
&= \frac{e^{-\lambda}\lambda^{m+n}}{(m+n)!} \binom{m+n}{m} p^m (1-p)^n \\
&= \frac{e^{-\lambda p}(\lambda p)^m}{m!} \times \frac{e^{-\lambda(1-p)}(\lambda(1-p))^n}{n!} \\
&= \mathbb{P}(S = m)\, \mathbb{P}(E = n).
\end{aligned}$$

Le variables aléatoires S et E sont indépendantes.

2. On suppose que S et E sont indépendantes. On a donc pour tout $(m,n) \in \mathbb{N}^2$,

$$\begin{aligned}
\mathbb{P}(S = m)\, \mathbb{P}(E = n) = \mathbb{P}(\{S = m\} \cap \{E = n\}) &= \mathbb{P}(\{S = m\} \cap \{N = m + n\}) \\
&= \mathbb{P}(N = m + n)\, \mathbb{P}(S = m \mid N = m + n) \\
&= \mathbb{P}(N = m + n) \binom{m+n}{m} p^m (1-p)^n.
\end{aligned}$$

On en déduit :

$$\mathbb{P}(N=m+n)(m+n)! = \frac{m!\,\mathbb{P}(S=m)}{p^m} \times \frac{n!\,\mathbb{P}(E=n)}{(1-p)^n} = u_m\, v_n,$$

en posant $u_m = \dfrac{m!\,\mathbb{P}(S=m)}{p^m}$ et $v_n = \dfrac{n!\,\mathbb{P}(E=n)}{(1-p)^n}$.

Pour $(m,n) \in \mathbb{N}^* \times \mathbb{N}$, on a :

$$u_{m-1}v_{n+1} = u_m v_n = \mathbb{P}(N=m+n) \neq 0$$

et donc $\dfrac{v_{n+1}}{v_n} = \dfrac{u_m}{u_{m-1}}\cdot$ Le rapport $\dfrac{v_{n+1}}{v_n}$ est égal à $\dfrac{u_m}{u_{m-1}}$ donc est indépendant de n : la suite (v_n) est donc géométrique. On note $\lambda > 0$ sa raison.

On en déduit, pour tout $n \in \mathbb{N}$:

$$\mathbb{P}(N=n) = \frac{u_0 v_n}{n!} = \frac{u_0 v_0 \lambda^n}{n!}\cdot$$

L'égalité $\displaystyle\sum_{n=0}^{+\infty} \mathbb{P}(N=n) = 1$ impose $u_0 v_0 = e^{-\lambda}$: N suit donc la loi de Poisson de paramètre λ.

16.7 1. La variable N suit la loi géométrique de paramètre p. Pour $n \in \mathbb{N}^*$, la loi conditionnelle de X sachant $\{N=n\}$ est la loi binomiale $\mathcal{B}(n,p)$. On a donc, pour tout $(n,k) \in \mathbb{N}^* \times \mathbb{N}$:

$$\begin{aligned}\mathbb{P}\big((N,X)=(n,k)\big) &= \mathbb{P}(N=n)\,\mathbb{P}(X=k \mid N=n)\\ &= \begin{cases} \binom{n}{k} p^{k+1}(1-p)^{2n-k-1} & \text{si } k \leqslant n\\ 0 & \text{si } k > n.\end{cases}\end{aligned}$$

Pour $k \in \mathbb{N}$, on en déduit $\mathbb{P}(X=k)$, en sommant sur n (avec $n \geqslant k$ et $n \geqslant 1$).

- Pour tout $k \in \mathbb{N}^*$, on a :

$$\begin{aligned}\mathbb{P}(X=k) &= \sum_{n=k}^{+\infty} \binom{n}{k} p^{k+1}(1-p)^{2n-k-1}\\ &= \frac{p^{k+1}(1-p)^{k-1}}{k!} \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!}(1-p)^{2(n-k)}.\end{aligned}$$

  On a, pour tout $x \in]-1,1[$:

$$\sum_{n=k}^{+\infty} \frac{n!}{(n-k)!} x^{n-k} = \frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\frac{1}{1-x}\right) = \frac{k!}{(1-x)^{k+1}}\cdot$$

  On en déduit :

$$\mathbb{P}(X=k) = \frac{p^{k+1}(1-p)^{k-1}}{k!} \frac{k!}{(1-(1-p)^2)^{k+1}} = \frac{(1-p)^{k-1}}{(2-p)^{k+1}}\cdot$$

- On a :

$$\mathbb{P}(X=0) = \sum_{n=1}^{+\infty} p(1-p)^{2n-1} = \frac{p(1-p)}{1-(1-p)^2} = \frac{1-p}{2-p}\cdot$$

2. Soit $p' \in]0,1[$. On considère deux variables aléatoires X et Y, indépendantes, Y suivant une loi de Bernoulli de paramètre p' et Z suivant une loi géométrique de paramètre p'. La variable aléatoire YZ est à valeurs dans $\mathbb{N}$. On a :
   - $\mathbb{P}(YZ = 0) = \mathbb{P}(Y = 0) = 1 - p'$;
   - pour $k \in \mathbb{N}^*$, $\mathbb{P}(YZ = k) = \mathbb{P}(Y = 1, Z = k) = {p'}^2(1-p')^{k-1}$.

   On voit que X et YZ ont même loi si, et seulement si, $p' = \dfrac{1}{2-p}\cdot$
3. On a, puisque Y et Z sont indépendantes :
$$\mathbb{E}(X) = \mathbb{E}(YZ) = \mathbb{E}(Y)\mathbb{E}(Z) = p'\frac{1}{p'} = 1.$$
   On a, de même, par indépendance de Y^2 et Z^2 :
$$\mathbb{E}(X^2) = \mathbb{E}(Y^2Z^2) = \mathbb{E}(Y^2)\mathbb{E}(Z^2) = \mathbb{E}(Y)\left(\mathbb{V}(Z) + (\mathbb{E}(Z))^2\right)$$
$$= p'\left(\frac{1-p'}{{p'}^2} + \frac{1}{{p'}^2}\right) = \frac{2-p'}{p'} = 3 - 2p.$$
   On en déduit $\mathbb{V}(X) = 2(1-p)$.

16.8 1. Les réels $\dfrac{n^{-s}}{\zeta(s)}$, pour $n \in \mathbb{N}^*$, sont positifs et $\displaystyle\sum_{n=1}^{+\infty} \frac{n^{-s}}{\zeta(s)} = 1$ par définition, donc on définit bien la loi d'une variable aléatoire.

2. Pour tout $n \in \mathbb{N}^*$, $A_n = \displaystyle\bigcup_{j \in \mathbb{N}^*} \{X = nj\}$ est un événement et :
$$\mathbb{P}(A_n) = \sum_{j=1}^{+\infty} \mathbb{P}(X = jn) = \frac{1}{\zeta(s)} \sum_{j=1}^{+\infty} (jn)^{-s} = \frac{n^{-s}}{\zeta(s)} \sum_{j=1}^{+\infty} j^{-s} = n^{-s}.$$
   Soit p_1, p_2, $\ldots$, p_k des nombres premiers distincts. Ces nombres sont premiers entre eux donc, d'après le théorème de Gauss :
$$A_{p_1} \cap A_{p_2} \cap \ldots \cap A_{p_k} = A_{p_1 p_2 \ldots p_k}.$$
   On en déduit :
$$\mathbb{P}\left(A_{p_1} \cap A_{p_2} \cap \ldots \cap A_{p_k}\right) = \mathbb{P}\left(A_{p_1 p_2 \ldots p_k}\right) = (p_1 p_2 \ldots p_k)^{-s}$$
$$= \prod_{i=1}^{k} p_i^{-s} = \prod_{i=1}^{k} \mathbb{P}(A_{p_i}).$$
   Les événements de la famille $(A_p)_{p \in \mathcal{P}}$ sont donc indépendants.

   On en déduit que $\left(\overline{A_p}\right)_{p \in \mathcal{P}}$ est aussi une famille d'événements indépendants. Notons $(p_n)_{n \geqslant 1}$ la suite des entiers premiers rangés par ordre croissant. On a :
$$\prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p^s}\right) = \prod_{n \in \mathbb{N}^*} \left(1 - \frac{1}{p_n^s}\right) = \lim_{N \to +\infty} \prod_{n=1}^{N} \left(1 - \frac{1}{p_n^s}\right)$$
$$= \lim_{N \to +\infty} \prod_{n=1}^{N} \mathbb{P}\left(\overline{A_{p_n}}\right) = \lim_{N \to +\infty} \mathbb{P}\left(\bigcap_{n=1}^{N} \overline{A_{p_n}}\right),$$

par indépendance des événements. Par ailleurs, on a, par continuité décroissante :

$$\lim_{N\to+\infty} \mathbb{P}\left(\bigcap_{n=1}^{N} \overline{A_{p_n}}\right) = \mathbb{P}\left(\bigcap_{n=1}^{+\infty} \overline{A_{p_n}}\right).$$

Mais $\bigcap\limits_{n=1}^{+\infty} \overline{A_{p_n}} = \{1\}$, car 1 est le seul entier qui n'ait pas de diviseur premier. Comme $\mathbb{P}(\{1\}) = \dfrac{1}{\zeta(s)}$, on en déduit :

$$\prod_{p\in\mathcal{P}} \left(1 - \frac{1}{p^s}\right) = \frac{1}{\zeta(s)}.$$

3. Notons E l'événement « aucun carré différent de 1 ne divise X ». L'événement E est réalisé si, et seulement si, le carré d'aucun nombre premier ne divise X. On a donc $E = \bigcap\limits_{p\in\mathcal{P}} \overline{A_{p^2}}$ et toujours par indépendance des événements $\overline{A_p}$:

$$\mathbb{P}(E) = \mathbb{P}\left(\bigcap_{p\in\mathcal{P}} \overline{A_{p^2}}\right) = \lim_{N\to+\infty} \mathbb{P}\left(\bigcap_{n=1}^{N} \overline{A_{p_n^2}}\right) = \lim_{N\to+\infty} \prod_{n=1}^{N} \mathbb{P}(\overline{A_{p_n^2}})$$
$$= \lim_{N\to+\infty} \prod_{n=1}^{N} \left(1 - \frac{1}{p_n^{2s}}\right) = \prod_{p\in\mathcal{P}} \left(1 - \frac{1}{p^{2s}}\right) = \frac{1}{\zeta(2s)},$$

d'après la question précédente.

16.9 1. On a, pour tout $n \in \mathbb{N}^*$, $x_n = \dfrac{\mathbb{P}(X = n)}{\mathbb{P}(X \geqslant n)}$ et donc :

$$x_n\mathbb{P}(X \geqslant n) = \mathbb{P}(X = n) = \mathbb{P}(X \geqslant n) - \mathbb{P}(X \geqslant n+1),$$

c'est-à-dire :

$$\mathbb{P}(X \geqslant n+1) = (1 - x_n)\mathbb{P}(X \geqslant n).$$

Comme $\mathbb{P}(X \geqslant 1) = 1$, on en déduit, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(X \geqslant n) = \prod_{k=1}^{n-1}(1 - x_k) \qquad \text{et} \qquad p_n = x_n \prod_{k=1}^{n-1}(1 - x_k).$$

2. (a) On a $x_n \in [0,1]$, puisque x_n est la probabilité d'un événement. Si $x_n = 1$, alors $\mathbb{P}(X \geqslant n+1) = (1-x_n)\mathbb{P}(X \geqslant n) = 0$, ce qui est contraire à l'hypothèse. On a donc $x_n \in [0,1[$.
On a pour tout $n \geqslant 1$:

$$\ln \mathbb{P}(X \geqslant n+1) = \sum_{k=1}^{n} \ln(1 - x_k).$$

De $\lim\limits_{n\to+\infty} \mathbb{P}(X \geqslant n+1) = 1 - \lim\limits_{n\to+\infty} \mathbb{P}(X \leqslant n) = 0$ (voir l'exercice 1 de la page 895), on déduit la divergence de la série de terme général $\ln(1 - x_n)$. Celle-ci implique la divergence de la série de terme général x_n, car sinon (x_n) tendrait vers 0 et on aurait $\ln(1 - x_n) \sim -x_n$ et la convergence de la série de terme général $\ln(1 - x_n)$ par théorème de comparaison.

(b) On pose, pour tout $n \in \mathbb{N}^*$, $p_n = x_n \prod\limits_{k=1}^{n-1} (1-x_k)$. Alors $p_n \in [0,1]$. On a pour tout $n \geqslant 1$:

$$p_n = \prod_{k=1}^{n-1}(1-x_k) - \prod_{k=1}^{n}(1-x_k).$$

On en déduit :

$$\sum_{k=1}^{n} p_k = 1 - \prod_{k=1}^{n}(1-x_k).$$

La série de terme général x_k diverge donc $\lim\limits_{n\to+\infty} \prod\limits_{k=1}^{n}(1-x_k) = 0$ car

$$\ln \prod_{k=1}^{n}(1-x_k) = \sum_{k=1}^{n} \ln(1-x_k) \leqslant -\sum_{k=1}^{n} x_k \to -\infty.$$

On a donc $\sum\limits_{k=1}^{+\infty} p_k = 1$, donc les p_k définissent la loi d'une variable aléatoire X à valeurs dans $\mathbb{N}^*$.

De plus, on a $x_k \neq 1$ pour tout $k \geqslant 1$, donc $\mathbb{P}(X < n) = \sum\limits_{k=1}^{n-1} p_k < 1$ pour tout $n \in \mathbb{N}^*$. On en déduit $\mathbb{P}(X \geqslant n) > 0$ pour tout $n \in \mathbb{N}^*$.

Enfin on a :

$$\begin{aligned}\mathbb{P}(X = n \mid X \geqslant n) &= \frac{\mathbb{P}(X=n)}{\mathbb{P}(X \geqslant n)} = \frac{p_n}{\mathbb{P}(X \geqslant n)} \\ &= \frac{p_n}{1-\mathbb{P}(X \leqslant n-1)} = \frac{x_n \prod\limits_{k=1}^{n-1}(1-x_k)}{\prod\limits_{k=1}^{n-1}(1-x_k)} = x_n,\end{aligned}$$

donc (x_n) est le taux de panne de la variable aléatoire X.

3. • Si X suit la loi géométrique de paramètre p, on a, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(X=n) = (1-p)^{n-1}p, \quad \mathbb{P}(X \geqslant n) = (1-p)^{n-1} \quad \text{et} \quad x_n = p.$$

• Si le taux de panne est constant, égal à p, alors $p \in [0,1[$ et on a, pour tout $n \geqslant 1$, $p_n = x_n \prod\limits_{k=1}^{n-1}(1-x_k) = p(1-p)^{n-1}$.

On voit que p ne peut pas être nul, car $\sum\limits_{n \in \mathbb{N}^*} p_n = 1$, donc $p \in]0,1[$ et X suit la loi géométrique de paramètre p.

16.10 1. On observe qu'en 1, il y toujours un maximum et un minimum provisoire et donc que X_n et Y_n sont à valeurs dans $[\![1,n]\!]$.

L'application $f : \Omega \to \Omega$ qui à la permutation σ associe la permutation $\sigma' : k \mapsto n+1-\sigma(k)$ est clairement bijective. Pour $i \in [\![1,n]\!]$, $\sigma(i)$ est un maximum provisoire de σ si, et seulement si, $\sigma'(i)$ est un minimum provisoire de σ. En effet $\sigma(i)=\max(\sigma(1),\ldots,\sigma(i))$ équivaut à :

$$\begin{aligned}n+1-\sigma'(i)&=\max(n+1-\sigma'(1),\ldots,n+1-\sigma'(i))\\&=n+1-\min(\sigma'(1),\ldots,\sigma'(i))\end{aligned}$$

et donc à $\sigma'(i)=\min(\sigma'(1),\ldots,\sigma'(i))$.

On en déduit que, pour tout $k\in[\![1,n]\!]$,

$$\sigma\in\{X_n=k\}\Longleftrightarrow\sigma'\in\{Y_n=k\}.$$

Comme f est bijective, on a $\operatorname{card}(\{X_n=k\})=\operatorname{card}(\{Y_n=k\})$, et donc $\mathbb{P}(X_n=k)=\mathbb{P}(Y_n=k)$, car Ω est muni de la probabilité uniforme.

2. (a) • La permutation σ est dans $\{X_3=1\}$ si, et seulement si, $\sigma(1)=3$. On en déduit $\mathbb{P}(X_3=1)=\dfrac{1}{3}\cdot$

• On a $\{X_3=3\}=\left\{\operatorname{Id}_{[\![1,3]\!]}\right\}$. On en déduit $\mathbb{P}(X_3=3)=\dfrac{1}{6}\cdot$

• Enfin, $\mathbb{P}(X_3=2)=1-\dfrac{1}{3}-\dfrac{1}{6}=\dfrac{1}{2}\cdot$

On obtient $\mathbb{E}(X_3)=\dfrac{11}{6}$ et $\mathbb{V}(X_3)=\dfrac{17}{36}\cdot$

(b) Pour $(k,\ell)\in[\![1,3]\!]^2$, on a $\mathbb{P}(X_3=k,Y_3=\ell)=0$ si $k+\ell\geqslant 4$. En effet, sauf pour $i=1$, $\sigma(i)$ ne peut pas être à la fois un maximum provisoire et un minimum provisoire. On obtient :

X \ Y	1	2	3
1	0	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{1}{3}$	0
3	$\frac{1}{6}$	0	0

On trouve $\mathbb{E}(X_3Y_3)=3$ et $\operatorname{Cov}(X_3,Y_3)=-\dfrac{13}{36}\cdot$

3. (a) Montrons que les variables $Z_1,\ldots,Z_n$ sont indépendantes.
Comme ce sont des variables de Bernoulli, il suffit de montrer que les événements $\{Z_1=1\},\ldots,\{Z_n=1\}$ sont indépendants. Soit $k\in[\![1,n]\!]$ et $i_1,\ldots,i_k$ tels que $1\leqslant i_1<\ldots<i_k\leqslant n$.
Pour déterminer un élément $\sigma\in\Omega$, pour lequel $\sigma(i_1)$, $\sigma(i_2)$, …, $\sigma(i_k)$ sont des maximums provisoires :

• on choisit l'ensemble $A_k=\{\sigma(i)\,;\ 1\leqslant i\leqslant i_k\}$: il y a $\binom{n}{i_k}$ choix possibles ; alors, le plus grand élément de cet ensemble est $\sigma(i_k)$;

• on choisit l'ensemble $A_{k-1}=\{\sigma(i)\,;\ 1\leqslant i\leqslant i_{k-1}\}$ sous-ensemble de $A_k\setminus\{\sigma(i_k)\}$: il y a $\binom{i_k-1}{i_{k-1}}$ choix possibles ; le plus grand élément de cet ensemble est $\sigma(i_{k-1})$; on attribue les images $\sigma(i)$ pour $i_{k-1}<i<i_k$ qui sont les $i_k-i_{k-1}-1$ éléments restants ; il y a $(i_k-i_{k-1}-1)!$ choix possibles ;

de plus on a $\binom{i_k-1}{i_{k-1}}(i_k - i_{k-1} - 1)! = \dfrac{(i_k-1)!}{i_{k-1}!}$;

- on continue ainsi jusqu'au choix $A_1 = \{\sigma(i)\,;\ 1 \leqslant i \leqslant i_1\}$ sous-ensemble de $A_2 \setminus \{\sigma(i_2)\}$: il y a $\binom{i_2-1}{i_1}$ choix possibles ; le plus grand élément de cet ensemble est $\sigma(i_1)$; on fixe ensuite les images des $\sigma(i)$ pour $i_1 \leqslant i < i_2$; il y a $(i_2 - i_1 - 1)!$ choix possibles ; de plus $\binom{i_2-1}{i_1}(i_2 - i_1 - 1)! = \dfrac{(i_2-1)!}{i_1!}$;
- il reste encore à fixer $\sigma(i)$ pour $i < i_1$ et pour $i > i_k$, ce qui donne $(i_1-1)!(n-i_k)!$ choix.

On obtient :

$$\mathbb{P}\left(\bigcap_{j=1}^{k}\{Z_{i_j}=1\}\right) = \frac{\binom{n}{i_k}\left(\prod\limits_{j=2}^{k}\frac{(i_j-1)!}{i_{j-1}!}\right)(i_1-1)!(n-i_k)!}{n!}$$

$$= \frac{\prod\limits_{j=1}^{k}(i_j-1)!}{\prod\limits_{j=1}^{k} i_j!} = \frac{1}{i_1\, i_2 \dots i_k}\cdot$$

Le cas où $k=1$, donne, pour tout $i \in [\![1,n]\!]$, $\mathbb{P}(Z_i = 1) = \dfrac{1}{i}\cdot$ On a donc

$$\mathbb{P}\left(\bigcap_{j=1}^{k}\{Z_{i_j}=1\}\right) = \prod_{j=1}^{k}\mathbb{P}(Z_{i_j}=1).$$

Les variables Z_1, Z_2, $\dots$, Z_n sont donc indépendantes.

(b) Il est clair que $X_n = \sum\limits_{k=1}^{n} Z_k$. De l'indépendance des variables Z_k, on déduit $g_n = \prod\limits_{k=1}^{n} h_k$, où h_k est la fonction génératrice de Z_k. Pour tout réel t, on a :

$$k_k(t) = \mathbb{P}(Z_k = 0) + t\,\mathbb{P}(Z_k = 1) = 1 - \frac{1}{k} + \frac{t}{k} = \frac{k-1+t}{k}\cdot$$

On en déduit :

$$g_n(t) = \prod_{k=1}^{n}\frac{k-1+t}{k} = \frac{1}{n!}\prod_{k=0}^{n-1}(k+t).$$

(c) Le coefficient de t^k dans $g_n(t)$ est $\mathbb{P}(X_n = k)$. On obtient :

$$\mathbb{P}(X_n = 1) = \frac{1}{n},\quad \mathbb{P}(X_n = 2) = \frac{1}{n}\left(\sum_{k=1}^{n-1}\frac{1}{k}\right),\ \ \mathbb{P}(X_n = n) = \frac{1}{n!}\cdot$$

(d) Pour $n \geqslant 1$ et $t > 0$, on a :

$$\ln g_n(t) = -\ln n! + \sum_{k=0}^{n-1}\ln(k+t) \ \text{ et donc }\ \frac{g_n'(t)}{g_n(t)} = \sum_{k=0}^{n-1}\frac{1}{k+t}\cdot$$

Comme $g_n(1) = 1$, on obtient :

$$\forall n \geqslant 1 \quad \mathbb{E}(X_n) = g_n'(1) = \sum_{k=1}^{n} \frac{1}{k}\cdot$$

En dérivant de nouveau, on trouve :

$$\frac{g_n''(t)}{g_n(t)} - \frac{(g_n'(t))^2}{(g_n(t))^2} = -\sum_{k=0}^{n-1} \frac{1}{(k+t)^2}\cdot$$

On en déduit :

$$g_n''(1) - \big(\mathbb{E}(X_n)\big)^2 = -\sum_{k=1}^{n} \frac{1}{k^2}\cdot$$

On a donc :

$$\mathbb{V}(X_n) = g_n''(1) + \mathbb{E}(X_n) - \big(\mathbb{E}(X_n)\big)^2 = \sum_{k=1}^{n} \frac{1}{k} - \sum_{k=1}^{n} \frac{1}{k^2}\cdot$$

On a $\mathbb{E}(X_n) \sim \ln n$ et $\mathbb{V}(X_n) \sim \ln n$ quand n tend vers $+\infty$.

16.11 1. L'application $1 - \prod\limits_{i=1}^{n}(1 - \mathbb{1}_{A_i})$ est à valeurs dans $\{0,1\}$. Pour tout $\omega \in \Omega$, on a :

$$1 - \prod_{i=1}^{n}\big(1 - \mathbb{1}_{A_i}(\omega)\big) = 1 \iff \exists i \in [\![1,n]\!] \quad \mathbb{1}_{A_i}(\omega) = 1 \iff \omega \in \bigcup_{i=1}^{n} A_i.$$

On en déduit $1 - \prod\limits_{i=1}^{n}(1 - \mathbb{1}_{A_i}) = \mathbb{1}_{\bigcup\limits_{i=1}^{n} A_i}$. En développant, on obtient :

$$\begin{aligned}
\mathbb{1}_{\bigcup\limits_{i=1}^{n} A_i} &= 1 - \sum_{I \subset [\![1,n]\!]} (-1)^{\operatorname{card} I} \prod_{i \in I} \mathbb{1}_{A_i} \\
&= \sum_{\substack{I \subset [\![1,n]\!] \\ I \neq \varnothing}} (-1)^{\operatorname{card} -1} \mathbb{1}_{\bigcap\limits_{i\in I} A_i} \qquad = \sum_{k=1}^{n} (-1)^{k-1} \sum_{\substack{I \subset [\![1,n]\!] \\ \operatorname{card} I = k}} \mathbb{1}_{\bigcap\limits_{i\in I} A_i}\cdot
\end{aligned}$$

Sachant que pour tout événement A, on a $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$, on en déduit, par linéarité de l'espérance :

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{k=1}^{n} \bigg((-1)^{k-1} \sum_{\substack{I \subset [\![1,n]\!] \\ \operatorname{card} I = k}} \mathbb{P}\big(\bigcap_{i \in I} A_i\big)\bigg).$$

2. (a) En posant $I = \{j_1, j_2, \ldots, j_k\}$, on a :

$$B_{j_1,m} \cap B_{j_2,m} \cap \cdots \cap B_{j_k,m} = \bigcap_{i=1}^{m} \{X_i \notin I\}.$$

Les variables X_i (pour $i \in \mathbb{N}^*$) étant indépendantes, on en déduit :

$$\mathbb{P}(B_{j_1,m} \cap B_{j_2,m} \cap \cdots \cap B_{j_k,m}) = \bigcap_{i=1}^{m} \mathbb{P}(X_i \notin I) = \left(\frac{n-k}{n}\right)^m.$$

(b) On a $\{X > m\} = \bigcup_{j=1}^{n} B_{j,m}$. On en déduit en appliquant la formule du crible :

$$\begin{aligned}\mathbb{P}(X > m) &= \sum_{k=1}^{n}(-1)^{k-1} \sum_{\substack{I\subset[\![1,n]\!] \\ \text{card } I=k}} \mathbb{P}\Big(\bigcap_{j\in I} B_{j,m}\Big) \\ &= \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}\left(\frac{n-k}{n}\right)^m.\end{aligned}$$

On en déduit $\lim\limits_{m\to+\infty} \mathbb{P}(X > m) = 0$ car on a une combinaison linéaire de suites géométriques de raison appartenant à $]0,1[$. La suite $(\{X > m\})$ décroît donc :

$$0 = \lim_{m\to+\infty} \mathbb{P}(X > m) = \mathbb{P}\Big(\bigcap_{m\in\mathbb{N}^*} \{X > m\}\Big) = \mathbb{P}(X = +\infty).$$

La variable aléatoire X est presque sûrement à valeurs dans $\mathbb{N}^*$.

(c) On a :

$$\begin{aligned}\mathbb{E}(X) = \sum_{m=0}^{+\infty} \mathbb{P}(X > m) &= \sum_{m=0}^{+\infty}\sum_{k=1}^{n}\left((-1)^{k-1}\binom{n}{k}\left(\frac{n-k}{n}\right)^m\right) \\ &= \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}\sum_{m=0}^{+\infty}\left(\frac{n-k}{n}\right)^m \\ &= \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}\frac{1}{1-\frac{n-k}{n}} \\ &= n\sum_{k=1}^{n}(-1)^{k-1}\frac{\binom{n}{k}}{k}\cdot\end{aligned}$$

(d) Pour tout $t \in [0,1]$, on pose $f(t) = n\sum\limits_{k=1}^{n}(-1)^{k-1}\dfrac{\binom{n}{k}t^k}{k}\cdot$ La fonction f est dérivable et pour $t \in]0,1]$, on obtient :

$$f'(t) = n\sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}t^{k-1} = -\frac{n}{t}\sum_{k=1}^{n}\binom{n}{k}(-t)^k = n\frac{1-(1-t)^n}{t}\cdot$$

On en déduit :

$$\begin{aligned}\mathbb{E}(X) = f(1) &= \int_0^1 f'(t)dt = n\int_0^1 \frac{1-(1-t)^n}{t}dt \\ &= n\int_0^1 \frac{1-u^n}{1-u}du = n\sum_{j=0}^{n-1}\int_0^1 u^j du \\ &= n\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right).\end{aligned}$$

On en déduit $\mathbb{E}(X) \sim n\ln n$.

Remarque

On retrouve par une autre méthode le résultat de l'exercice 19 de la page 913.

16.12 1. On a $X_i^{(m)} = \sum\limits_{k=1}^{m} \mathbb{1}_{\{U_k=i\}}$. Les variables U_k étant indépendantes, il en est de même des variables $\mathbb{1}_{\{U_k=i\}}$. Pour tout $k \in [\![1,m]\!]$, on a :

$$\mathbb{E}\left(\mathbb{1}_{\{U_k=i\}}\right) = \mathbb{P}(U_k = i) = \frac{1}{n}\cdot$$

On en déduit que $X_i^{(m)}$ suit la loi binomiale de paramètre $\left(m, \frac{1}{n}\right)$.

2. Par bilinéarité de la covariance, on obtient :

$$\operatorname{Cov}\left(X_i^{(m)}, X_j^{(m)}\right) = \sum_{1\leqslant k,\ell\leqslant m} \operatorname{Cov}\left(\mathbb{1}_{\{U_k=i\}}, \mathbb{1}_{\{U_\ell=j\}}\right).$$

- Pour $k \neq \ell$, les variables $\mathbb{1}_{\{U_k=i\}}$ et $\mathbb{1}_{\{U_\ell=j\}}$ sont indépendantes ; leur covariance est nulle.
- Pour $k = \ell$, on obtient :
$$\operatorname{Cov}\left(\mathbb{1}_{\{U_k=i\}}, \mathbb{1}_{\{U_k=j\}}\right) = \mathbb{E}\left(\mathbb{1}_{\{U_k=i\}}\mathbb{1}_{\{U_k=j\}}\right) - \mathbb{E}\left(\mathbb{1}_{\{U_k=i\}}\right)\mathbb{E}\left(\mathbb{1}_{\{U_k=j\}}\right).$$
La variable $\mathbb{1}_{\{U_k=i\}}\mathbb{1}_{\{U_k=j\}}$ est nulle donc son espérance est nulle. On a donc $\operatorname{Cov}\left(\mathbb{1}_{\{U_k=i\}}, \mathbb{1}_{\{U_k=j\}}\right) = -\frac{1}{n^2}\cdot$

On obtient :

$$\operatorname{Cov}\left(X_i^{(m)}, X_j^{(m)}\right) = \sum_{k=1}^{m} \operatorname{Cov}\left(\mathbb{1}_{\{U_k=i\}}, \mathbb{1}_{\{U_k=j\}}\right) = -\frac{m}{n^2}\cdot$$

Les variables $X_i^{(m)}$ et $X_j^{(m)}$ ne sont pas indépendantes, car leur covariance n'est pas nulle.

3. (a) Pour tout $m \in \mathbb{N}$, la loi conditionnelle de Y_i sachant $\{N = m\}$ est la loi de $X_i^{(m)}$. On a donc, pour tout $k \in \mathbb{N}$, d'après la formule des probabilités totales :

$$\begin{aligned}\mathbb{P}(Y_i = k) &= \sum_{m=0}^{+\infty} \mathbb{P}(Y_i = k \mid N = m)\mathbb{P}(N = m)\\ &= \sum_{m=0}^{+\infty} \mathbb{P}(X_i^{(m)} = k \mid N = m)\mathbb{P}(N = m).\end{aligned}$$

La variable aléatoire N est indépendante des U_k donc est indépendante des variables $X_i^{(m)}$, car $X_i^{(m)}$ est une fonction de U_1, U_2,… U_m. On obtient donc :

$$\begin{aligned}\mathbb{P}(Y_i = k) &= \sum_{m=0}^{+\infty} \mathbb{P}(X_i^{(m)} = k)\mathbb{P}(N = m)\\ &= \sum_{m=k}^{+\infty} \binom{m}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{m-k} e^{-\lambda}\frac{\lambda^m}{m!}\\ &= \frac{e^{-\lambda}}{k!}\left(\frac{\lambda}{n}\right)^k \sum_{m=k}^{+\infty} \lambda^{m-k}\left(1 - \frac{1}{n}\right)^{m-k} \frac{1}{(m-k)!}\\ &= \frac{e^{-\lambda}}{k!}\left(\frac{\lambda}{n}\right)^k e^{\lambda - \frac{\lambda}{n}} = e^{-\lambda/n}\frac{(\lambda/n)^k}{k!}\cdot\end{aligned}$$

La variable aléatoire Y_i suit la loi de Poisson de paramètre $\dfrac{\lambda}{n}\cdot$

(b) Il est clair que, par définition, on a $Y_1 + Y_2 + \cdots + Y_n = N$.
Soit $(k_1, \ldots, k_n) \in \mathbb{N}^n$. On pose $m = k_1 + \cdots + k_n$. On a alors :

$$\begin{aligned} p &= \mathbb{P}\big(Y_1 = k_1, \ldots, Y_n = k_n\big) \\ &= \mathbb{P}\big(Y_1 = k_1, \ldots, Y_n = k_n, N = m\big) \\ &= \mathbb{P}\big(Y_1 = k_1, \ldots, Y_n = k_n \mid N = m\big)\,(N = m) \\ &= \mathbb{P}\big(X_1^{(m)} = k_1, \ldots, X_n^{(m)} = k_n \mid N = m\big)\mathbb{P}(N = m) \\ &= \mathbb{P}\big(X_1^{(m)} = k_1, \ldots, X_n^{(m)} = k_n\big)\mathbb{P}(N = m), \end{aligned}$$

car la variable N est indépendante des U_k donc de $(X_1^{(m)}, \ldots, X_n^{(m)})$ qui est une fonction de $U_1, \ldots, U_m$.

L'événement $\{X_1^{(m)} = k_1, \ldots, X_n^{(m)} = k_n\}$ est réalisé si, et seulement si, parmi les variables aléatoires $U_1, \ldots, U_m$, k_1 prennent la valeur 1, k_2 la valeur 2, $\ldots$, k_n la valeur n. On obtient :

$$\mathbb{P}\left(X_1^{(m)} = k_1, \ldots, X_n^{(m)} = k_n\right) = \frac{m!}{\prod\limits_{i=1}^{n} k_i!}\left(\frac{1}{n}\right)^m,$$

car le nombre de façons de choisir dans $[\![1, m]\!]$, k_1 entiers, puis k_2 entiers parmi les entiers restants,..., et enfin k_n entiers est :

$$\binom{m}{k_1}\binom{m-k_1}{k_2}\cdots\binom{m-k_1-\ldots k_{n-1}}{k_n} = \frac{m!}{\prod\limits_{i=1}^{n} k_i!},$$

alors que la probabilité que chaque variable aléatoire prenne une valeur donnée est $\dfrac{1}{n}\cdot$ On a donc :

$$\begin{aligned} \mathbb{P}\big(Y_1 = k_1, \ldots, Y_n = k_n\big) &= \frac{m!}{\prod\limits_{i=1}^{n} k_i!}\left(\frac{1}{n}\right)^m e^{-\lambda}\frac{\lambda^m}{m!} \\ &= \frac{e^{-\lambda}}{\prod\limits_{i=1}^{n} k_i!}\left(\frac{\lambda}{n}\right)^m. \end{aligned}$$

16.13 1. L'application Z est à valeurs dans $\mathbb{N} \cup \{+\infty\}$, ensemble dénombrable.

- Soit $n \in \mathbb{N}$. Pour tout $\omega \in \Omega$, on a $Z(\omega) = n$ si, et seulement si, ω appartient exactement à n des ensembles E_k (pour $k \in \mathbb{N}$). On a donc :

$$\{Z = n\} = \bigcup_{I \in \mathcal{P}_n(\mathbb{N})}\left(\bigcap_{k \in I} E_k\right) \cap \left(\bigcap_{k \in \mathbb{N}\setminus I} \overline{E_k}\right),$$

où $\mathcal{P}_n(\mathbb{N})$ est l'ensemble des parties de $\mathbb{N}$ de cardinal n.
Pour tout $I \in \mathcal{P}_n(\mathbb{N})$, $\left(\bigcap\limits_{k \in I} E_k\right) \cap \left(\bigcap\limits_{k \in \mathbb{N}\setminus I} \overline{E_k}\right)$ est un événement, car c'est l'intersection d'un nombre dénombrable d'événements.

D'autre part, on peut écrire $\mathcal{P}_n(\mathbb{N}) = \bigcup_{N\in\mathbb{N}} \mathcal{P}_n([\![0,N]\!])$, où $\mathcal{P}_n([\![0,N]\!])$ est l'ensemble des parties de cardinal n de $[\![0,N]\!]$. Donc $\mathcal{P}_n(\mathbb{N})$ est dénombrable, car réunion dénombrable d'ensembles finis.
Enfin, $\{Z = n\}$ est un événement, car réunion dénombrable d'événements.

- D'autre part $\{Z = +\infty\}$ est l'ensemble des éléments de Ω qui appartiennent à une infinité d'ensembles E_k $(k \in \mathbb{N})$. On a donc :

$$\{Z = +\infty\} = \bigcap_{n\in\mathbb{N}} \bigcup_{p\geqslant n} E_p,$$

donc $\{Z = +\infty\}$ est un événement.

Ainsi, $\{Z = x\}$ est un événement, pour tout $x \in Z(\Omega)$, donc Z est une variable aléatoire discrète.

2. La série $\sum \mathbb{P}(E_n)$ converge. Il résulte du premier lemme de Borel-Cantelli (*cf.* l'exercice 5 de la page 853) que

$$\mathbb{P}(Z = +\infty) = \mathbb{P}\Big(\bigcap_{n\in\mathbb{N}} \bigcup_{p\geqslant n} E_p\Big) = 0.$$

3. Comme $Z \geqslant 0$, elle possède une espérance dans $[0,+\infty]$. Pour tout $n \in \mathbb{N}$, posons $Z_n = \sum\limits_{i=0}^{n} \mathbb{1}_{E_i}$. Alors $\mathbb{E}(Z_n) = \sum\limits_{i=0}^{n} \mathbb{E}(\mathbb{1}_{E_i}) = \sum\limits_{i=0}^{n} \mathbb{P}(E_i)$.

Les variables Z_n sont à valeurs dans $\mathbb{N}$, donc on a aussi (*cf.* l'exercice 18 de la page 911) $\mathbb{E}(Z_n) = \sum\limits_{k=0}^{+\infty} \mathbb{P}(Z_n > k)$ et, comme Z prend la valeur $+\infty$ avec une probabilité nulle, $\mathbb{E}(Z) = \sum\limits_{k=0}^{+\infty} \mathbb{P}(Z > k)$.

Soit $k \in \mathbb{N}$. La suite (Z_n) a pour limite Z. On a donc :

$$\{Z > k\} \subset \bigcup_{n\in\mathbb{N}} \{Z_n > k\}.$$

Comme la suite $\big(\{Z_n > k\}\big)_{n\in\mathbb{N}}$ est croissance, on a, par continuité monotone :

$$\lim_{n\to+\infty} \mathbb{P}(Z_n > k) = \mathbb{P}\bigg(\bigcup_{n\in\mathbb{N}} \{Z_n > k\}\bigg) \geqslant \mathbb{P}(Z > k).$$

Pour tout $N \in \mathbb{N}$, on a :

$$\sum_{k=0}^{N} \mathbb{P}(Z > k) \leqslant \lim_{n\to+\infty} \sum_{k=0}^{N} \mathbb{P}(Z_n > k) \leqslant \lim_{n\to+\infty} \mathbb{E}(Z_n) \leqslant \sum_{i=0}^{+\infty} \mathbb{P}(E_i).$$

Cela est valable pour tout $N \in \mathbb{N}$, donc Z est d'espérance finie et :

$$\mathbb{E}(Z) \leqslant \sum_{i=0}^{+\infty} \mathbb{P}(E_i).$$

Mais d'autre part, on a $Z \geqslant Z_n$ et donc $\mathbb{E}(Z) \geqslant \mathbb{E}(Z_n) = \sum_{i=0}^{n} \mathbb{P}(E_i)$ pour tout $n \in \mathbb{N}$. Par passage à la limite, on obtient $\mathbb{E}(Z) \geqslant \sum_{i=0}^{+\infty} \mathbb{P}(E_i)$ et donc :

$$\mathbb{E}(Z) = \sum_{i=0}^{+\infty} \mathbb{P}(E_i).$$

16.14 1. Pour tout entier n strictement positif, on a $X_n = \prod_{k=1}^{n} Z_k$. Pour tout $k \geqslant 1$, on a $\mathbb{E}(Z_k) = 1$ et, comme les variables Z_k sont indépendantes :

$$\mathbb{E}(X_n) = \prod_{k=1}^{n} \mathbb{E}(Z_k) = 1.$$

2. Pour tout $k \geqslant 1$, on a $\mathbb{E}(Z_k^2) = \dfrac{(1-a)^2 + (1+a)^2}{2} = 1 + a^2$. Par indépendance des variables Z_k^2, on en déduit :

$$\mathbb{E}(X_n^2) = \prod_{k=1}^{n} \mathbb{E}(Z_k^2) = (1+a^2)^n$$

puis $\mathbb{V}(X_n) = (1+a^2)^n - 1$. Ainsi, $\mathbb{V}(X_n)$ tend vers $+\infty$ quand n tend vers $+\infty$.

3. Les variables aléatoires sont finies donc possèdent une espérance finie et une variance. On obtient, pour tout $k \geqslant 1$, par la formule de transfert :

$$\mathbb{E}(Y_k) = \frac{1}{2}\big(\ln(1-a) + \ln(1+a)\big) = \ln\sqrt{1-a^2}.$$

Les variables Y_k sont indépendantes, car fonctions des variables Z_k qui sont indépendantes, de même loi, possèdent une espérance finie $\ln\sqrt{1-a^2}$ et une variance donc, d'après la loi faible des grands nombres, on a :

$$\forall \varepsilon > 0 \quad \lim_{n\to+\infty} \mathbb{P}\Big(\big|\widehat{Y}_n - \ln\sqrt{1-a^2}\big| \geqslant \varepsilon\Big) = 0.$$

Comme $0 \leqslant \mathbb{P}\big(\widehat{Y}_n > \ln\sqrt{1-a^2} + \varepsilon\big) \leqslant \mathbb{P}\big(\big|\widehat{Y}_n - \ln\sqrt{1-a^2}\big| \geqslant \varepsilon\big)$, on a *a fortiori* :

$$\forall \varepsilon > 0 \quad \lim_{n\to+\infty} \mathbb{P}\left(\widehat{Y}_n > \ln\sqrt{1-a^2} + \varepsilon\right) = 0.$$

On a $\ln(1-a^2) < 0$, donc on peut choisir $\varepsilon \in \left]0, -\ln(1-a^2)\right[$.

On pose alors $\delta = -\ln\sqrt{1-a^2} - \varepsilon$. On a bien $\delta > 0$ et $\mathbb{P}\big(\widehat{Y}_n > -\delta\big) \to 0$ quand n tend vers l'infini.

4. On remarque que $\widehat{Y}_n = \dfrac{\ln Z_1 + \cdots + \ln Z_n}{n} = \dfrac{\ln X_n}{n}\cdot$ On a donc :

$$\mathbb{P}\left(\widehat{Y}_n > -\delta\right) = \mathbb{P}(\ln X_n > -n\delta) = \mathbb{P}(X_n > e^{-n\delta}).$$

Comme $\lim\limits_{n\to+\infty} e^{-n\delta} = 0$, il existe $n_0 \in \mathbb{N}^*$ tel que, pour $n \geqslant n_0$, on ait $e^{-n\delta} \leqslant \varepsilon$.

Pour $n \geqslant n_0$, on a $\{X_n > \varepsilon\} \subset \{X_n > e^{-n\delta}\}$ et donc :

$$0 \leqslant \mathbb{P}(X_n > \varepsilon) \leqslant \mathbb{P}\big(\widehat{Y}_n > -\delta\big).$$

On en déduit que $\mathbb{P}(X_n > \varepsilon)$ tend vers 0 quand n tend vers l'infini.

16.15 1. Voir cours.

2. Numérotons les urnes de 1 à n et considérons pour $1 \leqslant i \leqslant n$, la variable X_i indicatrice de l'événement « la i-ème urne est vide ». Il est clair que $Y_n = \sum\limits_{i=1}^{n} X_i$.

L'événement $\{X_i = 1\}$ est réalisé si toutes les boules ont été placées dans une des $n-1$ autres urnes. On a donc $\mathbb{P}(X_i = 1) = \left(\dfrac{n-1}{n}\right)^{na}$. Comme X_i suit une loi de Bernoulli, on en déduit

$$\mathbb{E}(X_i) = \left(\frac{n-1}{n}\right)^{na} \quad \text{et} \quad \mathbb{V}(X_i) = \left(\frac{n-1}{n}\right)^{na}\left(1 - \left(\frac{n-1}{n}\right)^{na}\right).$$

Par linéarité de l'espérance, on a :

$$\mathbb{E}(S_n) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i) = \left(\frac{n-1}{n}\right)^{na}.$$

D'autre part, on a :

$$\mathbb{V}(S_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathbb{V}(X_i) + 2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i, X_j)\right).$$

Il faut calculer les covariances. Pour $i \neq j$, on a :

$$\mathbb{E}(X_iX_j) = \mathbb{P}(X_iX_j = 1) = \mathbb{P}\big(\{X_i = 1\} \cap \{X_j = 1\}\big) = \left(\frac{n-2}{n}\right)^{na},$$

car il faut que les na boules soient placées dans les $n-2$ urnes restantes. On a donc :

$$\mathrm{Cov}(X_i, X_j) = \left(\frac{n-2}{n}\right)^{na} - \left(\frac{n-1}{n}\right)^{2na}.$$

Toutes les variances sont égales, ainsi que toutes les covariances. On en déduit :

$$\begin{aligned}\mathbb{V}(S_n) &= \frac{1}{n^2}\left(n\left(\frac{n-1}{n}\right)^{na}\left(1-\left(\frac{n-1}{n}\right)^{na}\right)\right.\\ &\qquad\qquad \left. + \, n(n-1)\left(\left(\frac{n-2}{n}\right)^{na} - \left(\frac{n-1}{n}\right)^{2na}\right)\right)\\ &= \frac{1}{n}\left(\frac{n-1}{n}\right)^{na} - \left(\frac{n-1}{n}\right)^{2na} + \frac{n-1}{n}\left(\frac{n-2}{n}\right)^{na}.\end{aligned}$$

3. On a $\lim\limits_{n\to+\infty}\mathbb{E}(S_n) = \lim\limits_{n\to+\infty} e^{na\ln\left(1-\frac{1}{n}\right)} = e^{-a}$.

De $\lim\limits_{n\to+\infty}\left(\dfrac{n-1}{n}\right)^{2na} = \lim\limits_{n\to+\infty}\left(\dfrac{n-2}{n}\right)^{na} = e^{-2a}$, on déduit :

$$\lim_{n\to+\infty}\mathbb{V}(S_n) = 0.$$

Soit $\varepsilon > 0$. Puisque $\lim\limits_{n\to+\infty} \mathbb{E}(S_n) = e^{-a}$, il existe $n_0 \in \mathbb{N}^*$ tel que, pour $n \geqslant n_0$, on ait $|\mathbb{E}(S_n) - e^{-a}| \leqslant \frac{\varepsilon}{2}$. Pour $n \geqslant n_0$, on a par inégalité triangulaire :

$$|S_n - \mathbb{E}(S_n)| \geqslant |S_n - e^{-a}| - \left|\mathbb{E}(S_n) - e^{-a}\right| \geqslant |S_n - e^{-a}| - \frac{\varepsilon}{2}.$$

On en déduit :

$$\left\{|S_n - e^{-a}| \geqslant \varepsilon\right\} \subset \left\{|S_n - \mathbb{E}(S_n)| \geqslant \frac{\varepsilon}{2}\right\}.$$

On a donc, pour $n \geqslant n_0$:

$$\mathbb{P}\big(|S_n - e^{-a}| \geqslant \varepsilon\big) \leqslant \mathbb{P}\left(|S_n - \mathbb{E}(S_n)| \geqslant \frac{\varepsilon}{2}\right) \leqslant \frac{4\mathbb{V}(S_n)}{\varepsilon^2},$$

d'après l'inégalité de Bienaymé-Tchebychev. Comme $\lim\limits_{n\to+\infty} \mathbb{V}(S_n) = 0$, on en déduit :

$$\lim_{n\to+\infty} \mathbb{P}(|S_n - e^{-a}| \geqslant \varepsilon) = 0.$$

16.16 1. On a $u_0 = 0$ et $u_1 = \mathbb{P}(X_1 = 0) = \mathbb{P}(X = 0) = p_0$.

2. Il est clair que si $X_n = 0$, alors $X_{n+1} = 0$. On a donc

$$\{X_n = 0\} \subset \{X_{n+1} = 0\},$$

ce qui implique $u_n \leqslant u_{n+1}$. La suite $(u_n)_{n\in\mathbb{N}}$ est croissante et majorée par 1, donc elle converge.

3. On écrit, avec la formule des probabilités totales :

$$u_{n+1} = \mathbb{P}(X_{n+1} = 0) = \sum_{k=0}^{+\infty} \mathbb{P}(X_{n+1} = 0 \mid X_1 = k)\, \mathbb{P}(X_1 = k).$$

On a $\mathbb{P}(X_1 = k) = \mathbb{P}(X = k) = p_k$.

Supposons que $X_1 = k \geqslant 1$. Alors le nombre total de virus le $(n+1)$-ième jour est égal à $Y_1 + Y_2 + \cdots + Y_k$, où $Y_1, \ldots, Y_k$ sont des variables indépendantes représentant le nombre de descendants de la n-ième génération de chacun des k virus du premier jour. On a donc, sachant que les variables $Y_1, \ldots, Y_n$ sont également indépendantes de X_1 :

$$\begin{aligned}\mathbb{P}(X_{n+1} = 0 \mid X_1 = k) &= \mathbb{P}(Y_1 + \cdots + Y_k = 0 \mid X_1 = k)\\ &= \mathbb{P}(Y_1 + \cdots + Y_k = 0) = \prod_{i=1}^{k} \mathbb{P}(Y_i = 0).\end{aligned}$$

Comme la loi de reproduction est la même pour touts les virus, on a, pour $1 \leqslant i \leqslant k$, $\mathbb{P}(Y_i = 0) = \mathbb{P}(X_n = 0) = u_n$ et donc :

$$\mathbb{P}(X_{n+1} = 0 \mid X_1 = k) = u_n^k.$$

Cette formule reste vraie pour $k = 0$, car si $X_1 = 0$, alors $X_{n+1} = 0$. On obtient :

$$u_{n+1} = \sum_{k=0}^{+\infty} p_k u_n^k = f(u_n).$$

4. La limite de $(u_n)_{n\in\mathbb{N}}$ est un point fixe de f. On note que 1 est un point fixe de f. On considère $g = x \mapsto f(x) - x$. On sait que f et, donc g, est dérivable sur $[0,1[$:
$$\forall x \in [0,1[\quad g'(x) = f'(x) - 1 = \sum_{k=1}^{+\infty} kp_k x^{k-1} - 1.$$
La fonction g' est strictement croissante sur $[0,1[$, car
$$g''(x) = f''(x) = \sum_{k=2}^{+\infty} k(k-1)p_k x^{k-2} > 0$$
pour $x \in \,]0,1[$ (sinon $p_k = 0$ pour tout $k \geqslant 2$ et $p_0 + p_1 = 1$). De plus, on a $g'(1) = \mathbb{E}(X) - 1$ (si $\mathbb{E}(X) = +\infty$, la limite de g' en 1 est $+\infty$).
   - Si $\mathbb{E}(X) \leqslant 1$, on a $g'(1) \leqslant 0$ et $g'(x) < 0$ si $0 \leqslant x < 1$. La fonction g est strictement décroissante et ne s'annule qu'en 1. La suite (u_n) converge vers 1.
   - Si $\mathbb{E}(X) > 1$, on a $g'(1) > 0$. Comme $g'(0) = p_1 - 1 \leqslant p_0 + p_1 - 1 < 0$, la fonction g' change de signe en x_0 : g décroît puis croît.
     Comme $g(0) = p_0 > 0$ et $g(x_0) < g(1) = 0$, g s'annule une seconde fois sur $[0, x_0[$ en ℓ.
     Comme $u_0 = 0 \leqslant \ell$ et f croît, on a $u_n \leqslant \ell$ pour tout $n \in \mathbb{N}$. La limite de $(u_n)_{n\in\mathbb{N}}$ est ℓ.

   On a, comme la suite $(\{X_n = 0\})_{n\in\mathbb{N}}$ croît :
$$\lim_{n\to+\infty} u_n = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{X_n = 0\}\right).$$
La limite est donc la probabilité que la lignée de virus s'éteigne. Si $\mathbb{E}(X) \leqslant 1$, la lignée de virus s'éteint presque sûrement. Si $\mathbb{E}(X) > 1$, la probabilité que la lignée de virus s'éteigne est un réel $\ell \in \,]0,1[$.

16.17 1. On a $S_N(\Omega) \subset \{0\} \cup \bigcup_{n\in\mathbb{N}^*} S_n(\Omega)$. Les $S_n(\Omega)$ étant au plus dénombrables, il en est de même de leur union dénombrable, et *a fortiori* $S_N(\Omega)$ est au plus dénombrable. De plus :
$$\forall x \in S_N(\Omega) \quad \{S_N = x\} = \bigcup_{n\in\mathbb{N}} \{S_n = x\} \cap \{N = n\},$$
donc $\{S_N = x\}$, union dénombrable d'événements, est un événement et S_N est une variable aléatoire.

2. (a) On a $S_N(\Omega) = \mathbb{N}$ et d'après la formule des probabilités totales, pour tout $k \in \mathbb{N}$:
$$\mathbb{P}(S_N = k) = \sum_{n=0}^{+\infty} \mathbb{P}(S_N = k \mid N = n)\,\mathbb{P}(N = n).$$
On remarque que :
$$\mathbb{P}(S_N = k \mid N = n) = \mathbb{P}(S_n = k \mid N = n) = \mathbb{P}(S_n = k),$$

la dernière égalité résultant de l'indépendance des variables X_n par rapport à N. Pour $n \geqslant 1$, S_n est une somme de n variables de Bernoulli, indépendantes, de même paramètre p ; elle suit donc la loi binomiale de paramètre (n,p). On a donc :

$$\mathbb{P}(S_n = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si} \quad k \leqslant n \\ 0 & \text{si} \quad k > n. \end{cases}$$

On remarque que cette formule reste vérifiée si $n = 0$, car alors S_n prend la valeur 0, avec la probabilité 1. On en déduit :

$$\begin{aligned} \mathbb{P}(S_N = k) &= \sum_{n=k}^{+\infty} \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda} \frac{\lambda^n}{n!} \\ &= e^{-\lambda} \frac{(\lambda p)^k}{k!} \sum_{n=k}^{+\infty} \frac{((1-p)\lambda)^{n-k}}{(n-k)!} \\ &= e^{-\lambda} \frac{(\lambda p)^k}{k!} e^{(1-p)\lambda} = e^{-\lambda p} \frac{(\lambda p)^k}{k!}. \end{aligned}$$

La variable aléatoire S_N suit la loi de Poisson de paramètre λp.

(b) On procède comme dans la question précédente. On a $S_N(\Omega) = \mathbb{N}^*$ et, pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(S_N = k \mid N = n) = \mathbb{P}(S_n = k \mid N = n) = \mathbb{P}(S_n = k).$$

Comme les variables X_k sont à valeurs dans $\mathbb{N}^*$, il est clair que $\mathbb{P}(S_n = k) = 0$ si $k < n$. Supposons $k \geqslant n$. On peut alors écrire :

$$\{S_n = k\} = \bigcup_{(i_1, i_2, \dots, i_n) \in J_k} \{X_1 = i_1\} \cap \{X_2 = i_2\} \cap \dots \{X_n = i_n\},$$

où J_k est l'ensemble des n-listes d'entiers strictement positifs $(i_1, i_2, \dots, i_n)$ tels que $i_1 + i_2 + \dots + i_n = k$. Le cardinal de J_k est égal au nombre de $(n-1)$-listes $(j_1, j_2, \dots, j_{n-1})$ d'entiers tels que $1 \leqslant j_1 < j_2 \dots < j_{n-1} \leqslant k-1$, car à $(i_1, i_2, \dots, i_n)$, on peut associer bijectivement :

$$(j_1, j_2, \dots, j_{n-1}) = (i_1, i_1 + i_2, \dots, i_1 + i_2 + \dots + i_{n-1}).$$

On a donc $\operatorname{card}(J_k) = \binom{k-1}{n-1}$.

D'autre part, pour $(i_1, i_2, \dots, i_n) \in J_k$, on a, par indépendance des variables aléatoires X_i :

$$\begin{aligned} &\mathbb{P}\left(\{X_1 = i_1\} \cap \{X_2 = i_2\} \cap \dots \cap \{X_n = i_n\}\right) \\ &\qquad = \prod_{j=1}^{n} \mathbb{P}(X_j = i_j) = \prod_{j=1}^{n} p(1-p)^{i_j - 1} \\ &\qquad = p^n (1-p)^{\sum_{j=1}^{n} i_j - n} = p^n (1-p)^{k-n}. \end{aligned}$$

On obtient ainsi $\mathbb{P}(S_n = k) = \binom{k-1}{n-1}p^n(1-p)^{k-n}$. On en déduit, pour $k \in \mathbb{N}^*$:

$$\begin{aligned}\mathbb{P}(S_N = k) &= \sum_{n=1}^{+\infty} \mathbb{P}(S_N = k \mid N = n)\, \mathbb{P}(N = n)\\ &= \sum_{n=1}^{k} \binom{k-1}{n-1} p^n(1-p)^{k-n}p'(1-p')^{n-1}\\ &= pp' \sum_{j=0}^{k-1} \binom{k-1}{j} (1-p)^{k-1-j}(p(1-p'))^{n-1}\\ &= pp'(1-p+p(1-p'))^{k-1} = pp'(1-pp')^{k-1}.\end{aligned}$$

La variable S_N suit la loi géométrique de paramètre pp'.

3. (a) Pour tout $t \in [0,1]$,

$$G_{S_N}(t) = \sum_{k\in\mathbb{N}} t^k \mathbb{P}(S_N = k) = \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} t^k \mathbb{P}(\{S_N = k\} \cap \{N = n\}).$$

La famille de réels positifs $\big(t^k\mathbb{P}\left(\{S_N = k\} \cap \{N = n\}\right)\big)_{(k,n)\in\mathbb{N}^2}$ est donc sommable. On peut donc échanger l'ordre des sommations. Pour $n \in \mathbb{N}$, on a :

$$\begin{aligned}\sum_{k\in\mathbb{N}} t^k \mathbb{P}(\{S_N = k\} \cap \{N = n\}) &= \sum_{k\in\mathbb{N}} t^k \mathbb{P}(\{S_n = k\} \cap \{N = n\})\\ &= \sum_{k\in\mathbb{N}} t^k \mathbb{P}(S_n = k)\mathbb{P}(N = n),\end{aligned}$$

car S_n et N sont indépendantes. On en déduit :

$$\begin{aligned}\sum_{k\in\mathbb{N}} t^k \mathbb{P}(\{S_N = k\} \cap \{N = n\}) &= G_{S_n}(t)\mathbb{P}(N = n)\\ &= G_{X_1}(t)^n\, \mathbb{P}(N = n),\end{aligned}$$

car $G_{S_n}(t) = G_{X_1+\cdots+X_n}(t) = \prod\limits_{i=1}^{n} G_{X_i}(t) = G_{X_1}(t)^n$, par indépendance des variables X_i. On obtient finalement :

$$\forall t \in [0,1] \quad G_{S_N}(t) = \sum_{n\in\mathbb{N}} (G_{X_1}(t))^n\, \mathbb{P}(N = n) = G_N\big(G_{X_1}(t)\big),$$

ce qui est le résultat voulu.

(b) Les fonctions G_N et G_{X_1} sont dérivable en 1 et $G_{X_1}(1) = 1$. On en déduit que G_{S_N} est dérivable en 1 à gauche, donc dérivable en 1. Ainsi la variable aléatoire S_N est d'espérance finie :

$$\begin{aligned}\mathbb{E}(S_N) &= G'_{S_N}(1) = G'_N\left(G_{X_1}(1)\right) G'_{X_1}(1)\\ &= G'_N(1)G'_{X_1}(1) = \mathbb{E}(N)\mathbb{E}(X_1).\end{aligned}$$

(c) Les fonctions G_N et G_{X_1} sont deux fois dérivable en 1. On en déduit que G_{S_N}

est deux fois dérivable en 1 et :

$$\begin{aligned}G''_{S_N}(1) &= G'_N\left(G_{X_1}(1)\right)G''_{X_1}(1) + G''_N\left(G_{X_1}(1)\right)\left(G'_{X_1}(1)\right)^2\\ &= G'_N(1)G''_{X_1}(1) + G''_N(1)\left(G'_N(1)\right)^2.\end{aligned}$$

On a donc

$$\begin{aligned}\mathbb{E}\left(S_N(S_N-1)\right) &= \mathbb{E}(N)\mathbb{E}\left(X_1(X_1-1)\right) + \mathbb{E}\left(N(N-1)\right)\left(\mathbb{E}(X_1)\right)^2\\ &= \mathbb{E}(N)\left(\mathbb{E}\left(X_1^2\right) - \mathbb{E}(X_1)\right)\\ &\qquad + \left(\mathbb{E}(N^2) - \mathbb{E}(N)\right)\left(\mathbb{E}\left(X_1\right)\right)^2.\end{aligned}$$

On en déduit :

$$\begin{aligned}\mathbb{V}(S_N) &= \mathbb{E}\left(S_N(S_N-1)\right) + \mathbb{E}(S_N) - \left(\mathbb{E}(S_N)\right)^2\\ &= \mathbb{E}(N)\Big(\mathbb{V}(X_1) + \left(\mathbb{E}(X_1)\right)^2 - \mathbb{E}(X_1)\Big)\\ &\quad + \Big(\mathbb{V}(N) + \left(\mathbb{E}(N)\right)^2 - \mathbb{E}(N)\Big)\left(\mathbb{E}\left(X_1\right)\right)^2\\ &\quad + \mathbb{E}(X_1)\mathbb{E}(N) - \left(\mathbb{E}\left(X_1\right)\right)^2\left(\mathbb{E}\left(N\right)\right)^2\\ &= \mathbb{V}(X_1)\mathbb{E}(N) + \left(\mathbb{E}(X_1)\right)^2\mathbb{V}(N).\end{aligned}$$

4. (a) Pour tout $n \in \mathbb{N}$, la variable aléatoire S_n est d'espérance finie, comme somme de n variables aléatoires d'espérance finie, donc $|S_n|$ est aussi d'espérance finie. On a, d'après la formule de transfert :

$$\sum_{x \in S_N(\Omega)} |x|\mathbb{P}(S_n = x)\,\mathbb{P}(N = n) = \mathbb{E}(|S_n|)\,\mathbb{P}(N = n).$$

De l'inégalité $|S_n| \leqslant \sum\limits_{k=1}^{n} |X_k|$ on déduit, par croissance de l'espérance,

$$\mathbb{E}(|S_n|) \leqslant \sum_{k=1}^{n} \mathbb{E}(|X_k|) = n\mathbb{E}(|X_1|).$$

La famille $\left(n\mathbb{E}(|X_1|)\mathbb{P}(N = n)\right)_{n \in \mathbb{N}}$ est sommable, de somme $\mathbb{E}(|X_1|)\mathbb{E}(N)$. *A fortiori*, la famille $\left(\mathbb{E}(|S_n|)\,\mathbb{P}(N = n)\right)_{n \in \mathbb{N}}$ est sommable. On en déduit la sommabilité de la famille $\left(x\,\mathbb{P}(S_n = x)\mathbb{P}(N = n)\right)_{(x,n) \in S_N(\Omega) \times N(\Omega)}$.

(b) Notons S la somme de cette famille.
On obtient en refaisant le calcul précédent, mais sans valeur absolue :

$$\begin{aligned}S &= \sum_{n \in \mathbb{N}} \sum_{x \in S_N(\Omega)} x\mathbb{P}(S_n = x)\,\mathbb{P}(N = n) = \sum_{n \in \mathbb{N}} \mathbb{E}(S_n)\,\mathbb{P}(N = n)\\ &= \sum_{n \in \mathbb{N}} n\mathbb{E}(X_1)\,\mathbb{P}(N = n) = \mathbb{E}(X_1)\mathbb{E}(N).\end{aligned}$$

D'autre part, on a, en raisonnant comme dans la question 2 :

$$\begin{aligned}\sum_{n\in\mathbb{N}} x\,\mathbb{P}(S_n = x)\,\mathbb{P}(N=n) &= x\sum_{n\in\mathbb{N}} \mathbb{P}(S_n = x \mid N=n)\,\mathbb{P}(N=n)\\ &= x\sum_{n\in\mathbb{N}} \mathbb{P}(S_N = x \mid N=n)\,\mathbb{P}(N=n)\\ &= x\,\mathbb{P}(S_N = x).\end{aligned}$$

On sait qu'alors, la famille $(x\,\mathbb{P}(S_N = x))_{x\in S_N(\Omega)}$ est sommable de somme S. La variable S_N est donc d'espérance finie et :

$$\mathbb{E}(S_N) = \mathbb{E}(X_1)\mathbb{E}(N).$$

16.18 1. (a) On a $\Pi_0 = \Pi_1 = \dots \Pi_{r-1} = 0$ et $\Pi_r = p^r$.

(b) Pour $n \geqslant 1$, on note S_n l'événement « la n-ième épreuve est un succès » et A_n l'événement « au cours de n-premiers tirages, on a obtenus r succès consécutifs ». On a clairement $A_n \subset A_{n+1}$ et donc $\Pi_{n+1} - \Pi_n = \mathbb{P}(A_{n+1} \setminus A_n)$. L'événement $A_{n+1} \setminus A_n$ est réalisé si on obtient r succès consécutifs pour la première fois entre le $(n-r+2)$-ième et le $(n+1)$-ième tirage, ce qui impose que la $(n-r+1)$-ième épreuve soit un échec et qu'on n'ait pas avant obtenu r succès consécutifs. On a donc :

$$A_{n+1} \setminus A_n = \overline{A_{n-r}} \cap \overline{S_{n-r+1}} \cap S_{n-r+2} \cap \dots \cap S_{n+1}$$

et par indépendance des épreuves :

$$\Pi_{n+1} - \Pi_n = (1 - \Pi_{n-r})\,(1-p)\,p^r.$$

(c) La suite $(\Pi_n)_{n\in\mathbb{N}}$ est croissante et majorée par 1 donc convergente. On note L sa limite. Par passage à la limite dans la relation précédente, on obtient $L - L = (1-L)(1-p)p^r$, et donc $L = 1$, puisque $p \in\,]0,1[$.

2. (a) Comme la suite (A_n) est croissante, on a :

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}^*} A_n\right) = \lim_{n\to+\infty} \Pi_n = 1.$$

On obtient de manière presque sûre une suite de r succès consécutifs au bout d'un nombre fini d'épreuves. Sur un ensemble de probabilité 1, on peut définir l'application T.

On a, par définition $\{T = k\} = A_{k+1} \setminus A_k$, donc $\{T = k\}$ est un événement et T une variable aléatoire à valeurs dans $\mathbb{N}^*$.

(b) Pour tout $k \in \mathbb{N}$, on a $\{T > k] = \overline{A_k}$ et donc $\mathbb{P}(T > k) = 1 - \Pi_k$. D'après la question 1, on a pour tout $k \geqslant 0$:

$$1 - \Pi_k = \frac{\Pi_{k+r+1} - \Pi_{k+r}}{(1-p)p^r}\cdot$$

D'après l'exercice 18 de la page 911, on en déduit :

$$\mathbb{E}(T) = \sum_{k=0}^{+\infty} \frac{\Pi_{k+r+1} - \Pi_{k+r}}{(1-p)p^r} = \frac{\lim\limits_{n\to+\infty} \Pi_n - \Pi_r}{(1-p)p^r} = \frac{1-p^r}{(1-p)p^r}\cdot$$

16.19 1. (a) On a $S_n = 0$ si, et seulement si, parmi les n variables X_i pour $1 \leqslant i \leqslant n$, il y en a autant qui prennent la valeur 1 et la valeur -1. Cela nécessite n pair. On a donc $u_n = 0$ si n est impair et, si n est pair :

$$u_n = \binom{n}{\frac{n}{2}} p^{\frac{n}{2}} (1-p)^{\frac{n}{2}} = \binom{n}{\frac{n}{2}} \big(p(1-p)\big)^{\frac{n}{2}}.$$

(b) On a, pour tout $x \in \,]-1,1[$,

$$\frac{1}{\sqrt{1-x}} = \sum_{n=0}^{+\infty} \frac{\binom{2n}{n} x^n}{4^n}.$$

Or $4p(1-p) \leqslant 1$, donc, pour tout $x \in \,]-1,1[$:

$$\begin{aligned}\frac{1}{\sqrt{1-4p(1-p)x^2}} &= \sum_{n=0}^{+\infty} \frac{\binom{2n}{n}(4p(1-p))^n x^{2n}}{4^n} \\ &= \sum_{n=0}^{+\infty} u_{2n} x^{2n} = f(x),\end{aligned}$$

car $u_n = 0$ si n est impair.

2. (a) Les événements A_k pour $1 \leqslant k \leqslant n$ sont incompatibles et :

$$\{S_n = 0\} \subset \bigcup_{1 \leqslant k \leqslant n} A_k.$$

En effet si $S_n = 0$, en considérant k, le plus petit des entiers j tels que $S_j = 0$, l'événement A_k est réalisé et $1 \leqslant k \leqslant n$.

On a donc $\{S_n = 0\} = \bigcup\limits_{k=1}^{n} (\{S_n = 0\} \cap A_k)$ et comme on a une réunion finie d'événements incompatibles :

$$\mathbb{P}(S_n = 0) = \sum_{k=1}^{n} \mathbb{P}\left(\{S_n = 0\} \cap A_k\right).$$

(b) On suppose $1 \leqslant k \leqslant n$. Comme

$$S_1 = X_1, S_2 = X_1 + X_2, \ldots, S_k = X_1 + X_2 + \cdots X_k,$$

on peut écrire A_k sous la forme :

$$A_k = \{f(X_1, X_2, \ldots, X_k) \in N_k\},$$

en posant $f(x_1, x_2, \ldots, x_k) = (x_1, x_1 + x_2, \ldots, x_1 + \cdots + x_n)$ et $N_k = {\mathbb{N}^*}^{k-1} \times \{0\}$. On a alors :

$$\begin{aligned}\mathbb{P}\big(\{S_n = 0\} \cap A_k\big) &= \mathbb{P}\big(\{X_{k+1} + \cdots + X_n = 0\} \cap A_k\big) \\ &= \mathbb{P}\big(\{X_{k+1} + \cdots + X_n = 0\} \cap \{f(X_1, X_2, \ldots, X_k) \in N_k\}\big) \\ &= \mathbb{P}\big(X_{k+1} + \cdots + X_n = 0\big)\, \mathbb{P}(f(X_1, X_2, \ldots, X_k) \in N_k),\end{aligned}$$

car les variables aléatoires $X_{k+1} + \cdots + X_n$ et $f(X_1, X_2, \ldots, X_k)$ sont indépendantes (puisque de $X_1, \ldots, X_n$ sont indépendantes). On obtient :

$$\mathbb{P}\big(\{S_n = 0\} \cap A_k\big) = \mathbb{P}\big(X_{k+1} + \cdots + X_n = 0\big)\, \mathbb{P}(A_k).$$

Les variables aléatoires X_i sont indépendantes et suivent la même loi, donc la loi de $X_{k+1} + X_{k+2} + \cdots + X_n$ est la même que celle de $X_1 + X_2 + \cdots + X_{n-k} = S_{n-k}$. On a donc :

$$\mathbb{P}\big(\{S_n = 0\} \cap A_k\big) = \mathbb{P}\big(S_{n-k} = 0\big)\mathbb{P}(A_k) = u_{n-k}v_k.$$

Comme de plus $v_0 = 0$, on obtient, d'après la question précédente :

$$u_n = \mathbb{P}(S_n = 0) = \sum_{k=0}^{n} u_{n-k}v_k.$$

3. (a) Comme $0 \leqslant v_n \leqslant 1$, le rayon de la série entière définissant $g(x)$ est supérieur ou égal à 1. La fonction fg possède un développement en série entière dont les coefficients sont obtenus en faisant le produit de Cauchy des suites (u_n) et (v_n).

   Pour $n \geqslant 1$, on a $\displaystyle\sum_{k=0}^{n} u_{n-k}v_k = u_n$. Pour $n = 0$, $u_0v_0 = 0$. On obtient

   $$\forall x \in\,]-1,1[\quad f(x)g(x) = \sum_{n=1}^{+\infty} u_n x^n = f(x) - 1.$$

   On en déduit :

   $$\forall x \in\,]-1,1[\quad g(x) = \frac{f(x)-1}{f(x)} = 1 - \sqrt{1-4p(1-p)x^2}.$$

   (b) On a :

   $$A = \bigcup_{n\in\mathbb{N}^*} A_k.$$

   Comme les A_k sont deux à deux incompatibles, on en déduit :

   $$\mathbb{P}(A) = \sum_{k=1}^{+\infty} v_k = g(1).$$

   On a pour $x \in\,]-1,1[$, $g(x) = 1 - \sqrt{1-4p(1-p)x^2}$.
   La série $\sum v_n$ est convergente. Donc la série $\sum v_n x^n$ est une série de fonctions continues, normalement convergente sur $[-1,1]$. La fonction g est donc continue sur $[-1,1]$ et :

   $$\mathbb{P}(A) = g(1) = \lim_{x\to 1^-} g(x) = 1 - |2p-1|.$$

   - Si $p = \dfrac{1}{2}$, on a $\mathbb{P}(A) = 1$. Il est presque sûr que le mobile repassera à l'origine.
   - Si $p \neq \dfrac{1}{2}$, on a $\mathbb{P}(A) < 1$. Le mobile peut ne pas repasser à l'origine.

4. (a) Il est presque sûr qu'il existe $n \in \mathbb{N}$ tel que $\{S_n = 0\}$ et réalisé et donc un plus petit entier n pour lequel $\{S_n = 0\}$ est réalisé. On note T ce plus petit entier. On a donc $\{T = k\} = A_k$. Comme A_k est un événement, on définit bien une variable aléatoire T.

(b) On a, pour tout $x \in]-1,1[$:

$$g(x) = 1 - \sqrt{1-x^2} = 1 - \left(1 - \sum_{n=1}^{+\infty} \frac{(2n-2)!x^{2n}}{2^{2n-1}n!(n-1)!}\right)$$
$$= \sum_{n=1}^{+\infty} \frac{2\binom{2n-2}{n-1}x^{2n}}{n4^n}.$$

On a donc :

$$\forall n \in \mathbb{N}^* \quad v_{2n} = \frac{2\binom{2n-2}{n-1}}{n4^n}.$$

On a bien entendu $v_n = 0$ si n est impair.

(c) La variable T est d'espérance finie si, et seulement si, la série de terme général $2nv_{2n}$ converge. En utilisant la formule de Stirling, on obtient :

$$2nv_{2n} \sim \frac{1}{\sqrt{n\pi}}.$$

La série diverge, donc T n'est pas d'espérance finie.

5. (a) Pour $p = \dfrac{1}{2}$, on a $u_{2n} = \dfrac{\binom{2n}{n}}{4^n}$. En remplaçant u_{2n}, u_{2n-2} et v_{2n} par leur valeur, on montre facilement que $v_{2n} = u_{2n-2} - u_{2n}$ pour $n \geqslant 1$.

(b) La probabilité que le mobile ne soit jamais à l'origine à l'issue d'aucun des $2n$ premiers déplacements est :

$$1 - \mathbb{P}\left(\bigcup_{k=1}^{2n} A_k\right) = 1 - \sum_{k=1}^{n} v_{2k} = 1 - \sum_{k=1}^{n} (u_{2k-2} - u_{2k})$$
$$= 1 - (u_0 - u_{2n}) = u_{2n} = \mathbb{P}(S_{2n} = 0),$$

car $u_0 = 1$.

Comme les valeurs prises par $(S_{n+1} - S_n)$ sont ± 1, la suite (S_n) ne peut pas prendre une valeur strictement positive et une valeur strictement négative sans prendre entre les deux la valeur 0. On en déduit :

$$\mathbb{P}(S_1 \neq 0, S_2 \neq 0, S_{2n} \neq 0) = \mathbb{P}(S_1 > 0, S_2 > 0, S_{2n} > 0)$$
$$+ \mathbb{P}(S_1 < 0, S_2 < 0, S_{2n} < 0).$$

Comme $\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1)$ pour tout $i \in \mathbb{N}^*$, on a, pour des raisons de symétrie :

$$\mathbb{P}(S_1 > 0, S_2 > 0, S_{2n} > 0) = \mathbb{P}(S_1 < 0, S_2 < 0, S_{2n} < 0)$$

et donc :

$$\mathbb{P}(S_1 > 0, S_2 > 0, S_{2n} > 0) = \frac{1}{2}\mathbb{P}(S_1 \neq 0, S_2 \neq 0, S_{2n} \neq 0) = \frac{1}{2}u_{2n}.$$

16.20 1. (a) Par définition, on a $|g(X_1)| \leqslant c$: la variable aléatoire $g(X_1)$ est bornée donc d'espérance finie.

Les variables $g(X_1)$ et X_1 possèdent une espérance donc il en est de même de $|g(X_1)-X_1|$. Cette espérance est donnée par la formule de transfert :

$$\begin{aligned}\mathbb{E}\left(|g(X_1)-X_1|\right) &= \sum_{x\in X_1(\Omega)} |g(x)-x|\,\mathbb{P}(X_1=x)\\ &= \sum_{\substack{x\in X_1(\Omega)\\ |x|>c}} |x|\,\mathbb{P}(X_1=x).\end{aligned}$$

La famille $(|x|\,\mathbb{P}(X_1=x))_{x\in X_1(\Omega)}$ est sommable de somme $\mathbb{E}(|X_1|)$ donc il existe I, partie finie de $X_1(\Omega)$, telle que :

$$\sum_{x\in X_1(\Omega)\setminus I} |x|\,\mathbb{P}(X_1=x) \leqslant \frac{\varepsilon}{2}\cdot$$

On pose $c=\max\limits_{x\in I}|x|$. On a alors :

$$\mathbb{E}\left(|g(X_1)-X_1|\right) \leqslant \sum_{x\in X_1(\Omega)\setminus I} |x|\,\mathbb{P}(X_1=x) \leqslant \frac{\varepsilon}{2}\cdot$$

(b) On a, puisque X_1 est centrée :

$$|a| = |\mathbb{E}(g(X_1))\,| = |\mathbb{E}\left(g(X_1)-X_1\right)| \leqslant \mathbb{E}\left(|g(X_1)-X_1|\right) \leqslant \frac{\varepsilon}{2}\cdot$$

Par inégalité triangulaire et croissance de l'espérance, on en déduit :

$$\mathbb{E}\left(|g(X_1)-X_1-a|\right) \leqslant \mathbb{E}\left(|g(X_1)-X_1|\right)+|a| \leqslant \varepsilon.$$

(c) Toutes les variables X_n ont même loi, donc il en est de même des variables U_n. La fonction g est bornée, donc ce sont des variables bornées ; elles admettent donc une variance. Comme les variables X_n sont indépendantes, il en est de même des U_n et l'on a donc :

$$\mathbb{V}(Y'_n) = \frac{1}{n^2}\mathbb{V}(U_1+\cdots+U_n) = \frac{1}{n^2}\sum_{k=1}^n \mathbb{V}(U_k) = \frac{n\mathbb{V}(U_1)}{n^2} = \frac{\mathbb{V}(U_1)}{n}\cdot$$

On en déduit que $\lim\limits_{n\to+\infty}\mathbb{V}(Y'_n)=0$.

Par construction, les variables U_n sont centrées. On a donc $\mathbb{E}(Y'_n)=0$ et $\mathbb{V}(Y'_n)=\mathbb{E}\left({Y'_n}^2\right)$. De l'inégalité de Cauchy-Schwarz, on déduit :

$$\big(\mathbb{E}\left(|Y'_n|\right)\big)^2 = \big(\mathbb{E}(|Y'_n|\cdot 1)\big)^2 \leqslant \mathbb{E}\big({Y'_n}^2\big)\mathbb{E}(1) \leqslant \mathbb{V}(Y'_n).$$

Comme $\lim\limits_{n\to+\infty}\mathbb{V}(Y'_n)=0$, on a également $\lim\limits_{n\to+\infty}\mathbb{E}(|Y'_n|)=0$.

(d) On a, pour tout $n\in\mathbb{N}^*$:

$$|Y_n-Y'_n| = \frac{1}{n}\left|\sum_{k=1}^n (X_k-U_k)\right| \leqslant \frac{1}{n}\sum_{k=1}^n |X_k-U_k|.$$

Pour tout $k\in[\![1,n]\!]$, $\mathbb{E}\left(|X_k-U_k|\right)\leqslant\varepsilon$, car toutes les variables aléatoires X_k-U_k ont même loi que X_1-U_1. On en déduit, par croissance de l'espérance :

$$\mathbb{E}\big(|Y_n-Y'_n|\big) \leqslant \frac{1}{n}\sum_{k=1}^n \mathbb{E}\big(|X'_k-U_k|\big) \leqslant \varepsilon.$$

On a, pour tout $n \in \mathbb{N}^*$:
$$\mathbb{E}(|Y_n|) \leqslant \mathbb{E}(|Y_n - Y_n'|) + \mathbb{E}(|Y_n'|) \leqslant \varepsilon + \mathbb{E}(|Y_n'|).$$
Comme $(\mathbb{E}(|Y_n'|))$ converge vers 0, on en déduit pour n assez grand :
$$\mathbb{E}(|Y_n|) \leqslant 2\varepsilon.$$
Cela étant vrai pour tout $\varepsilon > 0$, on conclut :
$$\lim_{n\to+\infty} \mathbb{E}(|Y_n|) = 0.$$

2. On applique l'inégalité de Markov à la variable $|Y_n|$:
$$\forall \varepsilon > 0 \quad \mathbb{P}(|Y_n| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(|Y_n|)}{\varepsilon}.$$
Du résultat de la question précédente, on déduit $\lim\limits_{n\to+\infty} \mathbb{P}(|Y_n| \geqslant \varepsilon) = 0.$
3. Il suffit d'appliquer ce qui précède aux variables aléatoires $X_n - m$, pour $n \in \mathbb{N}^*$.

16.21 1. Les coefficients de P sont positifs car ce sont des probabilités et pour tout $i \in [\![1, N]\!]$, on a, pour tout $n \in \mathbb{N}$:
$$\sum_{j=1}^{N} p_{i,j} = \sum_{j=1}^{N} \mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_{n+1} \in [\![1, N]\!] \mid X_n = i) = 1.$$
Soit U le vecteur colonne à N lignes, dont tous les coefficients sont égaux à 1. On a $PU = U$, donc 1 est valeur propre de P.

2. (a) On démontre la propriété par récurrence sur n.
Pour $n = 1$, on a $\mathbb{P}(X_1 = x_1 \mid X_0 = x_0) = p_{x_0,x_1}$ par définition.
Supposons la propriété soit vraie au rang $n-1$ $(n \geqslant 2)$ et montrons-la au rang n.

- Si $\mathbb{P}\big(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}\big) \neq 0$, on peut écrire :
$$\begin{aligned}
&\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n \mid X_0 = x_0)\\
=&\frac{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)}{\mathbb{P}(X_0 = x_0)}\\
=&\frac{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)}{\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1})} \times \frac{\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1})}{\mathbb{P}(X_0 = x_0)}\\
=&\mathbb{P}(X_n = x_n \mid X_0 = x_0, \ldots, X_{n-1} = x_{n-1})\\
&\qquad \times \mathbb{P}(X_1 = x_1, \ldots, X_{n-1} = x_{n-1} \mid X_0 = x_0)\\
=&p_{x_{n-1},x_n} \times p_{x_0,x_1} p_{x_1,x_2} \cdots p_{x_{n-2},x_{n-1}}\\
=&p_{x_0,x_1} p_{x_1,x_2} \cdots p_{x_{n-2},x_{n-1}} p_{x_{n-1},x_n},
\end{aligned}$$
en utilisant la définition d'une chaîne de Markov et l'hypothèse de récurrence.
- Si $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}) = 0$, alors on a *a fortiori* :
$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n \mid X_0 = x_0) = 0$$
et l'égalité reste vérifiée car, $p_{x_0,x_1} \cdots p_{x_{n-2},x_{n-1}} = 0$ par hypothèse de récurrence.

La propriété est donc vérifiée pour tout $n \in \mathbb{N}$.
Pour obtenir la loi conditionnelle de X_n sachant $X_0 = x_0$, il faut sommer l'égalité précédente par rapport à $x_1, \ldots, x_{n-1}$. On obtient, pour x_0 et x_n dans $[\![1, N]\!]$:

$$\mathbb{P}(X_n = x_n \mid X_0 = x_0) = \sum_{(x_1,\ldots,x_{n-1})\in[\![1,N]\!]^{n-1}} p_{x_0,x_1} p_{x_1,x_2} \ldots p_{x_{n-1},x_n} = p^{(n)}_{x_0,x_n},$$

comme on le voit en réitérant la formule donnant le produit de deux matrices.

(b) L'événement $\{X_0 = x_0\}$ qui contient $\{X_0 = x_0, \ldots, X_n = x_n\}$ est *a fortiori* de probabilité non nulle et on a :

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = x_{n+1}, \ldots, X_{n+k} = x_{n+k} \mid X_0 = x_0, \ldots, X_n = x_n)\\
&\quad= \frac{\mathbb{P}(X_0 = x_0, \ldots, X_{n+k} = x_{n+k})}{\mathbb{P}(X_0 = x_0, \ldots, X_n = x_n)}\\
&\quad= \frac{\mathbb{P}(X_1 = x_1, \ldots, X_{n+k} = x_{n+k} \mid X_0 = x_0)}{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n \mid X_0 = x_0)}\\
&\quad= \frac{\prod\limits_{i=0}^{n+k-1} p_{x_i,x_{i+1}}}{\prod\limits_{i=0}^{n-1} p_{x_i,x_{i+1}}} = \prod_{i=n}^{n+k-1} p_{x_i,x_{i+1}}\\
&\quad= \mathbb{P}(X_1 = x_{n+1}, \ldots, X_k = x_{n+k} \mid X_0 = x_n),
\end{aligned}$$

d'après la question précédente.

3. (a) On note ℓ et ℓ' des éléments de $[\![1, N]\!]$ tels que $u_n = p^{(n)}_{\ell,j}$ et $v_n = p^{(n)}_{\ell',j}$. On a, pour tout $i \in [\![1, N]\!]$:

$$\begin{aligned}
p^{(n+1)}_{i,j} &= \sum_{k=1}^{N} p_{i,k} p^{(n)}_{k,j} = p_{i,\ell'} v_n + \sum_{k\in[\![1,N]\!]\setminus\{\ell'\}} p_{i,k} p^{(n)}_{k,j}\\
&\leqslant p_{i,\ell'} v_n + u_n \sum_{k\in[\![1,N]\!]\setminus\{\ell'\}} p_{i,k} = p_{i,\ell'} v_n + (1 - p_{i,\ell'}) u_n.
\end{aligned}$$

La fonction affine $x \mapsto x v_n + (1 - x) u_n$ décroît donc :

$$p^{(n+1)}_{i,j} \leqslant \varepsilon v_n + (1 - \varepsilon) u_n.$$

Comme cela est vrai pour tout $i \in [\![1, N]\!]$, on obtient :

$$u_{n+1} \leqslant (1 - \varepsilon) u_n + \varepsilon v_n.$$

On démontre de même, pour $i \in [\![1, N]\!]$:

$$\begin{aligned}
p^{(n+1)}_{i,j} &= p_{i,\ell} u_n + \sum_{k\in[\![1,N]\!]\setminus\{\ell\}} p_{i,k} p^{(n)}_{k,j} \geqslant p_{i,\ell} u_n + (1 - p_{i,\ell}) v_n\\
&\geqslant \varepsilon u_n + (1 - \varepsilon) v_n
\end{aligned}$$

et donc :

$$v_{n+1} \geqslant \varepsilon u_n + (1 - \varepsilon) v_n.$$

(b) Il résulte des inégalités démontrées dans la question précédente que $(u_n)_{n\in\mathbb{N}}$ décroît et $(v_n)_{n\in\mathbb{N}}$ croît. On a de plus

$$0 \leqslant u_{n+1} - v_{n+1} \leqslant (1-2\varepsilon)(u_n - v_n).$$

La suite $(u_n - v_n)$ est décroissante, car $\varepsilon > 0$, et minorée par 0. Sa limite L vérifie $0 \leqslant L \leqslant (1-2\varepsilon)L$ et donc $L = 0$. Les deux suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont donc adjacentes. Elles convergent vers la même limite.

(c) On note q_j la limite commune de $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$.

Soit $i \in [\![1, N]\!]$. Pour tout $n \in \mathbb{N}$, on a $u_n \leqslant p_{i,j}^{(n)} \leqslant v_n$. On en déduit que $\lim\limits_{n\to+\infty} p_{i,j}^{(n)} = q_j$.

On définit ainsi q_j pour tout $j \in [\![1, N]\!]$. Les coefficients q_j sont positifs car limite d'une suite à termes positifs. On a $PU = U$; on en déduit, pour tout $n \in \mathbb{N}$, $P^n U = U$, c'est-à-dire, pour tout $i \in [\![1, N]\!]$, $\sum\limits_{j=1}^{N} p_{i,j}^{(n)} = 1$. En faisant tendre n vers $+\infty$, on obtient :

$$\sum_{j=1}^{N} q_j = 1.$$

Si l'on note Z le vecteur ligne QP, on obtient, pour tout $i \in [\![1, N]\!]$:

$$z_j = \sum_{k=1}^{N} q_k p_{k,j} = \lim_{n\to+\infty} \sum_{k=1}^{N} p_{i,k}^{(n)} p_{k,j} = \lim_{n\to+\infty} p_{i,j}^{(n+1)} = q_j.$$

Ainsi, $QP = Q$.

Si R est une probabilité telle que $RP = R$, on a $RP^n = R$, pour tout $n \in \mathbb{N}$. La matrice P^n tend quand n tend vers $+\infty$ vers une matrice M dont les coefficients de la colonne d'indice j sont tous égaux à q_j. L'égalité $RM = R$ donne $r_j = \sum\limits_{i=1}^{N} r_i q_j = q_j$, donc $R = Q$.

16.22 1. (a) On a $\mathbb{P}(X < 0) = \sum\limits_{n\in\mathbb{N}^*} \mathbb{P}(X = -n) = \sum\limits_{n\in\mathbb{N}^*} \mathbb{P}(X = n) = \mathbb{P}(X > 0)$. On en déduit :

$$\mathbb{P}(X > 0) = \frac{1 - \mathbb{P}(X = 0)}{2} \leqslant \frac{1}{2} \text{ et } \mathbb{P}(X \geqslant 0) = \frac{1 + \mathbb{P}(X = 0)}{2} \geqslant \frac{1}{2}.$$

(b) Soit X et Y deux variables aléatoires indépendantes de même loi. La variable $X - Y$ est aussi à valeurs dans $\mathbb{Z}$ et, pour $n \in \mathbb{Z}$:

$$\begin{aligned}\mathbb{P}(X - Y = n) &= \sum_{\substack{(i,j)\in\mathbb{Z}^2 \\ i-j=n}} \mathbb{P}(X = i, Y = j) = \sum_{j\in\mathbb{Z}} \mathbb{P}(X = j + n, Y = j) \\ &= \sum_{j\in\mathbb{Z}} \mathbb{P}(X = j + n)\,\mathbb{P}(Y = j),\end{aligned}$$

par indépendance de X et Y. On en déduit, en faisant un changement d'indice, puis en utilisant le fait que X et Y ont même loi :

$$\begin{aligned}\mathbb{P}(X-Y=-n) &= \sum_{j\in\mathbb{Z}} \mathbb{P}(X=j-n)\,\mathbb{P}(Y=j)\\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}(X=i)\,\mathbb{P}(Y=i+n)\\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}(Y=i)\,\mathbb{P}(X=i+n).\end{aligned}$$

On a bien $\mathbb{P}(X-Y=-n)=\mathbb{P}(X-Y=n)$ pour tout $n\in\mathbb{Z}$.

(c) Soit X et Y deux variables aléatoires symétriques indépendantes. On a, pour tout $n\in\mathbb{Z}$:

$$\begin{aligned}\mathbb{P}(X+Y=n) &= \sum_{\substack{(i,j)\in\mathbb{Z}^2\\ i+j=n}} \mathbb{P}(X=i)\,\mathbb{P}(Y=j)\\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}(X=i)\,\mathbb{P}(Y=n-i).\end{aligned}$$

On en déduit, en utilisant le fait que X et Y sont symétriques, puis un changement d'indice :

$$\begin{aligned}\mathbb{P}(X+Y=-n) &= \sum_{i\in\mathbb{Z}} \mathbb{P}(X=i)\,\mathbb{P}(Y=-n-i)\\ &= \sum_{i\in\mathbb{Z}} \mathbb{P}(X=-i)\,\mathbb{P}(Y=n+i)\\ &= \sum_{j\in\mathbb{Z}} \mathbb{P}(X=j)\,\mathbb{P}(Y=n-j)\\ &= \mathbb{P}(X+Y=n).\end{aligned}$$

.

2. (a) Cela se déduit de la question 1.(b) par récurrence sur k.

(b) Soit $k\in[\![1,n]\!]$.

- On a $\{S_n-S_k\geqslant 0\}\cap\Omega_k\subset\{S_n\geqslant S_k\}\cap\{S_k>x\}\subset\{S_n>x\}$ et donc :
$$\{S_n-S_k\geqslant 0\}\cap\Omega_k\subset\{S_n>x\}\cap\Omega_k.$$
- On a :
$$\{S_n-S_k\geqslant 0\}\cap\Omega_k=\{X_{k+1}+\cdots X_n\geqslant 0\}\cap\Omega_k.$$
On peut écrire Ω_k sous la forme $\{f(X_1,\ldots,X_k)\in A_k\}$, où f est une application de $\mathbb{R}^k$ dans $\mathbb{R}$ et A_k une partie de $\mathbb{Z}$. Les variables X_1, X_2, …, X_n étant indépendantes, il en est de même de $X_{k+1}+\cdots X_n$ et de $f(X_1,\ldots,X_n)$. On en déduit l'indépendance des événements $\{S_n-S_k\geqslant 0\}$ et Ω_k. On a donc :
$$\mathbb{P}\left(\{S_n-S_k\geqslant 0\}\cap\Omega_k\right)=\mathbb{P}(S_n-S_k\geqslant 0)\,\mathbb{P}\left(\Omega_k\right).$$
La variable S_n-S_k est symétrique, comme somme de variables symétriques indépendantes, donc $\mathbb{P}(S_n-S_k\geqslant 0)\geqslant \frac{1}{2}\cdot$ D'où l'inégalité voulue.

(c) L'événement $\left\{\max\limits_{1\leqslant j\leqslant n} S_j > x\right\}$ est réalisé si, et seulement s'il existe $j \in [\![1,n]\!]$, pour lequel $\{S_j > x\}$ est réalisé. En considérant k, le plus petit des tels entiers j, cela signifie que l'événement Ω_k est réalisé. On a donc :

$$\bigcup_{k=1}^{n} \Omega_k = \left\{\max_{1\leqslant j\leqslant n} S_j > x\right\}.$$

(d) Les événements Ω_k, pour $1 \leqslant k \leqslant n$, sont incompatibles donc :

$$\mathbb{P}\left(\max_{1\leqslant j\leqslant n} S_j > x\right) = \sum_{k=1}^{n} \mathbb{P}(\Omega_k).$$

En utilisant la question (b), on obtient :

$$\mathbb{P}(\Omega_k) \leqslant 2\mathbb{P}\left(\{S_n - S_k \geqslant 0\} \cap \Omega_k\right) \leqslant 2\mathbb{P}\left(\{S_n > x\} \cap \Omega_k\right).$$

On en déduit :

$$\mathbb{P}\left(\max_{1\leqslant j\leqslant n} S_j > x\right) \leqslant 2\sum_{k=1}^{n} \mathbb{P}\left(\{S_n > x\} \cap \Omega_k\right).$$

Les événements $\{S_n > x\} \cap \Omega_k$, pour $1 \leqslant k \leqslant n$, sont incompatibles donc :

$$\begin{aligned}\sum_{k=1}^{n} \mathbb{P}\left(\{S_n > x\} \cap \Omega_k\right) &= \mathbb{P}\left(\{S_n > x\} \cap \bigcup_{k=1}^{n} \Omega_k\right)\\ &= \mathbb{P}\left(\{S_n > x\} \cap \left\{\max_{1\leqslant j\leqslant n} S_j > x\right\}\right) = \mathbb{P}\left(S_n > x\right).\end{aligned}$$

D'où le résultat.

16.23 1. Si $i = n$, on a $S_n - S_i = 0$ et cette variable aléatoire est indépendante de toute autre variable aléatoire. Supposons donc $i < n$. On a :

$$B_i = \{|S_1| < a\} \cap \ldots \cap \{|S_{i-1}| < a\} \cap \{|S_i| \geqslant a\},$$

donc $S_i \mathbb{1}_{B_i}$ est une fonction des variables aléatoires $X_1, \ldots, X_i$. D'autre part, on a $S_n - S_i = \sum\limits_{k=i+1}^{n} X_i$, donc $S_n - S_i$ est une fonction des variables aléatoires X_{i+1}, $\ldots$, X_n. Les variables aléatoires $X_1, \ldots, X_n$ étant indépendantes, il en est de même de $S_i \mathbb{1}_{B_i}$ et $S_n - S_i$, d'après la proposition 26 de la page 907.

Les variables aléatoires X_i sont dans $L_2(\Omega, \mathcal{A}, \mathbb{P})$, donc il en est de même des variables aléatoires S_i. Les variables aléatoires $S_n^2 \mathbb{1}_{B_i}$, $S_i^2 \mathbb{1}_{B_i}$ et $(S_n - S_i)^2 \mathbb{1}_{B_i}$ ont une espérance finie, car, elles sont positives et inférieures respectivement à S_n^2, S_i^2 et $(S_n - S_i)^2$, qui ont une espérance finie. On a, par linéarité de l'espérance :

$$\begin{aligned}\mathbb{E}\left(S_n^2 \mathbb{1}_{B_i}\right) - \mathbb{E}\left(S_i^2 \mathbb{1}_{B_i}\right) &= \mathbb{E}\left((S_n - S_i)(S_n + S_i) \mathbb{1}_{B_i}\right)\\ &= \mathbb{E}\left((S_n - S_i)^2 \mathbb{1}_{B_i}\right) + 2\mathbb{E}\left((S_n - S_i) S_i \mathbb{1}_{B_i}\right).\end{aligned}$$

Les variables aléatoires $S_i \mathbb{1}_{B_i}$ et $S_n - S_i$ sont indépendantes donc :

$$\mathbb{E}\left((S_n - S_i) S_i \mathbb{1}_{B_i}\right) = \mathbb{E}\left(S_i \mathbb{1}_{B_i}\right) \mathbb{E}(S_n - S_i) = 0,$$

car $\mathbb{E}(S_n - S_i) = \sum_{k=i+1}^{n} \mathbb{E}(X_k) = 0$. On a donc l'égalité voulue.

Soit $\omega \in \Omega$.

- Si $\omega \in B_i$, alors $S_i^2(\omega)\,\mathbb{1}_{B_i}(\omega) = S_i^2(\omega) \geqslant a^2 = a^2\mathbb{1}_{B_i}(\omega)$.
- Si $\omega \notin B_i$, alors $S_i^2(\omega)\,\mathbb{1}_{B_i}(\omega) = a^2\mathbb{1}_{B_i}(\omega) = 0$.

On a donc $S_i^2\mathbb{1}_{B_i} \geqslant a^2\mathbb{1}_{B_i}$, d'où l'on déduit :

$$\mathbb{E}\left(S_i^2\mathbb{1}_{B_i}\right) \geqslant a^2\mathbb{E}\left(\mathbb{1}_{B_i}\right) \geqslant a^2\mathbb{P}(B_i).$$

Comme $\mathbb{E}\left((S_n - S_i)^2\,\mathbb{1}_{B_i}\right) \geqslant 0$, car c'est l'espérance d'une variable positive, on a *a fortiori* $\mathbb{E}\left(S_n^2\,\mathbb{1}_{B_i}\right) \geqslant a^2\mathbb{P}(B_i)$.

2. (a) Soit $\omega \in \Omega$. Alors $\omega \in C$ si, et seulement s'il existe $k \in [\![1,n]\!]$ tel que $\omega \in \{|S_k| \geqslant a\}$. Alors i est le plus petit tel entier k si, et seulement si, $\omega \in B_i$. On en déduit que $C = B_1 \cup B_2 \ldots \cup B_n$. Les événements B_1, B_2, ..., B_n sont incompatibles, donc :

$$\mathbb{P}(C) = \sum_{i=1}^{n} \mathbb{P}(B_i).$$

(b) De la question 1, on déduit :

$$\mathbb{P}(C) \leqslant \frac{1}{a^2}\sum_{i=1}^{n} \mathbb{E}\left(S_n^2\,\mathbb{1}_{B_i}\right) \leqslant \frac{1}{a^2}\mathbb{E}\left(S_n^2 \sum_{i=1}^{n} \mathbb{1}_{B_i}\right).$$

Mais $\sum_{i=1}^{n} \mathbb{1}_{B_i} = \mathbb{1}_C \leqslant 1$, donc $\mathbb{E}\left(S_n^2 \sum_{i=1}^{n} \mathbb{1}_{B_i}\right) \leqslant \mathbb{E}\left(S_n^2\right) = \mathbb{V}(S_n)$, car $\mathbb{E}(S_n) = 0$. On obtient :

$$\mathbb{P}\left(\sup(|S_1|, |S_2|, \ldots, |S_n|) \geqslant a\right) = \mathbb{P}(C) \leqslant \frac{\mathbb{V}(S_n)}{a^2}.$$

16.24 1. On a, pour tout $k \in \mathbb{N}$, $|p_k - q_k| \leqslant p_k + q_k$. Les séries de termes généraux p_k et q_k convergent, donc la série de terme général $|p_k - q_k|$ converge également.

(a) Pour toute partie A de $\mathbb{N}$:

$$\begin{aligned}\big|\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)\big| &= \big|1 - \mathbb{P}(X \in \bar{A}) - 1 + \mathbb{P}(Y \in \bar{A})\big| \\ &= \big|\mathbb{P}(X \in \bar{A}) - \mathbb{P}(X \in \bar{A})\big|.\end{aligned}$$

On a donc :

$$\begin{aligned}2\big|\mathbb{P}\left(X \in A\right) - \mathbb{P}\left(Y \in A\right)\big| &= \big|\mathbb{P}(X \in A) - \mathbb{P}(X \in A)\big| \\ &\quad + \big|\mathbb{P}(X \in \bar{A}) - \mathbb{P}(X \in \bar{A})\big| \\ &= \left|\sum_{k \in A} p_k - \sum_{k \in A} q_k\right| + \left|\sum_{k \in \bar{A}} p_k - \sum_{k \in \bar{A}} q_k\right| \\ &\leqslant \sum_{k \in A} |p_k - q_k| + \sum_{k \in \bar{A}} |p_k - q_k| \\ &\leqslant 2\mathrm{d}(X, Y).\end{aligned}$$

(b) Pour tous réels a et b, $2\min(a,b) = a+b-|a-b|$. On a donc :

$$2\sum_{k\in\mathbb{N}}\min(p_k,q_k) = \sum_{k\in\mathbb{N}}p_k + \sum_{k\in\mathbb{N}}q_k - \sum_{k\in\mathbb{N}}|p_k-q_k| = 2-2\mathrm{d}(X,Y)$$

et donc :

$$\mathrm{d}(X,Y) = 1-\sum_{k\in\mathbb{N}}\min(p_k,q_k).$$

(c) L'événement contraire de $\{X\neq Y\}$ est :

$$\{X=Y\} = \bigcup_{k\in\mathbb{N}}\{X=k\}\cap\{Y=k\}.$$

Pour tout $k\in\mathbb{N}$, l'événement $\{X=k\}\cap\{Y=k\}$ est inclus dans $\{X=k\}$ et dans $\{Y=k\}$. Par croissance de $\mathbb{P}$, on obtient :

$$\mathbb{P}\left(\{X=k\}\cap\{Y=k\}\right) \leqslant \min(p_k,q_k).$$

On en déduit :

$$\mathbb{P}\left(X=Y\right) = \sum_{k\in\mathbb{N}}\mathbb{P}\left(\{X=k\}\cap\{Y=k\}\right) \leqslant \sum_{k\in\mathbb{N}}\min(p_k,q_k).$$

On a donc :

$$\mathrm{d}(X,Y) = 1-\sum_{k\in\mathbb{N}}\min(p_k,q_k) \leqslant 1-\mathbb{P}\left(X=Y\right) \leqslant \mathbb{P}\big(X\neq Y\big).$$

2. (a) On montre facilement que f est croissante sur $[0,1]$.
On a donc $f([0,1]) = [f(0),f(1)] = [0,1]$.

(b) Pour tout $i\in[\![1,n]\!]$, la variable X_i est à valeur dans $[0,1]$. On a par définition :

$$\mathbb{P}(X_i=0) = \mathbb{P}\big(\{Y_i=0\}\cap\{U_i=0\}\big)$$

et par indépendance de U_i et Y_i :

$$\mathbb{P}(X_i=0) = \exp\left(-\frac{\lambda}{n}\right)\left(1-f\left(\frac{\lambda}{n}\right)\right) = \left(1-\frac{\lambda}{n}\right).$$

On en déduit que $\mathbb{P}(X_i=1) = \dfrac{\lambda}{n}$: la variable X_i suit la loi de Bernoulli de paramètre $\dfrac{\lambda}{n}\cdot$

Les variables Y_1, U_1, Y_2, U_2,..., Y_n, U_n sont indépendantes. Pour tout $i\in[\![1,n]\!]$, X_i est une fonction de Y_i et U_i. D'après le lemme des coalitions, les variables $X_1,\ldots,X_n$ sont indépendantes. On a donc une somme de variables de Bernoulli indépendantes de même paramètre. On en déduit que $\displaystyle\sum_{i=1}^{n}X_i$ suit la loi binomiale de paramètre $\left(n,\dfrac{\lambda}{n}\right)$.

La variable aléatoire Y, somme de variables mutuellement indépendantes suivant la loi de Poisson de paramètre $\dfrac{\lambda}{n}$, suit la loi de Poisson de paramètre λ (voir l'exercice 14 de la page 908).

(c) L'événement $\{X_i = Y_i\}$ est réalisé si $X_i = 0$ et $Y_i = 0$ ou $X_i = 1$ et $Y_i = 1$. On remarque que si $X_i = 0$ alors $Y_i = 0$ et si $Y_i = 1$ alors $X_i = 1$. On a donc :

$$\mathbb{P}(X_i = Y_i) = \mathbb{P}(X_i = 0) + \mathbb{P}(Y_i = 1) = 1 - \frac{\lambda}{n} + \exp\left(-\frac{\lambda}{n}\right)\frac{\lambda}{n}\cdot$$

On en déduit que :

$$\mathbb{P}(X_i \neq Y_i) = 1 - \mathbb{P}(X_i = Y_i) = \frac{\lambda}{n}\left(1 - \exp\left(-\frac{\lambda}{n}\right)\right).$$

Une simple étude de fonction montre que $1 - e^{-t} \leqslant t$ pour tout $t \geqslant 0$. On obtient donc, pour tout $i \in [\![1, n]\!]$:

$$\mathbb{P}(X_i \neq Y_i) \leqslant \frac{\lambda^2}{n^2}\cdot$$

(d) On a montré dans la question 1 que $\mathrm{d}(X, Y) \leqslant \mathbb{P}(X \neq Y)$. Pour obtenir l'inégalité voulue, il suffit donc de démontrer :

$$\mathbb{P}(X \neq Y) \leqslant \frac{\lambda^2}{n}\cdot$$

On a $\displaystyle\bigcap_{i=1}^{n}\{X_i = Y_i\} \subset \left\{\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i\right\}$. On en déduit

$$\{X \neq Y\} = \overline{\left\{\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i\right\}} \subset \overline{\bigcap_{i=1}^{n}\{X_i = Y_i\}} = \bigcup_{i=1}^{n}\{X_i \neq Y_i\}.$$

Par croissance de $\mathbb{P}$, puis inégalité de Boole, on en déduit :

$$\begin{aligned}\mathbb{P}(X \neq Y) &\leqslant \mathbb{P}\left(\bigcup_{i=1}^{n}\{X_i \neq Y_i\}\right) \leqslant \sum_{i=1}^{n}\mathbb{P}(X_i \neq Y_i)\\ &\leqslant \sum_{i=1}^{n}\frac{\lambda^2}{n^2} \leqslant \frac{\lambda^2}{n}\cdot\end{aligned}$$

On a donc :

$$\mathrm{d}(X, Y) \leqslant \frac{\lambda^2}{n}\cdot$$

16.25 1. On remarque qu'une suite $(x_n)_{n\in\mathbb{N}}$ de réels converge vers x si, et seulement si :

$$\forall k \in \mathbb{N}^* \quad \exists n \in \mathbb{N} \quad \forall p \geqslant n \quad |x_p - x| \leqslant \frac{1}{k},$$

car pour tout $\varepsilon > 0$, il existe $k \in \mathbb{N}^*$ tel que $\dfrac{1}{k} \leqslant \varepsilon$.

On a donc, pour tout $\omega \in \Omega$,

$$\begin{aligned}\omega \in B &\iff \forall k \in \mathbb{N}^* \quad \exists n \in \mathbb{N} \quad \forall p \geqslant n \quad |X_p(\omega) - X(\omega)| \leqslant \frac{1}{k}\\ &\iff \forall k \in \mathbb{N}^* \quad \exists n \in \mathbb{N} \quad \forall p \geqslant n \quad \omega \in \left\{|X_p - X| \leqslant \frac{1}{k}\right\}\\ &\iff \forall k \in \mathbb{N}^* \quad \omega \in C_k\\ &\iff \omega \in \bigcap_{k\in\mathbb{N}^*} C_k.\end{aligned}$$

On a donc $B = \bigcap\limits_{k\in\mathbb{N}^*} C_k$. La suite $(C_k)_{k\in\mathbb{N}^*}$ étant décroissante, on en déduit :

$$\mathbb{P}(B) = \lim_{k\to+\infty} \mathbb{P}(C_k).$$

2. On obtient, en considérant l'événement contraire :

$$\forall \varepsilon > 0 \quad \mathbb{P}\Big(\bigcup_{n\in\mathbb{N}} \bigcap_{p\geqslant n} \big\{|X_p - X| \leqslant \varepsilon\big\}\Big) = 1.$$

On prenant $\varepsilon = \dfrac{1}{k}$, on obtient $\mathbb{P}(C_k) = 1$. En faisant tendre k vers l'infini, on en déduit $\mathbb{P}(B) = 1$: la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X.

3. Pour $\varepsilon > 0$ et $n \in \mathbb{N}$, on a, d'après l'inégalité de Boole :

$$\mathbb{P}\Big(\bigcup_{p\geqslant n} \big\{|X_p - X| > \varepsilon\big\}\Big) \leqslant \sum_{p=n}^{+\infty} \mathbb{P}(|X_p - X| > \varepsilon).$$

Le reste d'ordre $n-1$ de la série de terme général $\mathbb{P}\big(|X_n - X| > \varepsilon\big)$ tend vers 0 quand n tend vers $+\infty$ donc *a fortiori* :

$$\lim_{n\to+\infty} \mathbb{P}\Big(\bigcup_{p\geqslant n} \big\{|X_p - X| \leqslant \varepsilon\big\}\Big) = 0.$$

La suite $\Big(\bigcup\limits_{p\geqslant n} \big\{|X_p - X| \leqslant \varepsilon\big\}\Big)_{n\in\mathbb{N}}$ est décroissante donc, par continuité décroissante :

$$\mathbb{P}\Big(\bigcap_{n\in\mathbb{N}} \bigcup_{p\geqslant n} \big\{|X_p - X| > \varepsilon\big\}\Big) = \lim_{n\to+\infty} \mathbb{P}\Big(\bigcup_{p\geqslant n} \big\{|X_p - X| \leqslant \varepsilon\big\}\Big) = 0.$$

De la question précédente, on déduit que la suite $(X_n)_{n\in\mathbb{N}}$ converge presque sûrement vers X.

16.26 1. (a) Soit $t \in [a, b]$.

- $*$ Si $u \geqslant 0$, on a :

$$|u|^k e^{tu} \leqslant |u|^k e^{bu} = |u|^k e^{-u\delta} e^{(b+\delta)u}.$$

La fonction $u \mapsto |u|^k e^{-u\delta}$ est continue sur $\mathbb{R}_+$ et a pour limite 0 en $+\infty$, donc elle est bornée sur $\mathbb{R}_+$. Il existe $C_1 > 0$ tel que l'on ait $|u|^k e^{tu} \leqslant C_1 e^{(b+\delta)u}$ pour tout $u \geqslant 0$.

* Si $u \leqslant 0$, on a :

$$|u|^k e^{tu} \leqslant |u|^k e^{au} = |u|^k e^{u\delta} e^{(a-\delta)u}.$$

La fonction $u \mapsto |u|^k e^{u\delta}$ est continue sur $\mathbb{R}_-$ et a pour limite 0 en $-\infty$, donc elle est bornée sur $\mathbb{R}_-$. Il existe $C_2 > 0$ tel que l'on ait $|u^k| e^{tu} \leqslant C_2 e^{(a-\delta)u}$, pour tout $u \leqslant 0$.

En prenant $C = \max(C_1, C_2)$, on a le résultat voulu.

- On obtient, pour tout $t \in [a, b]$:

$$|X|^k e^{tX} \leqslant C\big(e^{(a-\delta)X} + e^{(b+\delta)X}\big).$$

Comme $a - \delta$ et $b + \delta$ sont dans $]\alpha, \beta[$, les variables $e^{(a-\delta)X}$ et $e^{(b+\delta)X}$ et donc $C\big(e^{(a-\delta)X} + e^{(b+\delta)X}\big)$ ont une espérance finie. On en déduit que $X^k e^{tX}$ est d'espérance finie. Cela est vrai pour tout $[a, b] \subset]\alpha, \beta[$, donc $X^k e^{tX}$ est d'espérance finie pour tout $t \in]\alpha, \beta[$.

(b) Supposons $X(\Omega)$ dénombrable et notons $X(\Omega) = \{x_n \,;\, n \in \mathbb{N}\}$. Pour tout $t \in]\alpha, \beta[$ et tout $k \in \mathbb{N}$, on pose :

$$f_k(t) = \mathbb{E}\left(X^k e^{tX}\right) = \sum_{n=0}^{+\infty} x_n^k e^{tx_n} \mathbb{P}(X = x_n)$$

et pour tout $n \in \mathbb{N}$, $u_{n,k}(t) = x_n^k e^{tx_n} \mathbb{P}(X = x_n)$.

Soit $k \in \mathbb{N}$ et $[a, b] \subset]\alpha, \beta[$. Il résulte de la question précédente qu'il existe un réel $C > 0$ tel que, pour $n \in \mathbb{N}$ et $t \in [a, b]$:

$$|x_n^k| e^{tx} \mathbb{P}(X = x_n) \leqslant C\big(e^{(a-\delta)x_n} \mathbb{P}(X = x_n) + e^{(b+\delta)x_n} \mathbb{P}(X = x_n)\big).$$

On en déduit que, pour tout $k \in \mathbb{N}$, la série $\sum\limits_n u_{n,k}(t)$ converge uniformément par rapport à t sur $[a, b]$.

D'autre part, on a pour tout $t \in]\alpha, \beta[$, $u'_{n,k}(t) = u_{n,k+1}(t)$.

Du théorème de dérivation des séries de fonctions (théorème 24 de la page 513), on déduit que la fonction f_k est dérivable sur $[a, b]$ de dérivée f_{k+1}. Comme cela est vrai pour tout $[a, b] \subset]\alpha, \beta[$, on conclut que f_k est dérivable sur $]\alpha, \beta[$, de dérivée f_{k+1}.

Comme $f_0 = L_X$, on en déduit que L_X est de classe $\mathcal{C}^\infty$ sur $]\alpha, \beta[$ et que, pour tout $k \in \mathbb{N}$, on a $L_X^{(k)} = f_k$.

En prenant $t = 0$, on en déduit que $\mathbb{E}(X^k) = L_X^{(k)}(0)$.

Si $X(\Omega)$ est fini, les résultats précédents restent vrais et sont évidents, car on dérive des sommes finies.

2. (a) Pour tout $t \in]\alpha, \beta[$, on a $L_X(t) > 0$, car c'est l'espérance d'une variable aléatoire strictement positive, donc $\Psi(t)$ est défini. On a :

$$\forall t \in]\alpha, \beta[\quad \Psi'(t) = \frac{L'_X(t)}{L_X(t)} \quad \text{et} \quad \Psi''(t) = \frac{L''_X(t) L_X(t) - \left(L'_X(t)\right)^2}{\left(L_X(t)\right)^2}.$$

On a $L''_X(t) L_X(t) - \left(L'_X(t)\right)^2 = \mathbb{E}\left(X^2 e^{tX}\right) \mathbb{E}\left(e^{tX}\right) - \left(\mathbb{E}\left(X e^{tX}\right)\right)^2$.

On écrit $e^{tX} = \left(e^{tX/2}\right)^2$ et $X^2e^{tX} = \left(Xe^{tX/2}\right)^2$, ce qui montre que $e^{tX/2}$ et $Xe^{tX/2}$ ont un moment d'ordre 2. D'après l'inégalité de Cauchy-Schwarz, on a :

$$\left(\mathbb{E}\left(Xe^{tX}\right)\right)^2 = \left(\mathbb{E}\left(Xe^{tX/2}e^{tX/2}\right)\right)^2 \leqslant \mathbb{E}\left(X^2e^{tX}\right)\mathbb{E}\left(e^{tX}\right).$$

On en déduit que, pour tout $t \in]\alpha,\beta[$, on a $\Psi''(t) \geqslant 0$.
De plus, s'il existe $t \in]\alpha,\beta[$ tel que $\Psi''(t) = 0$, alors X est presque sûrement constante (cas d'égalité dans l'inégalité de Cauchy-Schwarz : voir l'exercice 23 de la page 916), ce qui est contraire à l'hypothèse. Donc $\Psi'' > 0$ et Ψ est strictement convexe.

(b) On pose $h_c(t) = ct - \Psi(t)$. La fonction h_c' s'annule en $\gamma = (\Psi')^{-1}(c)$ et comme Ψ' croît, h_c' est positive sur $]\alpha,\gamma[$ et négative sur $]\gamma,\beta[$. On a donc $g(c) = h_c(\gamma)$.
On note que $m = \Psi'(0)$. Donc $m \in I$ et $g(m) = 0$, car :

$$\Psi(0) = \ln L_X(0) = \ln 1 = 0.$$

Si $c \neq m$, alors on a $\gamma \neq 0$ et $g(c) > \Psi(0) = 0$.

(c)
- Si $c < m$, c'est-à-dire $c < \Psi'(0)$, alors $\gamma < 0$. Le maximum de h_c sur $]\alpha,0[$, atteint en γ, est aussi le maximum de h_c sur $]\alpha,\beta[$, c'est-à-dire $g(c)$.
- De même, si $c > m$, alors $\gamma > 0$ et le maximum de h_c sur $]0,\beta[$ est le maximum de h_c sur $]\alpha,\beta[$, c'est-à-dire $g(c)$.

(d)
- Supposons $c < m$. Considérons $t \in]\alpha,0[$. On a, d'après l'inégalité de Markov :

$$\mathbb{P}(X \leqslant c) = \mathbb{P}(e^{tX} \geqslant e^{ct}) \leqslant \frac{\mathbb{E}\left(e^{tX}\right)}{e^{ct}} = e^{\Psi(t)-ct}.$$

En prenant $t = \gamma$, on obtient l'inégalité voulue.
- Si $c > m$, on obtient l'inégalité de la même manière, en prenant $t \in]0,\beta[$.

3. (a) On remarque que, pour tout $t \in]\alpha,\beta[$, on a $e^{tS_n} = \prod_{k=1}^{n} e^{tX_k}$. Les variables e^{tX_k} ont une espérance finie $L_X(t)$ et sont indépendantes, donc e^{tS_n} possède une espérance et :

$$\mathbb{E}\left(e^{tS_n}\right) = \prod_{k=1}^{n}\mathbb{E}\left(e^{tX_k}\right) = \left(L_X(t)\right)^n = e^{n\Psi_X(t)}.$$

On applique ce qui précède à la variable aléatoire S_n. On a, pour $t \in]\alpha,\beta[$, $\Psi_{S_n}(t) = n\Psi_X(t)$. Le maximum de $\Psi_{S_n}(t) - cnt$ est donc $ng(c)$ et $\mathbb{E}(S_n) = nm$.On a donc :

$$\mathbb{P}\left(\frac{S_n}{n} \leqslant c\right) = \mathbb{P}(S_n \leqslant nc) \leqslant e^{-ng(c)} \qquad \text{si} \qquad c < m$$
$$\mathbb{P}\left(\frac{S_n}{n} \geqslant c\right) = \mathbb{P}(S_n \geqslant nc \leqslant e^{-ng(c)} \qquad \text{si} \qquad c > m.$$

(b) On choisit $\varepsilon > 0$ assez petit pour que $m - \varepsilon$ et $m + \varepsilon$ soient dans I. On a alors :

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = \mathbb{P}\left(\frac{S_n}{n} \leqslant m - \varepsilon\right) + \mathbb{P}\left(\frac{S_n}{n} \geqslant m + \varepsilon\right).$$

En appliquant la question précédente, on en déduit :

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant e^{-ng(m-\varepsilon)} + e^{-ng(m+\varepsilon)} \leqslant 2e^{-n\min(g(m+\varepsilon),g(m-\varepsilon))}.$$

(c) Pour tout $\varepsilon > 0$ assez petit la série de terme général $\mathbb{P}\left(\left|\frac{S_n}{n} - m\right|\right)$ est donc convergente. D'après le résultat de l'exercice 16.25, on en déduit que la suite $\left(\frac{S_n}{n}\right)$ converge presque sûrement vers m.

16.27 1. (a) Soit $(t,u) \in [0,1]^2$.

- Si $|t-u| \leqslant \eta$, alors $|f(t) - f(u)| \leqslant \varepsilon$.
- Si $|t-u| > \eta$, on a $\dfrac{(t-u)^2}{\eta^2} > 1$ et donc :

$$|f(t) - f(u)| \leqslant 2\|f\|_\infty \leqslant \frac{2\|f\|_\infty (t-u)^2}{\eta^2}.$$

Comme ces majorants sont positifs, on a *a fortiori*, dans les deux cas :

$$|f(t) - f(u)| \leqslant \frac{2\|f\|_\infty (t-u)^2}{\eta^2} + \varepsilon.$$

(b) On en déduit que la variable aléatoire Y_n vérifie l'inégalité :

$$|f(Y_n) - f(x)| \leqslant \frac{2\|f\|_\infty (Y_n - x)^2}{\eta^2} + \varepsilon.$$

La variable aléatoire Y_n est finie. Il en est donc de même de $f(Y_n)$ et $(Y_n - x)^2$, qui possèdent donc une espérance. Par croissance, puis par linéarité de l'espérance, on obtient :

$$\begin{aligned}
|\mathbb{E}(f(Y_n) - f(x))| &\leqslant \mathbb{E}(|f(Y_n) - f(x)|) \\
&\leqslant \mathbb{E}\left(\frac{2\|f\|_\infty (Y_n - x)^2}{\eta^2} + \varepsilon\right) \\
&\leqslant \frac{2\|f\|_\infty}{\eta^2}\mathbb{E}\left((Y_n - x)^2\right) + \varepsilon.
\end{aligned}$$

On pose $S_n = X_1 + \cdots + X_n$. La variable S_n, somme de variables aléatoires indépendantes de même paramètre x suit la loi binomiale de paramètre (n,x).

On a donc $\mathbb{E}(Y_n) = \dfrac{1}{n}\mathbb{E}(S_n) = \dfrac{nx}{n} = x$.

On en déduit que $\mathbb{E}\left((Y_n - x)^2\right) = \mathbb{V}(Y_n)$. D'où l'inégalité :

$$|\mathbb{E}(f(Y_n)) - f(x)| \leqslant \frac{2\|f\|_\infty \mathbb{V}(Y_n)}{\eta^2} + \varepsilon.$$

2. On part de l'inégalité précédente.

   On a $f(Y_n) = f\left(\dfrac{S_n}{n}\right)$. Par la formule de transfert, appliquée à la variable aléatoire S_n qui suit une loi binomiale de paramètre (n, x), on obtient :

$$\begin{aligned}\mathbb{E}\left(f(Y_n)\right) = \mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right) &= \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\mathbb{P}(S_n = k)\\ &= \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k} = B_n(f)(x).\end{aligned}$$

   D'autre part, on a :

$$\mathbb{V}(Y_n) = \frac{1}{n^2}\mathbb{V}(S_n) = \frac{nx(1-x)}{n^2} = \frac{x(1-x)}{n}\cdot$$

   On a donc, pour $n \in \mathbb{N}^*$ et $x \in [0,1]$,

$$|B_n(f)(x) - f(x)| \leqslant \frac{2\|f\|_\infty x(1-x)}{n\eta^2} + \varepsilon \leqslant \frac{2\|f\|_\infty}{n\eta^2} + \varepsilon.$$

   Quand n tend vers $+\infty$, $\dfrac{2\|f\|_\infty}{n\eta^2}$ tend vers 0. Donc il existe $n_0 \in \mathbb{N}^*$ tel que $\dfrac{2\|f\|_\infty}{n\eta^2} \leqslant \varepsilon$, pour $n \geqslant n_0$. On obtient donc, pour $n \geqslant n_0$:

$$\forall x \in [0,1] \quad |B_n(f)(x) - f(x)| \leqslant 2\varepsilon.$$

   Comme ε est un réel strictement positif quelconque, on en déduit que la suite de fonctions polynomiales $(B_n(f))$ converge uniformément vers f sur $[0,1]$. Cela démontre le théorème de Weierstrass.

16.28 Cette expérience aléatoire peut être modélisée par une suite de variables aléatoires, indépendantes de même loi, correspondant aux choix successifs du guichetier, chaque variable prenant la valeur 1 avec la probabilité $\dfrac{1}{2}$ et chacune des autres valeurs avec la probabilité $\dfrac{1}{2(n-1)}\cdot$

1. À chaque instant le client en première position peut être choisi avec une probabilité $\dfrac{1}{2}\cdot$
   Dans le cas contraire, il reste en première position. L'ensemble des valeurs prises par T_1 est donc $\mathbb{N}^*$ et pour $k \in \mathbb{N}^*$, on a $\mathbb{P}(T_1 = k) = \left(\dfrac{1}{2}\right)^{k-1}\dfrac{1}{2}\cdot$

   En effet, cela signifie que $k-1$ fois, le guichetier ne l'a pas choisi, avant de le choisir la k-ième fois.

   Ainsi T_1 suit la loi géométrique de paramètre $\dfrac{1}{2}$ et l'on a :

$$\mathbb{E}(T_1) = \mathbb{V}(T_1) = 2.$$

2. Pour un client quelconque de la file d'attente, la probabilité d'être servi à un moment donné est supérieure ou égale à $\dfrac{1}{2(n-1)}$ et donc la probabilité de n'être pas servi est inférieure ou égal à $\left(1-\dfrac{1}{2(n-1)}\right)$. On en déduit que, pour tout $\ell \in \mathbb{N}^*$, on a $\mathbb{P}(T_k=\ell) \leqslant \left(1-\dfrac{1}{2(n-1)}\right)^{\ell-1}$, car il faut pour que $\{T_k=\ell\}$ soit réalisé que le client initialement à la k-ième place n'ait pas été choisi $\ell-1$ fois. Ainsi, $\mathbb{P}(T_k=\ell)$ est majorée par une suite géométrique convergente. On en déduit que la série de terme général $\ell\mathbb{P}(T_k=\ell)$ converge. Ainsi T_k est d'espérance finie.
3. Soit $k \geqslant 2$. On note X_1 la variable représentant le premier choix du guichetier. Après ce choix, le client qui est à la k-ième place avance d'une place si $1 \leqslant X_1 \leqslant k-1$, ne bouge pas si $X_1 \geqslant k+1$ et quitte la file si $X_1 = k$. On a donc, pour $\ell \geqslant 2$:
$$\begin{aligned}
&\mathbb{P}(T_k=\ell \mid X_1 \leqslant k-1) = \mathbb{P}(T_{k-1}=\ell-1),\\
&\mathbb{P}(T_k=\ell \mid X_1 \geqslant k+1) = \mathbb{P}(T_k=\ell-1)\\
&\mathbb{P}(T_k=\ell \mid X_1 = k) = 0.
\end{aligned}$$
Sachant que :
$$\mathbb{P}(1 \leqslant X_1 \leqslant k-1) = \frac{1}{2}+\frac{k-2}{2(n-1)} \qquad \text{et} \qquad \mathbb{P}(X_1 \geqslant k+1) = \frac{n-k}{2(n-k)},$$
on obtient, en appliquant la formule des probabilités totales :
$$\mathbb{P}(T_k=\ell) = \left(\frac{1}{2}+\frac{k-2}{2(n-1)}\right)\mathbb{P}(T_{k-1}=\ell-1) + \frac{n-k}{2(n-1)}\mathbb{P}(T_k=\ell-1).$$
Comme $\mathbb{P}(T_k=1) = \dfrac{1}{2(n-1)}$, on a :
$$\begin{aligned}
\mathbb{E}(T_k) &= \frac{1}{2(n-1)} + \left(\frac{1}{2}+\frac{k-2}{2(n-1)}\right)\sum_{\ell=2}^{+\infty} \ell\,\mathbb{P}(T_{k-1}=\ell-1)\\
&\qquad + \frac{n-k}{2(n-1)}\sum_{\ell=2}^{+\infty} \ell\,\mathbb{P}(T_k=\ell-1)\\
&= \frac{1}{2(n-1)} + \left(\frac{1}{2}+\frac{k-2}{2(n-1)}\right)(\mathbb{E}(T_{k-1})+1)\\
&\qquad + \frac{n-k}{2(n-1)}(\mathbb{E}(T_k)+1).
\end{aligned}$$
On en déduit :
$$(n-2+k)\mathbb{E}(T_k) = (2n-2)+(n-3+k)\mathbb{E}(T_{k-1}).$$
La suite $\big((n-2+k)\mathbb{E}(T_k)\big)$ est arithmétique de raison $2n-2$ et de premier terme $(n-1)\mathbb{E}(T_1) = 2n-2$.

On obtient, pour $1 \leqslant k \leqslant n$, $(n-2+k)\mathbb{E}(T_k) = (2n-2)k$ et :

$$\mathbb{E}(T_k) = \frac{(2n-2)k}{n-2+k}.$$

4. Dans une file d'attente classique, le temps d'attente du k-ième client serait k. On observe que, pour $k \in [\![1,n]\!]$, on a $\mathbb{E}(T_k) \geqslant k$, l'inégalité étant stricte pour $k < n$.

16.29 1. • Si X est presque sûrement constante, alors l'événement $\{X = x\}$ est de probabilité 0 ou 1, pour tout réel x, donc est indépendant de l'événement $\{Y = y\}$, pour tout y réel. Les variables X et Y sont indépendantes. Donc (i) implique (ii). On remarque qu'il suffit qu'une des deux variables aléatoires soit presque sûrement constante pour que (ii) soit vérifié.
La réciproque est fausse. Si (X, Y) est un couple de variables aléatoires suivant la loi uniforme sur $\{0,1\}^2$, les variables X et Y suivent une loi de Bernoulli de paramètre $\frac{1}{2}$ et sont indépendantes, sans être presque sûrement constantes.
• L'implication $(ii) \Rightarrow (iii)$ est un théorème du cours.
La réciproque est fausse. Si X est une variable réelle suivant la loi uniforme sur $\{-1,0,1\}$ et $Y = |X|$, on a $\mathbb{E}(X) = 0$ et $XY = X$, donc $\mathbb{E}(XY) = 0 = \mathbb{E}(X)\,\mathbb{E}(Y)$.
Mais X et Y ne sont pas indépendantes, car :

$$\mathbb{P}\big(\{X=0\} \cap \{Y=1\}\big) = 0 \quad \text{et} \quad \mathbb{P}(X=0)\,\mathbb{P}(Y=1) = \frac{1}{3} \times \frac{2}{3} = \frac{2}{9}.$$

2. On suppose que (iii) est vérifié et qu'il existe $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ strictement croissantes ainsi qu'une variable aléatoire réelle Z telle que $X = f(Z)$ et $Y = g(Z)$.
Comme f et g sont croissantes, $f(a) - f(b)$ et $g(b) - g(a)$ ont le signe de $a - b$, donc $\theta(a,b) \geqslant 0$, pour tout $(a,b) \in \mathbb{R}^2$. On en déduit que la variable $\theta(Z, Z')$ est à valeurs positives.
On écrit :

$$\theta(Z,Z') = f(Z)g(Z) + f(Z')g(Z') - f(Z)g(Z') - f(Z')g(Z).$$

La variable aléatoire $f(Z)g(Z) = XY$ est d'espérance finie. La variable aléatoire $f(Z')g(Z')$ a même loi que $f(Z)g(Z)$ et donc même espérance.
On a $f(Z) = X$ et $g(Z')$ a même loi que Y. Les variables $f(Z)$ et $g(Z')$ sont indépendantes car Z et Z' le sont, et d'espérance finie. On en déduit que $f(Z)g(Z')$ est d'espérance finie et

$$\mathbb{E}(f(Z)g(Z')) = \mathbb{E}(X)\,\mathbb{E}(Y).$$

On démontre de même que $\mathbb{E}(f(Z')g(Z)) = \mathbb{E}(X)\,\mathbb{E}(Y)$.
Ainsi $\theta(Z,Z')$ est une somme de variables aléatoires d'espérance finie. Elle est donc d'espérance finie et, par linéarité de l'espérance :

$$\mathbb{E}\big(\theta(Z,Z')\big) = 2\mathbb{E}(XY) - 2\mathbb{E}(X)\mathbb{E}(Y) = 0.$$

Comme $\theta(Z,Z')$ est une variable aléatoire positive, on en déduit qu'elle est presque sûrement nulle. Il existe un événement A de probabilité égale à 1 tel que l'on ait $\theta\big(Z(\omega),Z'(\omega)\big)=0$, pour tout $\omega\in A$. Si $\omega\in A$, on a $f(Z(\omega))=f(Z'(\omega))$ ou $g(Z(\omega))=g(Z'(\omega))$ et comme f et g sont strictement croissantes, on a $Z(\omega)=Z'(\omega)$.

Si a et b sont deux réels distincts, l'événement $\{Z=a\}\cap\{Z'=b\}$ est inclus dans $\{Z\neq Z'\}$ et donc dans $\Omega\setminus A$. Il est donc de probabilité nulle. Les variables aléatoires Z et Z' sont indépendantes et Z' a même loi que Z donc :

$$\mathbb{P}(Z=a)\,\mathbb{P}(Z=b)=\mathbb{P}(Z=a)\,\mathbb{P}(Z'=b)=\mathbb{P}\big(\{Z=a\}\cap\{Z'=b\}\big)=0.$$

Soit a tel que $\mathbb{P}(Z=a)\neq 0$ (un tel a existe car $\sum\limits_{a\in Z(\Omega)}\mathbb{P}(Z=a)=1$). On a alors $\mathbb{P}(Z=b)=0$ pour tout $b\neq a$ et donc $\mathbb{P}(Z=a)=1$. La variable Z est donc presque sûrement constante. On en déduit qu'il en est de même des variables aléatoires $X=f(Z)$ et $Y=g(Z)$.

16.30 1. Si f est bornée, il en est de même de la variable aléatoire $f(X_1)$. En notant K un majorant de $|f(X_1)|$, on a $|X_2f(X_1)|\leqslant K|X_2|$, et comme $|X_2|$ est d'espérance finie, il en est de même de$X_2\,f(X_1)$. Le théorème de transfert appliquée à la variable (X_1,X_2) à valeurs dans $\mathbb{Z}^2$ et à la fonction $(x,y)\mapsto yf(x)$ donne :

$$\mathbb{E}(X_2\,f(X_1))=\sum_{(m,n)\in\mathbf{Z}^2} nf(m)\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right).$$

La famille $\left(nf(m)\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right)\right)_{(m,n)\in\mathbf{Z}^2}$ est sommable donc on a :

$$\begin{aligned}\mathbb{E}(X_2f(X))&=\sum_{m\in\mathbf{Z}}\sum_{n\in\mathbf{Z}}nf(m)\mathbb{P}\left(\{X_1=n\}\cap\{X_2=n\}\right)\\&=\sum_{m\in\mathbf{Z}}f(m)\sum_{n\in\mathbf{Z}}n\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right).\end{aligned}$$

Si $\mathbb{P}(X_1=m)\neq 0$, alors on peut écrire

$$\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right)=\mathbb{P}(X_1=m)\,\mathbb{P}(X_2=n\mid X_1=m),$$

et on a alors :

$$\sum_{n\in\mathbf{Z}}n\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right)=\mathbb{P}(X_1=m)\sum_{n\in\mathbf{Z}}n\mathbb{P}\left(X_2=n\mid X_1=m\right).$$

La dernière somme est l'espérance de X_2 pour la probabilité conditionnelle $\mathbb{P}_{\{X_1=m\}}$, que l'on note $\mathbb{E}(X_2\mid X_1=m)$.

Par ailleurs $\sum\limits_{n\in\mathbf{Z}}n\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right)=0$ si $\mathbb{P}(X_1=m)=0$ car $\mathbb{P}\left(\{X_1=m\}\cap\{X_2=n\}\right)=0$ pour tout $n\in\mathbb{Z}$. En posant :

$$\forall m\in\mathbb{Z}\quad h(m)=\begin{cases}\mathbb{E}(X_2\mid X_1=m) & \text{si}\quad \mathbb{P}(X_1=m)\neq 0\\ 0 & \text{sinon,}\end{cases}$$

on obtient :

$$\mathbb{E}(X_2\,f(X_1))=\sum_{m\in\mathbf{Z}}\mathbb{P}(X_1=m)f(m)h(m).$$

D'après le théorème de transfert, on a :

$$\mathbb{E}(X_2 f(X_1)) = \mathbb{E}\big(h(X_1)f(X_1)\big).$$

Cela est valable pour toute fonction bornée f. En prenant pour f la fonction constante égale à 1, on obtient que $h(X_1)$ possède une espérance égale à $\mathbb{E}(X_2)$. La variable aléatoire $Y_1 = h(X_1)$ a les propriétés voulues.

2. Soit $h'(X_1)$ une autre variable aléatoire ayant les propriétés requises. Considérons $m \in \mathbb{Z}$ tel que $\mathbb{P}(X_1 = m) \neq 0$. En prenant $f = \mathbb{1}_m$, on obtient $f(X_1) = \mathbb{1}_{\{X_1=m\}}$ et l'égalité $\mathbb{E}(X_2 \mathbb{1}_{\{X_1=m\}}) = \mathbb{E}\big(h'(X_1)\mathbb{1}_{\{X_1=m\}}\big)$. Le théorème de transfert donne :

$$\begin{aligned}
\mathbb{E}(X_2\, \mathbb{1}_{\{X_1=m\}}) &= \sum_{(k,n)\in\mathbf{Z}^2} \mathbb{1}_m(k)\, n\, \mathbb{P}\left(\{X_1 = k\} \cap \{X_2 = n\}\right) \\
&= \sum_{n\in\mathbf{Z}} n\mathbb{P}\left(\{X_1 = m\} \cap \{X_2 = n\}\right) \\
&= \mathbb{P}(X_1 = m) \sum_{n\in\mathbf{Z}} n\, \mathbb{P}(X_2 = n \mid X_1 = m) \\
&= \mathbb{P}(X_1 = m)\mathbb{E}(X_2 \mid X_1 = m)
\end{aligned}$$

et $\mathbb{E}\big(h'(X_1)\mathbb{1}_{\{X_1=m\}}\big) = h'(m)\mathbb{P}(X_1 = m)$. On obtient :

$$h'(m) = \mathbb{E}(X_2 \mid X_1 = m) = h(m).$$

Posons $C = \{m \in \mathbb{Z} \mid \mathbb{P}(X_1 = m) = 0\}$ et $A = \bigcup_{m\in C} \{X_1 = m\}$. Alors A est un événement négligeable, car réunion au plus dénombrable d'ensembles négligeables. Les variables aléatoires $h(X_1)$ et $h'(X_1)$ coïncident sur le complémentaire de A, donc sont presque sûrement égales.

3. Par définition de l'espérance conditionnelle, il existe une fonction k définie sur $\mathbb{Z}$ telle que :

$$\mathbb{E}(X_2\, g(X_1) \mid X_1) = k(X_1).$$

D'après les deux premières questions, on a $k(m) = \mathbb{E}(X_2 g(X_1) \mid X_1 = m)$ pour tout $m \in \mathbb{Z}$ tel que $\mathbb{P}(X_1 = m) \neq 0$. Le théorème de transfert, appliquée à (X_1, X_2) et à la fonction $(x, y) \mapsto yg(x)$ donne :

$$\begin{aligned}
\mathbb{E}(X_2\, g(X_1) \mid X_1 = m) &= \sum_{(k,n)\in\mathbf{Z}^2} ng(k)\mathbb{P}\left(\{X_1 = k\} \cap \{X_2 = n\} \mid X_1 = m\right) \\
&= \sum_{n\in\mathbf{Z}} ng(m)\mathbb{P}\left(\{X_1 = m\} \cap \{X_2 = n\} \mid X_1 = m\right) \\
&= g(m) \sum_{n\in\mathbf{Z}} n\mathbb{P}\left(X_2 = n \mid X_1 = m\right) \\
&= g(m)\mathbb{E}(X_2 \mid X_1 = m) = g(m)h(m).
\end{aligned}$$

On a donc $\mathbb{E}(X_2\, g(X_1) \mid X_1) = g(X_1)h(X_1) = g(X_1)\mathbb{E}(X_2 \mid X_1)$ sur le complémentaire de A.